Are you slacking? Where do you and your country stand in the happiness pursuit?

Badunenko, Oleg and Cordero, Jose M. and Kumbhakar, Subal C.

16 June 2021

Are you slacking? Where do you and your country stand in the happiness pursuit?

Oleg Badunenko[*] José M. Cordero[**] Subal C. Kumbhakar[⋆]

May 26, 2021

Abstract

This paper contributes to the literature on subjective well-being of individuals by estimating a "happiness frontier" under the premise that individuals seek to maximize happiness. The happiness frontier is defined as the pinnacle of happiness, given the resources and personal characteristics of individuals. However, individuals may not only strive to maximize their happiness but also to achieve higher levels of income, ceteris paribus. Thus, there is a trade-off between these two objectives, and the happiness-income frontier is attained when one can neither increase her happiness nor income, given her resources and personal characteristics. If one fails to attain the happiness-income frontier, there is a shortfall. The shortfall measures the rate at which both happiness and income for an individual, who has failed to attain the happiness frontier, could be increased given everything else. We also explore the determinants of this shortfall and conduct an empirical analysis using micro-level data gathered from the World Values Survey (WVS) for 74 countries. Our results suggest that the age as well as certain personal circumstances, such as being unemployed or having a partner, have a strong influence on the levels the estimated well-being and income attainment measures. Likewise, we also find that the quality of the government of the country of residence also greatly affects individuals' capacity to convert their resources into higher levels of happiness and income.

JEL: C13, I31, H53

Keywords: Happiness (well-being), Income, Frontier, Happiness shortfall, Cross-country analysis

[*] Oleg Badunenko, Department of Economics and Finance, College of Business, Arts and Social Sciences, Brunel University London, UK, E-mail: oleg.badunenko@brunel.ac.uk.

[**] José M. Cordero Universidad de Extremadura, Badajoz, Spain, E-mail: jmcordero@unex.es.

[⋆] Subal C. Kumbhakar, Department of Economics, State University of New York, Binghamton, NY 13902-6000, USA and Stavanger Business School, Norway, E-mail: kkar@binghamton.edu.

1 Introduction

In the last few decades there has been a growing interest within social sciences in the study of happiness based on the approach of subjective well-being and, more specifically, in analyzing the factors that influence whether an individual defines himself as happy or satisfied with life.[1] This interest has spread mainly among economists, causing the birth of the so-called "economics of happiness", which has become one of the most relevant fields of research within this discipline resulting in a significant increase in the number of applied studies being published in the most prestigious economic journals (Veenhoven, 2012; MacKerron, 2012). The methodological approaches used in those studies differ in details, although most are, in the end, based on using an equation where the dependent variable is a measure of the level of subjective well-being and different econometric techniques are used to identify explanatory variables associated with this indicator (Ferrer-i Carbonell and Frijters, 2004; Di Tella and MacCulloch, 2006; Powdthavee, 2010).

Studies agree that subjective well-being is explained mainly by individual characteristics like age, health, employment status, education, gender or marital status (Bonini, 2008; Helliwell and Huang, 2008).[2] These determinants are similar across different societies and levels of development (Helliwell and Barrington-Leigh, 2010). Nevertheless, there is clear evidence that measures of well-being vary widely across countries as highlighted by the World Happiness Reports (e.g., Helliwell et al., 2020). It is therefore not surprising that there is also a vast literature devoted to analyzing other indicators at country or regional level that have more influence on determining individual life satisfaction (e.g., Bjørnskov et al., 2010; O'Connor, 2017).

Recently, a small body of empirical literature has emerged within this field of research with the aim of examining the efficiency with which individuals convert their resources into well-being (life satisfaction). This approach was originally developed by Rayo and Becker (2007) who argue that individuals are not primarily concerned with their absolute level of happiness, but rather with their relative situation with respect to other individuals. This argument is based on the fact that each person adopts a specific conceptual reference that affects her definition of happiness. This may affect the subjective appreciation of the same experiences and personal circumstances, that is, the ability to obtain happiness from similar attributes or resources, even if they have the same conditions or resources. Therefore, some individuals are intrinsically happier than others. Hence, there might be some slacks in the conversion process of resources into well-being, because some individuals might not be able to reach the maximum levels of satisfaction given a certain level of resources at their disposal. That is, there are slacks meaning that the achieved levels of satisfaction is less

[1] The literature on well-being is based on individuals' self-reported data about life satisfaction, happiness or subjective well-being. Although there are significant differences between these constructs, we will use the words happiness, satisfaction and (subjective) well-being indistinctly throughout the paper (see, for example, Veenhoven, 1991; Easterlin, 1995).

[2] For instance, Bonini (2008) attributes almost 80% of variation in well-being measures to these factors.

than their maximum possible level. Alternatively, they might need more resources than other individuals to attain certain levels of well-being.

To address this issue, we borrow a tool from the stochastic frontier approach that is commonly applied to examine production or cost efficiency. Here we estimate well-being efficiency (attainment) on the basis of the capability approach (Sen, 1985; Narayan et al., 2000). Therefore, the frontier is defined as the pinnacle of happiness, given the resources and characteristics of the individuals. We call this happiness frontier (HF). This frontier serves as a reference for assessing happiness (well-beings) of all individuals. Some may be on the frontier thereby attaining maximum happiness, ceteris paribus; and, thus, relative measures of happiness can be estimated as the distance to the HF – the way it was defined in Farrell (1957).

However, so far, only a few empirical studies have estimated happiness attainment (efficiency) measures and explore the potential causes of happiness shortfall. Note that by inefficiency we mean happiness shortfall from the maximum possible (happiness frontier), given resources and individual characteristics.[3] The majority of the applications have applied nonparametric techniques such as data envelopment analysis (DEA) (e.g., Debnath and Shankar, 2014; Carboni and Russu, 2015; Mizobuchi, 2017) or partial frontiers (Binder and Broeckel, 2012; Cordero et al., 2017; Nikolova and Popova, 2021). The main argument for using these methods is that they are very flexible, since they do not require assuming a specific functional form of the happiness (production) function. However, while using this approach it is not possible to statistically test (the way it is done in a standard regression) whether resources (inputs) are really contributing to life satisfaction or not. Moreover, they do not recognize whether deviations from the frontier are due to inefficiency or statistical noise attributable to measurement errors. Actually, when they attempt to identify the factors that might explain happiness shortfalls, it is common that they rely on regressions (e.g., two-stage approach), which convert these methods into semi-parametric models.

This paper uses an innovative methodological approach in order to provide evidence about shortfalls of well-being across individuals from different countries and the factors affecting the process of converting resources into well-being using longitudinal data. First, we rely on a parametric approach to specify the HF, which allows for identifying both shortfalls and noise in the estimation of the happiness frontier.[4] Second, we take into account that the outcome variable is discrete. Third, in our model specification we consider two outcome variables, represented by subjective well-being (life satisfaction) and incomes, since we consider that resources and personal characteristics contribute to not only well-being or happiness, but also income levels. This is an important novelty with respect to previous studies in this line of research, in which the

[3] We will be using the term frontier and inefficiency borrowed from the stochastic frontier production models and label them happiness frontier and happiness shortfall. Thus, the output is happiness or life satisfaction and inputs are resources and individual characteristics.

[4] In the context of a production theory the HF is the production frontier, resources are inputs, and shortfall is inefficiency. Thus, determinants of production inefficiency is translated to determinants of happiness shortfall.

level of income is normally considered as a determinant of subjective well-being despite the fact that the relationship between these two variables is far more complex (Easterlin, 1995; Oswald, 1997). We argue that individuals may not only strive to reach their maximum happiness given their resources and personal characteristics but also pursue to achieve higher level of income. Thus, there is a trade-off between these two objectives, and the frontier is attained when one can neither increase her happiness nor income, given her resources and personal characteristics. To accommodate two outcome variables we borrow the multi-output production function with inefficiency concept from Kumbhakar (2013). Moreover, since both selected output variables are defined in terms of ordered categorical indicators, we have transformed this function into an ordered probit model, which required us to perform some additional calculations to interpret the marginal effects of resources and determinants of the shortfall measure. To the best of our knowledge, this approach has hardly been used in applications with stochastic frontiers.[5]

Our empirical analysis uses information on a large sample of over 210,000 individuals from the World Value Survey (WVS) longitudinal database (Inglehart et al., 2014). This dataset has a pseudo-panel structure and it includes information about individuals from 74 different countries participating in different waves of the survey. This information has been combined with data at the country level on some key economic, social, and institutional factors identified in the previous literature as determinants of subjective well-being. By exploiting the variations from a large number of countries over the span of three decades, we can determine more precisely the potential factors behind the cross-country differences in well-being attainment.

In summary, this paper presents important methodological and empirical contributions to the analysis of subjective well-being. From a methodological perspective, we apply a stochastic frontier model adapted to the presence of two categorical output variables (wellbeing and income) using a nonlinear ordered probit model to explore the determinants of shortfall in both dimensions. This process is somewhat complex since the magnitude of the marginal effects of those potential determinants needs to be calculated differently for continuous and discrete variables. We provide a detailed explanation of the procedure employed to facilitate the interpretation of the results in this specific context. Likewise, from the empirical point of view, the possibility of having a large volume of information on individuals from countries all over the world allows us to explore the potential personal and national characteristics having the greatest impact on the shortfalls in well-being and income.

The remainder of this paper is structured as follows. Section 2 reviews the previous literature on determinants of happiness distinguishing between studies based on simple regression techniques and those that apply frontier techniques to estimate well-being efficiencies. Section 3 explains our estimation strategy and describes the methodology adopted to estimate attainment (efficiency) measures and explore the factors

[5] Griffiths et al. (2014) is an exception. In contrast to Griffiths et al. (2014), we derive a closed form for the likelihood function as well as inefficiency.

affecting the shortfalls. Section 4 presents the main characteristics of the dataset and the variables used in our empirical study. Section 5 reports and discusses the main results. Finally, the paper ends with some concluding remarks in Section 6.

2 Previous research on subjective well-being

The study of the determinants of subjective well-being has received increasing attention in the economic literature in recent decades. In this section we do not intend to offer a comprehensive review of those studies,[6] but merely offer some guidelines on the different methodological approaches that can be used to address this issue and briefly review the main findings regarding the variables identified as the main factors influencing well-being and also well-being attainment. Since our main interest lies in the methodological aspects, we have divided the studies into two blocks, depending on whether the analysis uses regressions or frontier models.

2.1 Studies focused on exploring the determinants of subjective well-being

Most empirical researches on the determinants of well-being use a simple additive function in which the self-reported level of subjective well-being, derived from the responses provided by individuals to questions about their current level of happiness or their satisfaction with their lives,[7] depend on a range of individual, economic, socio-demographic and institutional factors. Most empirical studies exploring those factors rely on conventional econometric methods such as ordinary least squares, probit or logit regression models. Some authors consider subjective well-being as a continuous variable, thus they apply OLS, whilst in most cases it is considered as a latent variable and, therefore, they rely on an ordered logit or probit model. Those can then be estimated by examining within-person deviations from means when only cross sectional data is available, although the use of panel data is becoming more frequent in recent analyses, since it allows controlling for time-invariant individual factors, such as the personality of individuals.

Most of the initial studies conducted within this line of research are from a single country (Ferrer-i Carbonell, 2005; Dittmann and Goebel, 2010). Later, with the development of international databases that provide information on the levels of life satisfaction or happiness of individuals, such as the World Values Survey (WVS), the European Social Survey (ESS) or the Gallup World Poll, it is increasingly common to find studies that adopt a cross-country approach (e.g., Schyns, 1998; Veenhoven, 2005; Exton et al., 2015).

---

6 Extensive reviews on the 'happiness economics' literature can be found in Bruni and Porta (2007); Frey (2008); Dolan et al. (2008) or, more recently, in Diener et al. (2018).

7 The validity of this measure was initially questioned, but recent evidence has proved their reliability (Krueger and Schkade, 2008; Diener et al., 2013), so it is common to find this variable in national and international surveys.

Studies using data from these sources frequently apply multilevel regression techniques or, at least, control country-fixed effects. Thus they can take into account the nested or hierarchical nature of the data, allowing for the joint inclusion of variables representative of the individual characteristics and data referring to the country (or region) where they live.

As mentioned in the introduction, there is a wide range of variables related to the personal characteristics of individuals. These variables present a strong and consistent correlation with measures of life satisfaction. Among them, the most relevant ones are the health status Van Praag and Ferrer-i Carbonell (2004); Blanchflower and Oswald (2011), the level of incomes (Easterlin, 1995; Clark et al., 2008)[8] and the education level completed (Argyle, 1999). Other socio-demographic variables that have proved to be significantly related to subjective well-being are: age (subjective well-being is higher among the young and the old, with a dip occurring around forties, Blanchflower and Oswald (2008), gender (in the sense that women are happier than men) (Gerdham and Johannesson, 2001; Salinas-Jiménez et al., 2013), employment (the negative influence of being unemployed is particularly relevant) (Winkelmann, 2009). Likewise, there is also evidence supporting that married people, or people living with a partner, are more satisfied with their lives (Diener et al., 2000; Stutzer and Frey, 2006) as well as those who are engaged in religious activities (Helliwell, 2003).

Regarding country-level indicators, although their influence might vary across different studies, the most relevant ones are those representing economic prosperity (e.g., GDP per capita or life expectancy at birth) (Peiro, 2006; Stutzer and Frey, 2012). Likewise, social support (having someone to depend on in times of bad times), freedom to make life choices and generosity and perceptions of corruption can also explain a large proportion of the variation in the national average levels of well-being (Helliwell et al., 2016). Other variables that also play a role are the levels of unemployment and inequality (Di Tella et al., 2003; Schneider, 2016), the proportion of expenditure in social services (Haller and Hadler, 2006), as well as different aspects related to institutional quality and welfare-state policies (Pacek and Radcliff, 2008; Ott, 2010, 2011).

2.2 Studies focused on measuring well-being attainment

In the most recent literature, we find several studies that apply frontier techniques originally developed for the analysis of production efficiency to estimate how efficiently individuals (or countries) transform the resources they have at their disposal into well-being. Assuming that it is possible to determine a measure of happiness efficiency, the use of this measure of individual efficiency will allow one to distinguish between individuals who are able to reach certain levels of well-being and others having difficulties to achieve those levels given the same level of resources and individual characteristics. Lovell et al. (1994) pioneered this

[8] The complex relationship between well-being and income is discussed in greater detail in Section 3.

approach to estimate the standard of living, the quality of life and the efficiency in transforming resources into achieved functioning. Since then, several empirical studies have adopted frontier techniques for similar purposes.

This line of research started with the work of Binder and Broeckel (2012), who assessed relative happiness efficiency for a sample of British individuals using a robust nonparametric approach known as order-m (Cazals et al., 2002). Subsequently, these authors explore the potential influence of a set of individual characteristics on the happiness efficiency score using a second-stage panel regression framework with fixed effects. Cordero et al. (2017) also adopted a nonparametric order-m approach to estimate happiness efficiency measures for individuals from 26 OECD countries, albeit using the conditional nonparametric approach proposed by Daraio and Simar (2005) to incorporate the influence of a set of individual- and country-level contextual factors into their models. Both studies conclude that traditional well-being determinants such as age, marital status, religion and unemployment also have influence on happiness attainment.

In addition, we find other studies using aggregated data at the country or regional level. Debnath and Shankar (2014) applied DEA to calculate relative happiness attainment measures for a sample of 130 countries considering several indicators of governance policies as input variables. They found that similar policies might affect happiness attainment across countries differently. Similarly, they provide evidence demonstrating that developed countries are more inefficient in terms of happiness than developing countries. Mizobuchi (2017) also uses DEA to estimate a happiness function for a sample of 36 countries considering several well-being dimensions and a set of socio-economic variables. One of the most relevant results is that the health factor explains the largest part of the cross-country variation in subjective well-being. Nikolova and Popova (2021) estimate well-being efficiencies for a sample of 91 countries over the 2009-2014 period using a robust nonparametric order-α approach (Aragon et al., 2005). They then apply second-stage panel data fixed effects regressions to examine the influence of several country-related institutional and social characteristics on happiness efficiency measures. Their results show that countries with high-quality institutions where citizens perceive that they have the freedom to choose their way of life are more efficient in terms of well-being. Carboni and Russu (2015) apply DEA to assess the performance of the 20 Italian regions considering different dimensions of well-being and Malmquist indices to examine their evolution from 2005 to 2011. Their results indicate that northern regions outperform southern regions in terms of well-being attainment, although none improved their well-being over the evaluated period.

In all the aforementioned studies, the variable used as output (well-being or happiness) is treated as a continuous in the analysis when in fact it is an ordered categorial indicator. Moreover, as noted in the introduction, combining the use of non-parametric techniques to estimate well-being attainment measures with regressions to explore the potential influence of different explanatory factors on those measures might entail some severe problems in the estimation, since they are not considering the potential influence of noise or random effects on estimates. The only exceptions to this second limitation are represented by the recent studies conducted by Cordero et al. (2021) and Mamatzakis and Tsionas (2021). In the former, the authors adopt a novel approach known as stochastic semi-nonparametric envelopment of data (StoNED) that accounts for both inefficiency and noise in their estimations (Kuosmanen and Kortelainen, 2012). However, in their application they use aggregate data at the country level for a single year, thus they cannot explore the potential influence of personal characteristics on individual well-being. In the latter, the authors apply a Markov Chain Monte Carlo (MCMC) procedure to estimate happiness attainment measures for a longitudinal sample of British individuals. They identify that a large part of the observed happiness levels of inefficiency can be explained by personality traits.

The present work constitutes the first empirical study estimating subjective well-being attainment measures at individual level using a stochastic frontier framework. In addition, the use of a multi-output approach, which allows us to consider the obtaining of income as an alternative to the attainment of well-being, is incorporated as a novelty as well as adopting a probit function to take into account that the two output variables considered are defined through ordered categorical indicators. By applying this approach to a cross-country database, we can estimate happiness attainment (efficiency) measures and explore the potential personal and national characteristics having the greatest impact on happiness shortfall.

3 Modeling and explaining shortfall in happiness and income

3.1 Happiness and income: A simultaneous relationship

Do people strive for attaining maximum levels of happiness and income, given the resources at their disposal? What is the relationship between them? This is important for formulating the model to examine their maxima and therefore the slacks in them. In this section we try to shed some light on this issue based on the evidence available in previous literature on subjective well-being, thus justifying the approach adopted in our study.

The study of the relationship between income and happiness or well-being has been one of the most discussed and debated topics in this strand of literature since the early 1970s (for an overview, see Senik, 2004). When this relationship is examined at a particular point in time and within the same region or country, many empirical studies conclude that individuals with higher income have, on average, higher levels of well-being (e.g., Diener et al., 1995; McBride, 2001). However, there are also many researchers claiming that, once basic needs are met, there is only a small effect of income on life satisfaction relative to other life circumstances such as unemployment, health or marital status (Diener and Biswas-Diener, 2002; Deaton and Stone, 2013).[9] This phenomenon constitutes what is popularly known in literature as the Easterlin paradox.

Several explanations have been proposed for what seems to be a contradiction.[10] First, subjective well-being does not only depend on income in absolute terms but in relative terms, i.e., individuals compare themselves with other individuals whom they take as their reference group. Therefore, an increase in the income of all individuals does not imply an automatic increase in their level of well-being; since the relative position of individuals in the income distribution is not changing, thus this increase does not affect their level of satisfaction with their lives (Easterlin, 1995). This reflects the importance of the relative position of individuals in society when analyzing both life satisfaction and income level. Second, the level of well-being depends on the subjective perception of whether income is enough to satisfy the existing needs, which will depend largely on the personal characteristics and the circumstances of the individuals. Third, it is often argued that individuals adapt to new situations by changing their expectations and it is common that those expectations are related to the level of income. Therefore, it is not surprising that people strive for high incomes, even if these only lead to a temporary or small increase in well-being (Ferrer-i Carbonell, 2005).

An alternative viewpoint to explain the relationship between the two variables is that there is a certain trade-off between income and well-being since individuals' level of well-being depends on their allocation of time between work and leisure (Haworth and Lewis, 2005). Considering that leisure can be broadly interpreted as time not occupied by paid work or personal chores and obligations (Harvey and Mukhopadhyay, 2007; Roberts, 1999), the possibility of achieving higher levels of well-being through leisure time will implicitly imply giving up work time and, therefore, losing potential earnings (Gratton and Taylor, 2004). Therefore, in principle, it is to be expected that if individuals behave rationally, they will work more hours as long as the benefits in terms of obtaining higher levels of income compensate for giving up their leisure time, so that when a certain level of income is reached, individuals will no longer be interested in working more. Nevertheless, this relationship is rather more complex, since upper income people do not tend to spend more of their day in enjoyable activities as compared to lower income people, but that upper income people tend to have lower levels of well-being because they suffer more stress (Kahneman and Deaton, 2010).

Another argument for not considering income as an explanatory factor for well-being is that there may be an inverse causal relationship between these two variables that might bias the estimates, since some aspects of a happy personality, such as optimism, may contribute to achieve higher levels of incomes (Diener

---

9 The findings of several empirical studies suggest that income explains only about 2–5 percent of the variance in well-being when they are conducted in developed countries (Hsieh, 2004; Ahuvia, 2008).

10 For an extensive review on different explanations to the 'Easterlin paradox', see Easterlin (1995) or Clark et al. (2008). A complete study on the relationship between income and subjective well-being can also be found in Stevenson and Wolfers (2008).

and Lucas, 1999). Therefore, it seems reasonable to rethink the inclusion of income level as an input in our HF framework.

In view of all the above, we propose a function with a multi-output structure. This approach relies on some previous evidence showing that several input variables and determinants of well-being can also affect individual incomes. For instance, Hartog and Oosterbeek (1998) suggest that the impact of education on well-being may be manifested indirectly through its effects on health and wealth. Similarly, unemployment reduces both income and well-being (Oswald, 1997). Therefore, when estimating relative measures of individuals' well-being attainment using stochastic frontiers, we consider income as an alternative output instead of being an input, thus differentiating us from previous studies using a HF approach.

3.2 Happiness-income frontier and their shortfalls

We evaluate the shortfall in happiness based on a presumption that there exist a correspondence between characteristics of a person, their income and happiness. To model this we use the notion of a multi-output production technology with inefficiency (Kumbhakar, 2013). In the terminology of this paper, we define the transformation function (subscripts are omitted to streamline the exposition),

$$A = f(\boldsymbol{X}, \theta \boldsymbol{Y}), \tag{1}$$

where $\boldsymbol{X}$ is the vector of personal characteristics and $\boldsymbol{Y} = [y_1 \; y_2]$ contains happiness (life satisfaction) (y_1) and income (y_2). A can include variables other than $\boldsymbol{X}$, if any, as well as random noise. Finally, $\theta \geq 1$ is a scalar and $(1/\theta) \leq 1$ represents shortfall of both happiness and income from their maximum values. The maximum attainable value of $\boldsymbol{Y} = \theta \boldsymbol{Y}$ and the shortfalls in both happiness and income are measured radially from $y_1/\max(y_1) = y_2/\max(y_2) = 1/\theta \leq 1$. The advantage of using a transformation function is that we do not need to make any behavioral assumption to estimate the model in (1). For identification, we use the normalizing assumption that the transformation function in (1) is homogeneous of degree 1 in $\boldsymbol{Y}$. This helps us to write (1) as

$$A/(\theta y_1) = f(\boldsymbol{X}, y_2/y_1). \tag{2}$$

Taking log of (2), we obtain

$$\ln A - \ln y_1 - \ln \theta = \ln [f(\boldsymbol{X}, y_2/y_1)]. \tag{3}$$

We write $\ln A = v$ and $\ln \theta = u \geq 0$. With these (3) becomes

$$\ln y_1 = -\ln\left[f(\boldsymbol{X}, y_2/y_1)\right] + v - u \equiv g(\boldsymbol{X}, y_2/y_1) + v - u, \tag{4}$$

which defines the happiness frontier when $u = 0$. This frontier is stochastic because of the presence of v. For a small value of u, we can interpret it as the percentage shortfall (when multiplied by 100) in happiness (y_1). Since $f(.)$ is homogeneous of degree one in $\boldsymbol{Y}$, the percentage shortfall of income (y_2) is also u. Put differently, u can be viewed as potential increase in both happiness and income, holding everything else unchanged. Conversely, it is the percentage shortfall of happiness and income from their maximum possible values.[11] Note that one can impose the linear homogeneity restriction in terms of y_2 which will give the income frontier with y_1 used as a numeraire. The choice of the numeraire does not affect either the estimates of the parameters or estimate of u (and its interpretation). Thus, one can interpret u as slack in the pursuit of happiness as well as income. Therefore, if $u = .10$ happiness (income) could be increased by 10%, which means happiness efficiency (happiness attained relative to the maximum possible happiness) is 0.90%.

Once an assumption about the functional form $f()$, and distributional assumptions on the error term v and the happiness shortfall term u are made, the parameters of $g(\boldsymbol{X}, y_2/y_1)$ in (4) can be estimated using the same tool that is used for estimating (in)efficiency in stochastic frontier models.

Since in our case the outcome variable (y_1) is an ordered variable, the standard production frontier approach needs to be modified. In the section below, we consider the case of a frontier model where the dependent variable is ordered.

3.3 Modeling discrete dependent variable

We make some amendments to the model in the previous section to allow for discrete outcome variables, viz., y_1. The model that we will be using is expressed as

$$y_i^* = \boldsymbol{x}_i'\boldsymbol{\beta} + v_i - u_i = \boldsymbol{x}_i'\boldsymbol{\beta} + \epsilon_i, \quad i = 1, \ldots, N, \tag{5}$$

where y^* is a latent continuous variable. In terms of the model in section 3.2, $g(\boldsymbol{X}, y_2/y_1) = \boldsymbol{x}_i'\boldsymbol{\beta}$ and $\ln y_{1i} = y_i^*$ so that we do not need to introduce new notations. What we observe is

$$y_i = m \quad \text{if} \quad \mu_{m-1} < y_i^* \leq \mu_m, \tag{6}$$

[11] This interpretation follows from the single as well as multi-output stochastic frontier production model (formally known as output distance function) in which inefficiency u is interpreted as the potential increase in all the outputs, ceteris paribus. Alternatively, $u \times 100$ is the percentage by which all the outputs could be increased if inefficiency is eliminated. Thus, this is also the output shortfall measure.

where $m = 1, \ldots, M$, $\mu_0 = -\infty$ and $\mu_M = \infty$. The model in (5) is an ordered probit model.

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

$$\begin{aligned} p_{im} &= \texttt{Prob}(y_i = m) \\ &= \texttt{Prob}(\mu_{m-1} < \boldsymbol{x}_i'\boldsymbol{\beta} + v_i - u_i \leq \mu_m) \\ &= \texttt{Prob}(v_i \leq \mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} - u_i) + \texttt{Prob}(v_i > \mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} - u_i) \\ &= \texttt{Prob}(v_i \leq \mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} - u_i) - \texttt{Prob}(v_i \leq \mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} - u_i) \\ &= \Phi(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} - u_i) - \Phi(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} - u_i), \end{aligned} \tag{7}$$

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

$$f(y_i, \boldsymbol{\mu}, \boldsymbol{x}_i'\boldsymbol{\beta} - u_i) = \sum_{m=1}^{M} I_m(y_i) p_{im}, \tag{8}$$

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

$$I_m(y_i) = \begin{cases} 1 & \text{if} \quad y_i = m \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

Given that $u \geq 0$, 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 defined both at individual and country-level. The determinants of shortfall in happiness and income are modeled by making variance of u heteroskedastic.[12]

Thus, the pdf of u_i is

$$f(u_i) = \frac{2}{\sqrt{2\pi}\sigma_{u_i}} \exp\left[-\frac{u_i^2}{2\sigma_{u_i}^2}\right], \tag{10}$$

12 Another way of modeling effects of determinants is to assume u to normally distributed with a non-zero mean truncated at zero (truncated normal) and make the mean and or the variance a function of determinants. This is more general and also very complicated to estimate.

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

3.4 Log-likelihood function

To derive the log-likelihood in a closed-form, we expand (7) for the ordered *probit* case

$$p_{im} = \texttt{Prob}(y_i = m) = \begin{cases} \texttt{Prob}(v_i \leq \mu_1 - \boldsymbol{x}_i'\boldsymbol{\beta} - u_i) & \text{if } y_i = 1 \\ \left[\begin{array}{l} \texttt{Prob}(v_i \leq \mu_m - \boldsymbol{x}_i'\boldsymbol{\beta}u_i) \\ -\texttt{Prob}(v_i \leq \mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} - u_i) \end{array} \right] & \text{if } 1 < y_i = m < M \\ 1 - \texttt{Prob}(v_i \leq \mu_{M-1} - \boldsymbol{x}_i'\boldsymbol{\beta} - u_i) & \text{if } y_i = M \end{cases}$$

$$= \begin{cases} \Phi(\mu_1 - \boldsymbol{x}_i'\boldsymbol{\beta} - u_i) & \text{if } y_i = 1 \\ \left[\begin{array}{l} \Phi(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} - u_i) \\ -\Phi(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} - u_i) \end{array} \right] & \text{if } 1 < y_i = m < M \\ 1 - \Phi(\mu_{M-1} - \boldsymbol{x}_i'\boldsymbol{\beta} - u_i) & \text{if } y_i = M. \end{cases} \tag{11}$$

The probability density function of the convolution of two error components, because of their independence, is

$$f(u_i, \epsilon_i) = \frac{2}{\sigma_u}\phi\left(\frac{u}{\sigma_u}\right) \times \sum_{m=1}^{M} I_m(y_i)p_{im}, \tag{12}$$

where $\phi(.)$ is the standard normal probability density function. To find the density of ϵ_i, $f(\epsilon_i)$, we need to integrate u_i out, i.e.,

$$f(\epsilon_i) = \int_0^\infty f(u_i, \epsilon_i)du_i \tag{13}$$

and to do that we need an integral of the type

$$\int_0^\infty \frac{2}{\sigma_u}\phi\left(\frac{u}{\sigma_u}\right)\Phi(\mu^* - \boldsymbol{x}_i'\boldsymbol{\beta} - u_i)du_i. \tag{14}$$

There is a closed form expression of the above integral. That is,

$$f(\epsilon_i) = \Phi\left(\frac{\mu^* - \boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}\right) + 2T\left(\frac{\mu^* - \boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}, -\sigma_u\right), \tag{15}$$

where T is the Owen's T function (Owen, 1956). Thus the log-likelihood is given by

$$ll_i\left(\boldsymbol{\beta},\boldsymbol{\mu},\boldsymbol{\gamma}\right)=\ln\left\{\sum_{m=1}^{M}I_m(y_i)\mathcal{L}_m\right\}, \tag{16}$$

where

$$\mathcal{L}_m=\begin{cases}\Phi\left(\frac{\mu_1-\boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}\right)+2T\left(\frac{\mu_1-\boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}},-\sigma_u\right) & \text{if } y_i=1\\ \left[\begin{array}{l}\Phi\left(\frac{\mu_m-\boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}\right)+2T\left(\frac{\mu_m-\boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}},-\sigma_u\right)\\ -\Phi\left(\frac{\mu_{m-1}-\boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}\right)-2T\left(\frac{\mu_{m-1}-\boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}},-\sigma_u\right)\end{array}\right] & \text{if } 1<y_i=m<M\\ \Phi\left(-\frac{\mu_{M-1}-\boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}\right)+2T\left(-\frac{\mu_{M-1}-\boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}},\sigma_u\right) & \text{if } y_i=M.\end{cases} \tag{17}$$

To maximize the above log-likelihood function we need its gradient. Fortunately, we can derive closed form expressions of these gradients. The derivatives of the log-likelihood function with respect to the parameters are

$$\frac{\partial ll_i}{\partial\boldsymbol{\beta}}=-\boldsymbol{x}_i\frac{1}{\mathcal{L}_m}\mathcal{D}_{1m}, \tag{18}$$

where

$$\mathcal{D}_{1m}=\begin{cases}\frac{2}{\sqrt{1+\sigma_u^2}}\phi\left(\frac{\mu_1-\boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}\right)\Phi\left(\frac{-\sigma_u(\mu_1-\boldsymbol{x}_i'\boldsymbol{\beta})}{\sqrt{1+\sigma_u^2}}\right) & \text{if } y_i=1\\ \left[\begin{array}{l}\frac{2}{\sqrt{1+\sigma_u^2}}\phi\left(\frac{\mu_m-\boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}\right)\Phi\left(\frac{-\sigma_u(\mu_m-\boldsymbol{x}_i'\boldsymbol{\beta})}{\sqrt{1+\sigma_u^2}}\right)\\ -\frac{2}{\sqrt{1+\sigma_u^2}}\phi\left(\frac{\mu_{m-1}-\boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}\right)\Phi\left(\frac{-\sigma_u(\mu_{m-1}-\boldsymbol{x}_i'\boldsymbol{\beta})}{\sqrt{1+\sigma_u^2}}\right)\end{array}\right] & \text{if } 1<y_i=m<M\\ -\frac{2}{\sqrt{1+\sigma_u^2}}\phi\left(-\frac{\mu_{M-1}-\boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}\right)\Phi\left(\frac{-\sigma_u(\mu_{M-1}-\boldsymbol{x}_i'\boldsymbol{\beta})}{\sqrt{1+\sigma_u^2}}\right) & \text{if } y_i=M,\end{cases} \tag{19}$$

and

$$\frac{\partial ll_i}{\partial\mu_i}=-\frac{1}{\mathcal{L}_m}\mathcal{D}_{1m}. \tag{20}$$

Finally, given the specification of $\sigma_{u_i}^2=\exp\left(\boldsymbol{z}_i\boldsymbol{\gamma}\right)$,

$$\frac{\partial ll_i}{\partial\boldsymbol{\gamma}}=-\boldsymbol{z}_i\frac{1}{\mathcal{A}}\mathcal{D}_{2m}, \tag{21}$$

where

$$\mathcal{D}_{2m} = \begin{cases} \dfrac{\sigma^2_{u_i}}{1+\sigma^2_{u_i}} \dfrac{h_1 f(h_1)}{2} - \exp\left(-\dfrac{h_1^2\left(1+a^2\right)}{2}\right) \dfrac{1}{2\pi} \dfrac{a}{1+a^2} & \text{if } y_i = 1 \\ \left[\begin{array}{l} \dfrac{\sigma^2_{u_i}}{1+\sigma^2_{u_i}} \dfrac{h_m f(h_m)}{2} - \exp\left(-\dfrac{h_m^2\left(1+a^2\right)}{2}\right) \dfrac{1}{2\pi} \dfrac{a}{1+a^2} \\ -\dfrac{\sigma^2_{u_i}}{1+\sigma^2_{u_i}} \dfrac{h_{m-1} f(h_{m-1})}{2} + \exp\left(-\dfrac{h_{m-1}^2\left(1+a^2\right)}{2}\right) \dfrac{1}{2\pi} \dfrac{a}{1+a^2} \end{array} \right] & \text{if } 1 < y_i = m < M \\ \dfrac{\sigma^2_{u_i}}{1+\sigma^2_{u_i}} \dfrac{h_{M-1} f(h_{M-1})}{2} - \exp\left(-\dfrac{h_{M-1}^2\left(1+a^2\right)}{2}\right) \dfrac{1}{2\pi} \dfrac{a}{1+a^2} & \text{if } y_i = M, \end{cases} \tag{22}$$

$$h_i = \begin{cases} \dfrac{\mu_1 - \boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}} & \text{if } y_i = 1 \\ \dfrac{\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}} & \text{if } 1 < y_i = m < M \\ -\dfrac{\mu_{M-1} - \boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}} & \text{if } y_i = M, \end{cases} \tag{23}$$

$$a = \begin{cases} -\sigma_u & \text{if } y_i = 1 \\ -\sigma_u & \text{if } 1 < y_i = m < M \\ \sigma_u & \text{if } y_i = M, \end{cases} \tag{24}$$

and

$$f(h) = 2\phi\left(h\right)\Phi\left(-ha\right) \tag{25}$$

4 Data and Empirical results

4.1 Data and variables

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 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 five more waves, covering the period 1981 to 2014.[13] 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. After dropping the observations due to the presence of missing data in some relevant variables, our final sample consists of 210,218 observations from 74 different countries. Table A1 in the Appendix shows the number of observations available for each country and participation in different waves.

The WVS dataset provides information on the two main variables of our interest, subjective well-being or life satisfaction and income level. 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).[14] Responses are based on a scale from 1, which means 'completely dissatisfied', to 10, meaning 'completely satisfied'. The latter is represented by the declared relative position of the individuals in the income distribution of their country.[15] This variable is expressed in deciles, thus it is also possible to distinguish ten different income levels.

In addition, the WVS contains information on individual socioeconomic and demographic characteristics that we incorporate into our analysis. We have chosen two variables that fulfil the requirement of isotonicity or monotonicity, which have been also used in some previous studies. The first one is the level of education, which is grouped into eight different categories according to the total number of years of completed education, and the second one is an indicator of the health status perceived by the individuals in a five-level scale (very poor, poor, fair, good or very good).

As explanatory variables, we have selected several variables that previous literature identifies as the most common predictors of subjective well-being (e.g., Helliwell et al., 2016), 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 value 1), their employment status (unemployed takes value 1), marital status (married or living together as married takes value 1) and being religious. At country level, we have 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,

13 The second-, third-, fourth-, fifth- and sixth-wave data were collected from 1990 to 1994, from 1995 to 1998, from 1999 to 2004, from 2005 to 2009 and from 2010 to 2014, respectively. The seventh wave is planned to be conducted worldwide in the period 2017-2020.

14 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).

15 Considering relative income is more frequent in the literature than absolute income. See Clark et al. (2008) for details.

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

Variable	Type	Mean	SD	Min	Max
Life satisfaction/well-being	Output	6.5314	2.4617	1	10
Incomes	Output	4.6354	2.2845	1	10
Education level	Input	4.6981	2.2249	1	8
Health status	Input	3.8361	0.8786	1	5
Individual characteristics					
Age	Covariate	40.5124	15.971	16	99
Gender	Covariate	0.5085	0.4992	0	1
Unemployed	Covariate	0.0964	0.2951	0	1
Married	Covariate	0.6410	0.4796	0	1
Religious	Covariate	0.6837	0.4650	0	1
Country-level variables					
Social protection expenditure	Covariate	12.518	7.965	0.35	32.07
GDP pc	Covariate	14,678	17,428	272.39	91,617
Unemployment rate	Covariate	8.412	6.167	0.48	36
Gini index	Covariate	37.78	8.91	21	64.8
Corruption Perception Index	Covariate	46.60	22.71	12	95
Quality of governance (WGI)	Covariate	0.1857	0.9043	-1.752	1.841

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).[16]

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, on average, individuals included in the sample seem to be quite satisfied with their life (mean value of 6.53 out of 10). Figure 1 shows frequencies of observing individuals in happiness 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.

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

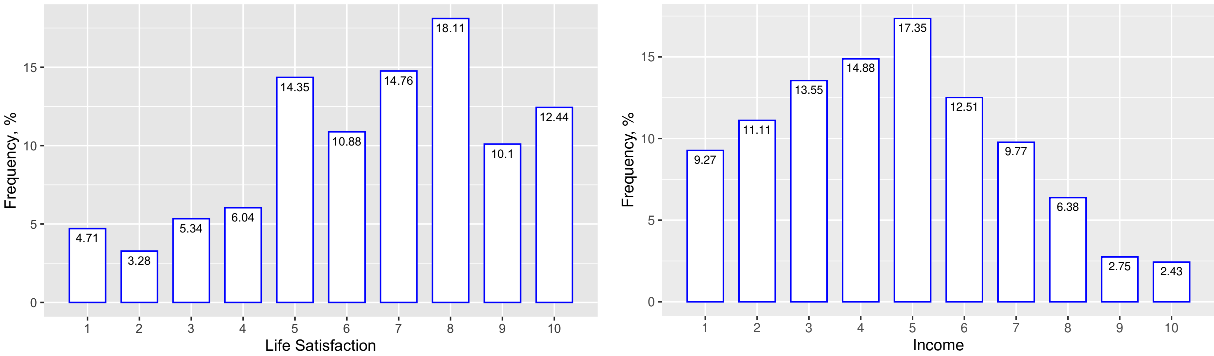

Figure 1: The Frequency of Life satisfaction/Happiness and Income

4.2 Regression results

Table 2 shows the estimated coefficients of the happiness-income frontier. All 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 effects of regressors is the change in probability of being in a group defined in a latent way by estimated parameters $\boldsymbol{\mu}$. Due to nonlinear nature of our model the coefficients are not marginal effects. In the next section we show the calculation that are required to interpret the coefficients presented in Table 2.

Table 2: SF-ordered probit regression estimates. Dependent variable is Happiness

Variable	Coefficient	z-value	Variable	Coefficient	z-value
(health)2	−0.2153	(−5.66)	log(income/happiness)	1.0177	(220.38)
(health)3	−0.6819	(−18.28)	log(gdppc)	−0.1388	(−39.22)
(health)4	−1.0829	(−28.94)	(wave)3	−0.2374	(−11.30)
(health)5	−1.4604	(−38.69)	(wave)4	−0.2984	(−14.08)
(education)2	−0.1319	(−12.65)	(wave)5	−0.3369	(−16.17)
(education)3	−0.2411	(−19.90)	(wave)6	−0.4763	(−22.59)
(education)4	−0.3403	(−33.33)	Unemployment rate	0.0165	(35.85)
(education)5	−0.3065	(−26.05)	Gini index	−0.0112	(−36.06)
(education)6	−0.3641	(−35.00)	Social protection expenditure	0.0010	(1.95)
(education)7	−0.5190	(−41.36)	Corruption Perception Index	0.0069	(18.41)
(education)8	−0.6430	(−59.17)	Quality of governance (WGI)	−0.1611	(−18.60)
$\ln \sigma^2_{u_i}$			$\boldsymbol{\mu}$		
(Intercept)	−2.1055	(−11.13)	μ_1	1.1757	(21.15)
Age	−0.0496	(−10.45)	μ_2	1.6231	(28.99)
Age Squared	0.0006	(12.63)	μ_3	2.0977	(37.14)
Female	0.1714	(6.63)	μ_4	2.4734	(43.45)
Religious	0.4708	(14.92)	μ_5	3.1102	(53.83)
Unemployed	−4.8111	(−4.29)	μ_6	3.5014	(59.99)
Married	1.2509	(18.52)	μ_7	4.0003	(67.57)
			μ_8	4.6686	(77.18)
			μ_9	5.1694	(83.91)
N = 210218, log-likelihood = −403011.98					

4.3 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 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 (7),[17]

$$\begin{aligned}\frac{\partial \texttt{Prob}(y_i = m)}{x_k} &= \frac{\partial\left[\Phi(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} - u_i) - \Phi(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} - u_i)\right]}{\partial x_k} \\ &\approx \frac{\partial\left[\Phi(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} - E(u_i)) - \Phi(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} - E(u_i))\right]}{\partial x_k}. \end{aligned} \tag{26}$$

Now we can spell out the marginal effect for each group, viz.,

$$\frac{\partial \texttt{Prob}(y_i = m)}{z_k} = \begin{cases} -\dfrac{\partial E(u_i)}{\partial z_k}\phi(\mu_1 - \boldsymbol{x}_i'\boldsymbol{\beta} - E(u_i)) & \text{if } y_i = 1 \\ -\dfrac{\partial E(u_i)}{\partial z_k}\begin{bmatrix} \phi(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} - E(u_i)) \\ -\phi(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} - E(u_i)) \end{bmatrix} & \text{if } 1 < y_i = m < M \\ \dfrac{\partial E(u_i)}{\partial z_k}\phi(\mu_{M-1} - \boldsymbol{x}_i'\boldsymbol{\beta} - E(u_i)) & \text{if } y_i = M \end{cases}. \tag{27}$$

The values in (27) 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

$$\begin{aligned}\Delta p_{im} &= \texttt{Prob}(y_i = m | z_k = 1) - \texttt{Prob}(y_i = m | z_k = 0) \\ &\approx \begin{cases} \begin{bmatrix} \Phi(\mu_1 - \boldsymbol{x}_i'\boldsymbol{\beta} - E(u_i)|x_k = 1) \\ -\Phi(\mu_1 - \boldsymbol{x}_i'\boldsymbol{\beta} - E(u_i)|x_k = 0) \end{bmatrix} & \text{if } y_i = 1 \\ \begin{bmatrix} \Phi(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} - E(u_i)|x_k = 1)) \\ -\Phi(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} - E(u_i)|x_k = 1)) \\ -\Phi(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} - E(u_i)|x_k = 0)) \\ +\Phi(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} - E(u_i)|x_k = 0)) \end{bmatrix} & \text{if } 1 < y_i = m < M \\ \begin{bmatrix} \Phi(\mu_{M-1} - \boldsymbol{x}_i'\boldsymbol{\beta} - E(u_i)|x_k = 0))) \\ -\Phi(\mu_{M-1} - \boldsymbol{x}_i'\boldsymbol{\beta} - E(u_i)|x_k = 1))) \end{bmatrix} & \text{if } y_i = M \end{cases}. \end{aligned} \tag{28}$$

[17] Calculation of $E(u_i)$ is introduced below.

Table 3: Marginal effects on probability of belonging to a happiness category

Group	Unemployment rate	Gini index	Social protection expenditure	Corruption Perception Index	Quality of governance (WGI)
10	−0.26	0.17	−0.016	−0.11	2.52
9	−0.12	0.08	−0.007	−0.05	1.13
8	−0.11	0.07	−0.007	−0.05	1.07
7	−0.01	0.00	−0.000	−0.00	0.07
6	0.04	−0.03	0.003	0.02	−0.42
5	0.13	−0.09	0.008	0.05	−1.23
4	0.08	−0.06	0.005	0.03	−0.79
3	0.09	−0.06	0.005	0.04	−0.86
2	0.06	−0.04	0.004	0.03	−0.59
1	0.09	−0.06	0.006	0.04	−0.90

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

Similar to the continuous case, the values in (28) are calculated for all observations and then the APE is reported.

Table 3 shows the marginal effects of continuous regressors on probability of belonging to a happiness category. Reported are the average partial effects for each category. The marginal effects are multiplied by 100 to facilitate interpretation in percent. First, note that the numbers in each column sum to zero since the probabilities sum to 1. The marginal effects in this table have the following interpretation. If, for example, the unemployment rate in the country increases by 1 percent, holding everything else fixed, the probability of observing a person in happiness category 1 (the lowest) increases by 0.09%. The probability of observing an individual in higher happiness 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 happiness category 10 (the highest) decreases by 0.26%. Thus, we conclude that the unemployment rate has a negative effect on happiness and this effect is non-linear.

The evidence from the column with Gini index suggests that the more inequality, holding everything else constant, increases the probability of being in the highest happiness category. Thus reduction in inequality predicts less happiness. Increase in social protection expenditure implies lower probability in higher happiness categories and higher probabilities in lower happiness categories. These results contradict to some extent those obtained in previous studies that only analyze the determinants of well-being (e.g., Pacek and Radcliff, 2008; Verme, 2011). However, the values of the coefficients are quite small, so that the influence of these variables on the estimated measures of attainment is practically zero.

Table 3 also reveals that the CPI has a detrimental effect on happiness. This is not surprising 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 leads to decrease in the probabilities to be in higher

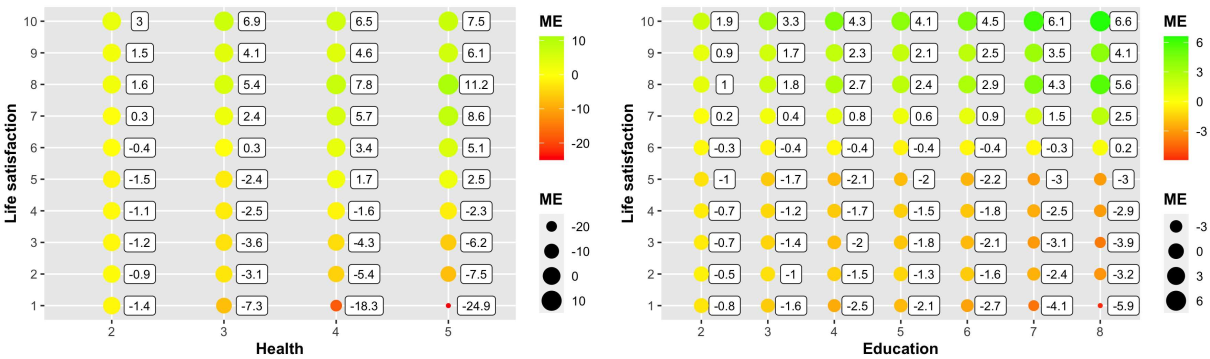

Figure 2: Marginal effects of health and education on probability of belonging to a happiness category. Reported are average marginal effects over all individuals in their respective happiness category

happiness categories and increases the probability of being in lower happiness categories, although again the values are quite small. The WGI index acts in the opposite direction and, in this case, the marginal effects do represent relevant percentages. Thus, the improvement in quality of governance has a positive effect on happiness 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 happiness category 1. The marginal effects remain virtually the same for happiness 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 the category 10. The trend for the marginal effects that are negative for category 1 is the opposite.

Figure 2 summarizes the marginal effects (multiplied by 100) of dummy variables. The size of the marginal effects is shown by color and size of a marker. The actual value is shown to the left of the marker. As with continuous variables the sum of average marginal effects is 0. Consider first the effect of health. As health status increases, i.e., move rightwards in the left panel of Figure 2, the probability of being in lower happiness categories gets smaller. Having health status of either 4 or 5 (the highest) holding everything else fixed, decreases the probability of being in the lowest happiness category by 18 and 25 percent, respectively. The same changes in health improve the odds of being in the highest happiness category by 6.5 and 7.5 percent, respectively. When we consider changing probabilities of belonging to happiness categories within the same dummy variables, for example health status is 2, the change in probabilities are negative for lower happiness categories and positive for the higher categories. As the health status is higher, the changes in probabilities follow the same patterns, however the magnitudes are amplified. The effects are strong and health has unquestionably positive effect on happiness.

Consider the right panel of Figure 2. Education has a similar effect on happiness. Better education (movement left to right) increases probability of belonging to a higher happiness category and decreases probability of being unhappy. If we consider the original frequencies of observing individuals in happiness categories shown in Figure 1, the marginal effects of both health and education are quite substantial. For example, there are 12.5 percent of individuals in the happiest category. The highest education alone (holding everything else fixed) could boost this number to 19 percent.

5 Prediction of happiness-income shortfall

5.1 Happiness attainment

The task after estimating the parameters of the model and their marginal effects on probability of belonging to a category is to predict happiness and income 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., $f(u_i|\epsilon_i)$ which is

$$f(u_i|\epsilon_i) = \frac{f(u,\epsilon_i)}{f(\epsilon_i)} = \frac{\frac{2}{\sigma_u}\phi\left(\frac{u}{\sigma_u}\right) \times \sum_{m=1}^{M} I_m(y_i)p_{im}}{\sum_{m=1}^{M} I_m(y_i)\mathcal{L}_m}. \tag{29}$$

Note that after estimating the parameters of the model, we can compute ϵ_i from the residuals. Since the denominator in (29) does not depend on u_i, the expected value of $u_i|\epsilon_i$ is

$$E(u_i|\epsilon_i) = \frac{1}{\sum_{m=1}^{M} I_m(y_i)\mathcal{L}_m} \int_0^\infty u_i \frac{2}{\sigma_u}\phi\left(\frac{u_i}{\sigma_u}\right) \times \sum_{m=1}^{M} I_m(y_i)p_{im}du_i. \tag{30}$$

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

$$\int_0^\infty u\frac{2}{\sigma_u}\phi\left(\frac{u}{\sigma_u}\right)\Phi(\mu^* - \boldsymbol{x}_i'\boldsymbol{\beta} - u_i)du_i, \tag{31}$$

which can be expressed as

$$2\sigma_u\left[\frac{-\sigma_u}{\sqrt{1+\sigma_u^2}}\phi\left(\frac{\mu^* - \boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}\right)\Phi\left(\frac{\sigma_u(\mu^* - \boldsymbol{x}_i'\boldsymbol{\beta})}{\sqrt{1+\sigma_u^2}}\right)\right] + \frac{1}{\sqrt{2\pi}}\Phi(\mu^* - \boldsymbol{x}_i'\boldsymbol{\beta}). \tag{32}$$

Thus, the prediction of u_i is

$$E(u_i|\epsilon_i) = 2\sigma_u\frac{\sum_{m=1}^{M} I_m(y_i)\mathcal{B}_m}{\sum_{m=1}^{M} I_m(y_i)\mathcal{L}_m}, \tag{33}$$

Table 4: Summary statistics of happiness and income attainment by group

Group	N	sd	min	mean	median	max
Africa	33946	0.133	0.202	0.746	0.739	0.984
Asia	21256	0.118	0.190	0.690	0.682	0.984
Eastern Europe	17353	0.124	0.190	0.706	0.698	0.984
Latin and Central America	34929	0.125	0.235	0.710	0.714	0.984
Middle East	14208	0.123	0.267	0.715	0.706	0.984
OECD	25033	0.112	0.188	0.708	0.706	0.984
Other	9409	0.107	0.268	0.713	0.703	0.984
Post Soviet	32304	0.118	0.173	0.719	0.711	0.984
Western Europe	21780	0.120	0.267	0.704	0.693	0.984
Total	210218	0.123	0.173	0.714	0.708	0.984

where

$$\mathcal{B}_m = \begin{cases} \left[\frac{-\sigma_u}{\sqrt{1+\sigma_u^2}} \phi\left(\frac{\mu_1 - \boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}\right) \Phi\left(\frac{\sigma_u(\mu_1 - \boldsymbol{x}_i'\boldsymbol{\beta})}{\sqrt{1+\sigma_u^2}}\right) + \frac{1}{\sqrt{2\pi}}\Phi(\mu_1 - \boldsymbol{x}_i'\boldsymbol{\beta}) \right] & \text{if } y_i = 1 \\ \left[\frac{-\sigma_u}{\sqrt{1+\sigma_u^2}} \phi\left(\frac{\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}\right) \Phi\left(\frac{\sigma_u(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta})}{\sqrt{1+\sigma_u^2}}\right) + \frac{1}{\sqrt{2\pi}}\Phi(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta}) - \frac{-\sigma_u}{\sqrt{1+\sigma_u^2}} \phi\left(\frac{\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}\right) \Phi\left(\frac{\sigma_u(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta})}{\sqrt{1+\sigma_u^2}}\right) - \frac{1}{\sqrt{2\pi}}\Phi(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta}) \right] & \text{if } 1 < y_i = m < M \\ \left[\frac{\sigma_u}{\sqrt{1+\sigma_u^2}} \phi\left(-\frac{\mu_{M-1} - \boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}\right) \Phi\left(\frac{\sigma_u(\mu_{M-1} - \boldsymbol{x}_i'\boldsymbol{\beta})}{\sqrt{1+\sigma_u^2}}\right) + \frac{1}{\sqrt{2\pi}}\Phi(-(\mu_{M-1} - \boldsymbol{x}_i'\boldsymbol{\beta})) \right] & \text{if } y_i = M. \end{cases} \tag{34}$$

To predict happiness and income shortfall, u_i, we use $E(u_i|\epsilon_i)$ as it predictor. These numbers for each observation gives an estimate of happiness and income shortfall. We can use them to rank the shortfall or use $1 - E(u_i|\epsilon_i)$ to measure happiness and income attainment (efficiency).

Table 4 present descriptive statistics of happiness attainment after classifying the countries into different groups according to the geographical area to which they belong (see Table A2 in Annex). This information is complemented by the estimated densities reported in Figure 3. The main conclusion to be drawn from both is that there are no significant differences among groups. Nevertheless, the happiness attainment seems to be slightly lower in Asia than elsewhere, while the highest median values are registered in Africa. Specifically, the median value is 0.682 in Asia vs 0.739 in Africa. The possible explanation for these results lies mainly in the lower and higher values of the inputs presented by African and Asian individuals, since the values of the two outputs are close to the average of the sample as a whole, as can be seen in Figure 4.

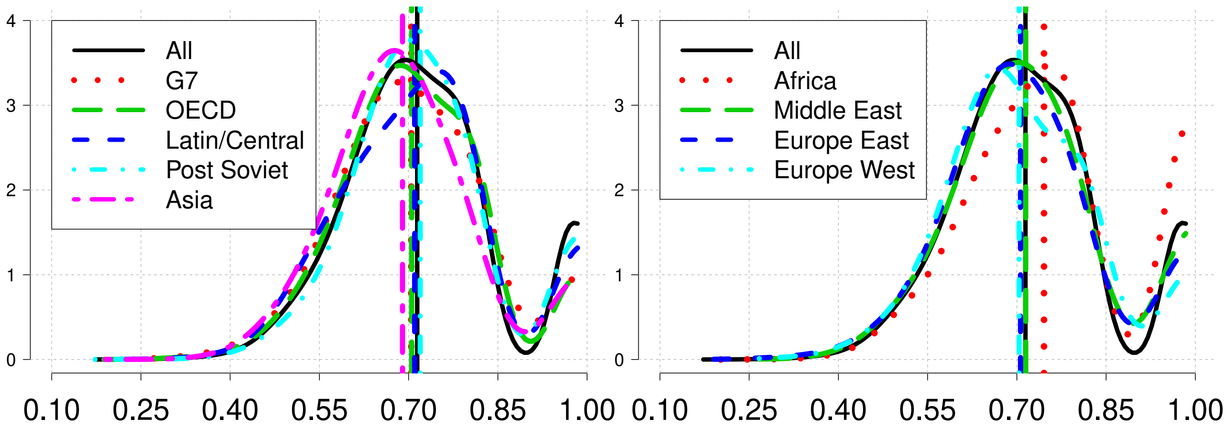

Figure 3: Kernel densities of estimated happiness attainment by groups of countries

Another noteworthy result is that obtained for the countries belonging to Latin and Central America, which several previous studies using frontier methods in this framework identified as top performers in terms of wellbeing efficiency measures (Debnath and Shankar, 2014; Cordero et al., 2021), but in our case are placed very close to the average. This is probably due to the consideration of income as an additional output, with the mean values of this variable being considerably lower than the average, which compensates to some extent for the fact that individuals in these countries report the highest welfare values in the sample, as can be seen in Figure 4. In contrast, Western European countries have above-average values for both outputs, but since they also have high values for the inputs, they are slightly below the average in terms of attainment (efficiency).

While considering marginal effects of determinants of both attainment and probability of belonging to happiness category seems a more interesting exercise, it is astonishing to observe large happiness and income shortfall in all countries. To be precise, given personal characteristics, this shortfall is 30 percent on average, so there is clearly ample room for improvement in all countries.

5.2 Marginal effects of determinants on happiness and income shortfall

Our next goal is the examine the marginal effect of z_{ki} on happiness and income shortfall. We do this from the derivative of $E(u_i)$ with respect to z_{ki}.[18]

Similar to the marginal effects of x_k on $\texttt{Prob}(y_i = m)$, the marginal effects of a variable z_k on $E(u)$ are calculated differently for continuous and discrete (binary) variables, and the magnitude of the marginal effect is not constant. By analogy, it depends on values of all the determinants $\boldsymbol{z}$, the regressors $\boldsymbol{x}$, and estimated parameters $\boldsymbol{\beta}$, $\boldsymbol{\mu}$ and $\boldsymbol{\gamma}$.

[18] It is also possible to compute the marginal effects of z_{ki} from the conditional mean $E(u_i|\epsilon_i)$, instead of the unconditional mean of u_i, as in Kumbhakar and Sun (2013).

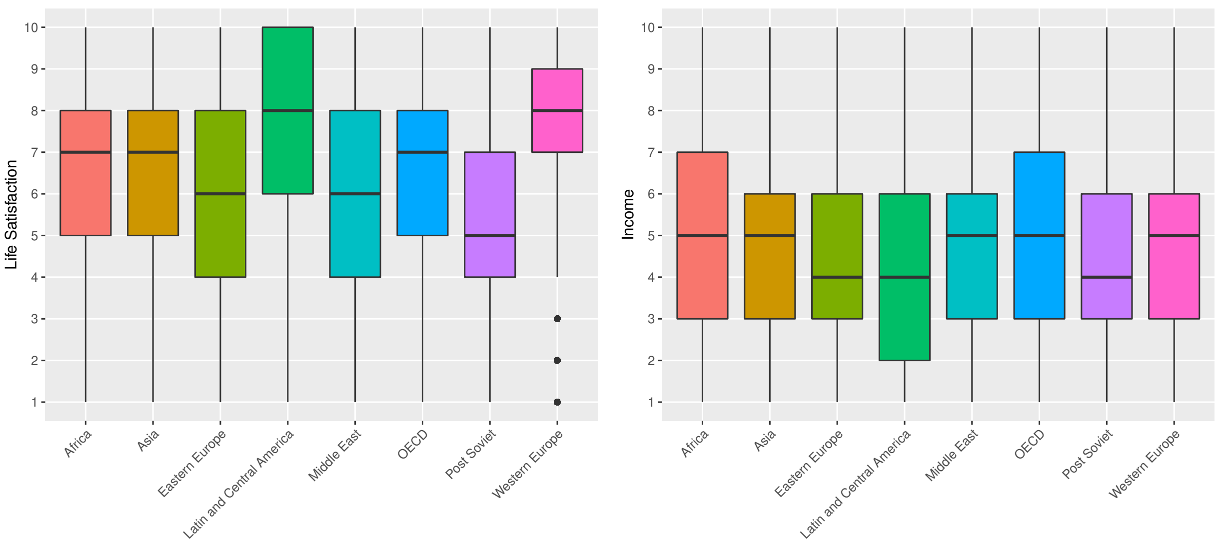

Figure 4: Life satisfaction/Happiness and Income by group of countries

If z_k is a continuous variable, its marginal effect on happiness and income shortfall is

$$\frac{\partial E(u_i)}{\partial z_k} = 2\frac{\mathcal{B}}{\mathcal{L}}d_1 + 2\sigma_u \frac{(\mathcal{B}_z\mathcal{L} - \mathcal{B}\mathcal{L}_z)}{\mathcal{L}^2}, \tag{35}$$

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

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

In the last equation,

$$d_0 = \frac{\partial(\boldsymbol{z\gamma})}{\partial z_k}. \tag{36}$$

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

$$\mathcal{L}_z = \begin{cases} d_2(\mu_1 - \boldsymbol{x}_i'\boldsymbol{\beta})f(h_1) - d_1\dfrac{\exp\left(-h_1^2(1+\sigma_u^2)/2\right)}{\pi(1+\sigma_u^2)} & \text{if } y_i = 1 \\ \left[\begin{array}{l} d_2(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta})f(h_m) - d_1\dfrac{\exp\left(-h_m^2(1+\sigma_u^2)/2\right)}{\pi(1+\sigma_u^2)} \\ -d_2(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta})f(h_{m-1}) + d_1\dfrac{\exp\left(-h_{m-1}^2(1+\sigma_u^2)/2\right)}{\pi(1+\sigma_u^2)} \end{array}\right] & \text{if } 1 < y_i = m < M\ , \\ -d_2(\mu_{M-1} - \boldsymbol{x}_i'\boldsymbol{\beta})f(h_{M-1}) + d_1\dfrac{\exp\left(-h_{M-1}^2(1+\sigma_u^2)/2\right)}{\pi(1+\sigma_u^2)} & \text{if } y_i = M \end{cases} \tag{37}$$

where $f(h)$ is defined in (25) and h_i is defined in (23).

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

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

$$\mathcal{B}_z = \begin{cases} 0.5f(h_1)\left(d_3 - h_1^2 a d_2\right) - h_1 a g(h_1) d_3 & \text{if } y_i = 1 \\ \begin{bmatrix} 0.5f(h_m)\left(d_3 - h_m^2 a d_2\right) - h_m a g(h_m) d_3 \\ -0.5f(h_{m-1})\left(d_3 - h_{m-1}^2 a d_2\right) + h_{m-1} a g(h_{m-1}) d_3 \end{bmatrix} & \text{if } 1 < y_i = m < M\ , \\ 0.5f(h_{M-1})\left(-d_3 - h_{M-1}^2 a d_2\right) + h_{M-1} a g(h_{M-1}) d_3 & \text{if } y_i = M \end{cases} \quad (38)$$

where

$$g(h) = \phi\left(h\right)\phi\left(-ha\right)$$

and

$$d_3 = \frac{\partial\left\{\frac{-\sigma_u}{\sqrt{1+\sigma_u^2}}\right\}}{\partial z_k} = -\frac{1}{\sqrt{1+\sigma_u^2}}d_1 - \sigma_u d_2.$$

For a binary variable z_k, the marginal effect of z_k on happiness and income shortfall is calculated as the difference in $E(u_i)$ given in (33) evaluated at $z_k = 1$ and $z_k = 0$.

5.3 Marginal effects of z_k on happiness and income attainment

For a continuous z_k, the marginal effects of z_k on happiness and income attainment $\exp\left(-u\right)$ can be obtained by differentiating it, which is

$$\frac{\partial\exp\left(-u\right)}{\partial z_k} \approx \frac{\partial\exp\left(-E(u)\right)}{\partial z_k} = -\exp\left(-E(u)\right)\frac{\partial E(u)}{\partial z_k} \quad (39)$$

We have observed large shortfall in happiness and income attainment. Now we would like to explore possible explanatory factors for that. Table 5 summarizes marginal effects of determinants $\boldsymbol{z}$ on happiness and income attainment. 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 for the probabilities.

The shortfall is bigger for women in all happiness categories. The largest shortfall is in the happiest category. Everything else being the same, women's happiness and income attainment is 0.31 lower than that for men. The same effect is for individuals that are religious, however the magnitude of the effect is three

Table 5: Marginal effects on happiness and income attainment

Happiness		Marginal effect (multiplied by 100)				
Category	Frequency, %	Female	Religious	Unemployed	Married	Age
10	12.44	−0.31	−0.89	3.63	−2.25	−0.04
9	10.10	−0.21	−0.59	2.63	−1.51	−0.01
8	18.11	−0.33	−0.94	4.38	−2.40	−0.02
7	14.76	−0.24	−0.68	3.36	−1.74	−0.01
6	10.88	−0.16	−0.46	2.41	−1.19	−0.01
5	14.35	−0.19	−0.55	3.06	−1.42	−0.01
4	6.04	−0.07	−0.21	1.23	−0.55	−0.02
3	5.34	−0.06	−0.18	1.07	−0.46	−0.02
2	3.28	−0.04	−0.11	0.66	−0.28	−0.02
1	4.71	−0.05	−0.14	0.90	−0.37	−0.03

Note: Reported are average marginal effects over all individuals in their respective happiness category. The marginal effects do not sum to zero.

times that of the gender dummy variable. Married is the third dummy variable which exhibits negative effect on attainment. For example, being married reduces attainment of the most happy (category 10) by 2.25. This is quite a big effect. In contrast being married reduces attainment only 0.37 for the least happy individuals.

The effect of age on attainment deserves special attention as we model nonlinear relationship. By looking at the last column in Table 5, we can conclude that as person gets older the attainment is necessarily decreasing. The marginal effect for a continuous variable can provide more information than that of a dummy variable due to its greater variability. Figure 5 shows the marginal effect (multiplied by 100) of age plotted for each individual in our sample – which is larger than 200,000 – against the age of an individual. We observe that the happier is the individual the larger is the marginal effect. Up to the age of 40 years (considering coefficients of age and ages squared in Table 2, $0.0496/2/0.0006 = 40.59$), the marginal effect of age on attainment is positive albeit diminishing. After the age of 41, the effect of age is negative and is getting bigger. By the age of 75, the attainment could be reduced by 1 (out of 100) with additional year.

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

The marginal effects of a variable z_k on $\texttt{Prob}(y_i = m)$ are differently calculated for continuous and discrete variables. For a continuous z_k we use (7) to obtain,

$$\begin{aligned} \frac{\partial \texttt{Prob}(y_i = m)}{z_k} &= \frac{\partial \left[\Phi(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} - u_i) - \Phi(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} - u_i)\right]}{\partial z_k} \\ &\approx \frac{\partial \left[\Phi(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} - E(u_i)) - \Phi(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} - E(u_i))\right]}{\partial z_k}, \end{aligned} \tag{40}$$

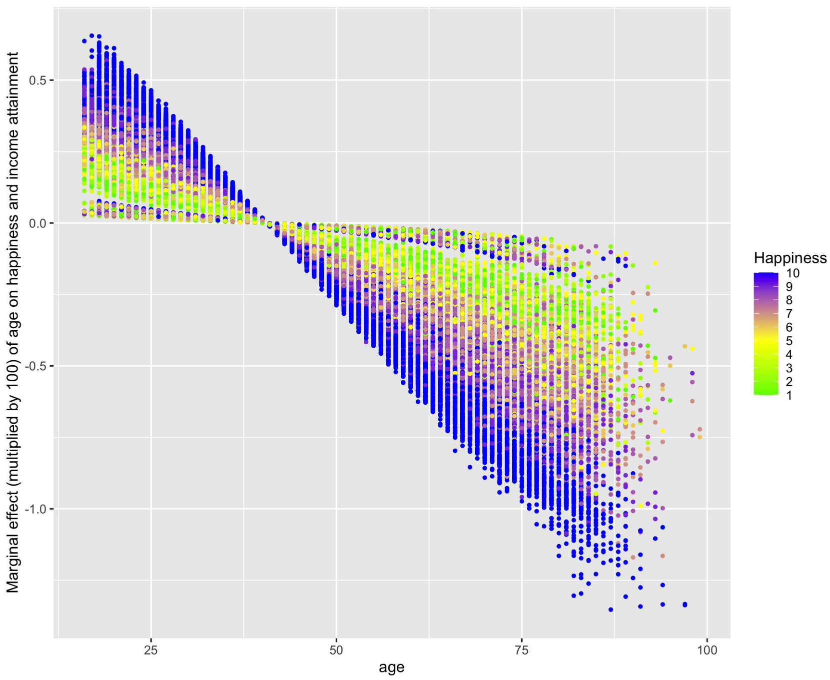

Figure 5: Marginal effect of age on happiness and income attainment attainment

and therefore,

$$\frac{\partial \texttt{Prob}(y_i = m)}{z_k} = \begin{cases} -\dfrac{\partial E(u_i)}{\partial z_k}\phi(\mu_1 - \boldsymbol{x}_i'\boldsymbol{\beta} - u_i) & \text{if } y_i = 1 \\ -\dfrac{\partial E(u_i)}{\partial z_k}\left[\begin{array}{l} \phi(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} - u_i) \\ -\phi(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} - u_i) \end{array}\right] & \text{if } 1 < y_i = m < M\ . \\ \dfrac{\partial E(u_i)}{\partial z_k}\phi(\mu_{M-1} - \boldsymbol{x}_i'\boldsymbol{\beta} - u_i) & \text{if } y_i = M \end{cases} \tag{41}$$

The values in (41) are calculated for all observations and then the average is reported, aka APE.

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

$$\Delta p_{im} = \texttt{Prob}(y_i = m|z_k = 1) - \texttt{Prob}(y_i = m|z_k = 0)$$

Table 6: Marginal effects on $\texttt{Prob}(y_i = m)$,

Happiness		Marginal effect (multiplied by 100)				
Category	Frequency, %	Female	Religious	Unemployed	Married	Age
10	12.44	0.421	1.21	−4.73	2.93	0.0026
9	10.10	0.191	0.54	−2.48	1.36	0.0021
8	18.11	0.173	0.48	−2.74	1.26	0.0035
7	14.76	−0.001	−0.01	−0.59	0.04	0.0023
6	10.88	−0.081	−0.24	0.64	−0.55	0.0011
5	14.35	−0.218	−0.62	2.53	−1.53	−0.0002
4	6.04	−0.133	−0.38	1.79	−0.95	−0.0012
3	5.34	−0.138	−0.39	2.02	−1.00	−0.0023
2	3.28	−0.090	−0.25	1.42	−0.66	−0.0023
1	4.71	−0.123	−0.34	2.15	−0.91	−0.0056

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

$$\approx \begin{cases} \begin{bmatrix} \Phi(\mu_1 - \boldsymbol{x}_i'\boldsymbol{\beta} - E(u_i)|z_k = 1) \\ -\Phi(\mu_1 - \boldsymbol{x}_i'\boldsymbol{\beta} - E(u_i)|z_k = 0) \end{bmatrix} & \text{if } y_i = 1 \\ \begin{bmatrix} \Phi(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} - E(u_i)|z_k = 1)) \\ -\Phi(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} - E(u_i)|z_k = 1)) \\ -\Phi(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} - E(u_i)|z_k = 0)) \\ +\Phi(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} - E(u_i)|z_k = 0)) \end{bmatrix} & \text{if } 1 < y_i = m < M\ . \\ \begin{bmatrix} \Phi(\mu_{M-1} - \boldsymbol{x}_i'\boldsymbol{\beta} - E(u_i)|z_k = 0))) \\ -\Phi(\mu_{M-1} - \boldsymbol{x}_i'\boldsymbol{\beta} - E(u_i)|z_k = 1))) \end{bmatrix} & \text{if } y_i = M \end{cases} \quad (42)$$

Similar to the continuous case, the values in (42) are calculated for all observations and then the APE is reported.

Table 6 shows the marginal effects of shortfall determinants on probability of belonging to a happiness category. Reported are the average effects in the respective happiness category. The 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 2.1. For example, with regard to gender, women are 0.421 percent more likely than men to be among happiest individuals. On the contrary, men are 0.123 percent more likely than women to be among the least happy individuals. 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 6 also suggests that being married has also a relevant effect on happiness. More specifically, married individual is 2.93 percent more likely to be in the happiest category, 1.36 percent more likely to

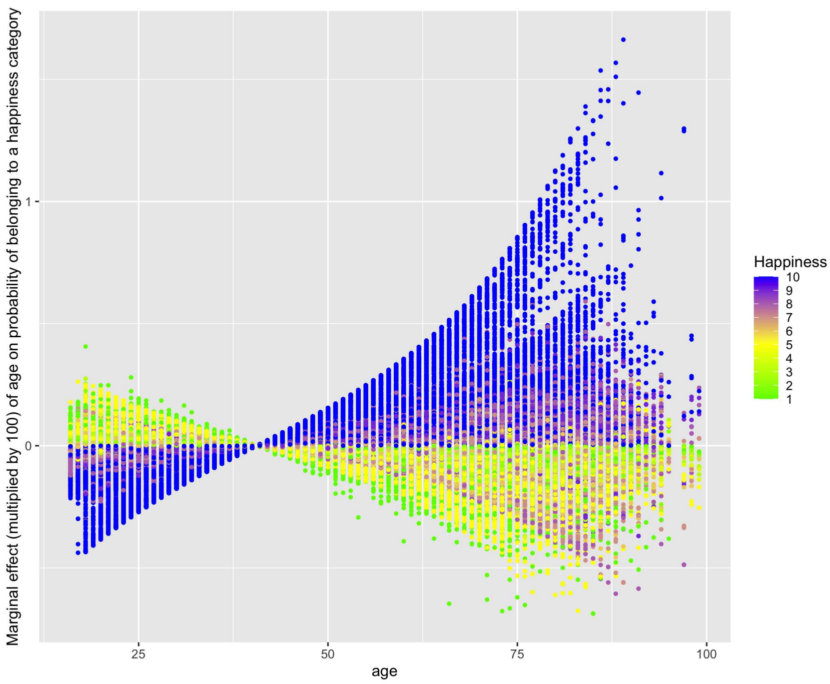

Figure 6: Marginal effect of age on probability to be in a given happiness group

belong to category 9. being married (or people living with a partner) reduces the probability of being the least happy by 0.91 percent. Given that there are only 4.71 percent in the least happy category (see column 2), the effect of being married seems to be also substantial. Finally, being unemployed has by far the most sizeable effect on happiness. An unemployed individual is 4.73 percent less likely to belong to the happiest group of individuals. Recall that there only 12.44 percent of those. On the other side of the happiness spectrum, being unemployed increases the probability by 2.15 percent to be among the unhappiest, whose proportion is 4.71 percent. The last column of Table 6 reveals that a person is getting happier with age. Since variable age is continuous, we can see the effect of age (multiplied by 100) in greater detail in Figure 6. Which up to 40 years of age, each additional year has a detrimental effect on happiness, after the age of 41, the effect is increasingly positive. This result practically coincides with the idea of a U-shaped relationship suggested by Blanchflower and Oswald (2008).

6 Conclusion

This paper focuses on the analysis of subjective well-being from the perspective of a stochastic frontier technique to estimate the extent to which individuals are being fully efficient in converting their resources and/or personal characteristics into well-being or life satisfaction. Within this framework, we made two main contributions to the existing literature. First, we consider that income is generated along with happiness by individual characteristics, thus we measure individual's ability to achieve their maximum happiness along with the highest possible level of income. As a result, it is possible for them to exhibit certain level of shortfall in both areas, i.e., not attain happiness and income to their full potential. Second, in order to explore the potential causes of this shortfall we rely on a stochastic frontier approach adapted to the presence of ordered categorical indicators as dependent variables representing happiness and income. The happiness relationship has therefore been transformed into an ordered probit model. We further did required additional algebraic calculations to properly interpret the estimated coefficients and derived marginal effects of resources and the determinants of happiness shortfall.

In our empirical study, we have estimated a happiness frontier for a huge sample of almost 210,000 individuals from 74 different countries participating in different waves of the World Value Survey. Our results show that the overall average happiness efficiency is around 70 percent. This value is close to those estimated in some recent studies with a similar approach (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 happiness frontier.

If all the countries analyzed are grouped by geographic areas, we find no relevant differences among individuals average happiness efficiency levels, although slightly lower values are observed for the Asian countries and somewhat higher for African countries. This is a striking result considering that there are notable differences between countries in terms of both average levels of well-being and income reported by individuals. Thus, it seems that the process of converting resources into outputs (happiness and income) is not so much affected by where individuals live. In this sense, our results indicate that most of the indicators representing countries´ economic and social characteristics have quite relatively small marginal effects on our estimated measures of happiness attainment. The only exception in this regard is represented by the quality of governance, which has a positive and quite relevant effect on happiness attainment, corroborating evidence widely demonstrated in previous literature.

The marginal effects estimated for the individual determinants of the shortfall also point in the same direction as that of other previous studies focused on analyzing the determinants of subjective well-being. Therefore, it seems to be a clear link between both lines of research. Specifically, we find that women, religious people and those being married have higher probabilities of reporting the higher levels of well-being and less likely to be lower to be among the most unhappy individuals. 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 the age the relationship is U-shaped, reaching the lowest values between the ages of 40 and 45.

Despite these interesting results, it is worth mentioning that there might be unobserved latent factors influencing the production process that have been ignored in our empirical analysis. For instance, previous research has demonstrated that personality traits might also have an impact on the process of converting resources into well-being and/or incomes (Lucas and Diener, 2015; Mamatzakis and Tsionas, 2021). Unfortunately, data regarding those traits is not available in our main source of data (World Value Survey), thus we have not been able to account for this. However, our choice of using a stochastic frontier model in which unobserved heterogeneity is captured by the error term mitigates to some extent the potential problems of bias that might arise due to not considering these variables in the analysis. In any case, we believe that it will be necessary to further study the determinants of well-being in greater depth in future studies. In this regard, the recent development of longitudinal international databases that provide information on a multitude of factors that may influence well-being is very useful for researchers concerned with analyzing the causes of happiness shortfalls.

References

Abdallah, S., Thompson, S. and Marks, N., "Estimating worldwide life satisfaction," *Ecological Economics*, 2008, 65(1), pp. 35–47.

Ahuvia, A.C., "Wealth, consumption and happiness," in "The Cambridge handbook of psychology and economic behaviour," , A. Lewis, ed., Cambridge University Press, 2008.

Aragon, Y., Daouia, A. and Thomas-Agnan, C., "Nonparametric Frontier Estimation: A Conditional Quantile-Based Approach," *Econometric Theory*, 2005, 21(02), pp. 358–389.

Argyle, M., "Causes and Correlates of Happiness," in "Well-being: The foundations of hedonic psychology," , D. Kahneman, E. Diener and N. Schwarz, eds., New York: Russell Sage, 1999, pp. 353–373.

Binder, M. and Broeckel, T., "Happiness no matter the cost? An examination on how efficiently individuals reach their happiness levels," *Journal of Happiness Studies*, 2012, 13, pp. 621–645.

Bjørnskov, Christian, Dreher, Axel and Fischer, Justina A V, "Cross-country determinants of life satisfaction: exploring different determinants across groups in society," *Social Choice and Welfare*, 2008, 30(1), pp. 119–173, ISSN 1432-217X.

——, "Formal institutions and subjective well-being: Revisiting the cross-country evidence," *European Journal of Political Economy*, 2010, 26(4), pp. 419–430.

Blanchflower, D.G. and Oswald, A.J., "Is Well-Being U-Shaped over the Life Cycle?" *Social Science & Medicine*, 2008, 66(8), pp. 1733–1749.

——, "International happiness: A new view on the measure of performance," *The Academy of Management Perspectives*, 2011, 25(1), pp. 6–22.

Bonini, A.N., "Cross-national variation in individual life satisfaction: Effects of national wealth, human development, and environmental conditions," *Social Indicators Research*, 2008, 87(2), pp. 223–236.

Bruni, L. and Porta, P.L., *Handbook of Economics of Happiness*, London, Elgar, 2007.

Ferrer-i Carbonell, A., "Income and well-being: An empirical analysis of the comparison income effect," *Journal of Public Economics*, 2005, 89, pp. 997–1019.

Ferrer-i Carbonell, A. and Frijters, P., "How Important is Methodology for the Estimates of the Determinants of Happiness?" *The Economic Journal*, 2004, 114, pp. 641–659.

Carboni, O.A. and Russu, P., "Assessing regional well-being in Italy: An application of Malmquist-DEA and self-organizing map neural clustering," *Social Indicators Research*, 2015, 122(3), pp. 677–700.

Cazals, C., Florens, J.P. and Simar, L., "Nonparametric frontier estimation: a robust approach," *Journal of Econometrics*, 2002, 106(1), pp. 1–25.

Clark, A.E., Frijters, P. and Shields, M.A., "Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles," *Journal of Economic Literature*, 2008, 46(1), pp. 95–144.

Cordero, Jose Manuel, Polo, Cristina and Salinas-Jiménez, Javier, "Subjective Well-Being and Heterogeneous Contexts: A Cross-National Study Using Semi-Nonparametric Frontier Methods," *Journal of Happiness Studies*, 2021, 22(2), pp. 867–886.

Cordero, José Manuel, Salinas-Jiménez, Javier and Salinas-Jiménez, M Mar, "Exploring factors affecting the level of happiness across countries: A conditional robust nonparametric frontier analysis," *European Journal of Operational Research*, 2017, 256(2), pp. 663–672.

Daraio, C. and Simar, L., "Introducing environmental variables in nonparametric frontier models: a probabilistic approach," *Journal of Productivity Analysis*, 2005, 24(1), pp. 93–121.

Deaton, A. and Stone, A.A., "Two happiness puzzles," *American Economic Review*, 2013, 103(3), pp. 591–97.

Debnath, R.M. and Shankar, R., "Does good governance enhance happiness: A cross nation study," *Social Indicators Research*, 2014, 116(1), pp. 235–253.

Di Tella, R. and MacCulloch, R., "Some uses of happiness data in economics," *Journal of Economic Perspectives*, 2006, 20, pp. 25–46.

Di Tella, R., MacCulloch, R.J. and Oswald, A.J., "The macroeconomics of happiness," *Review of Economics and Statistics*, 2003, 85(4), pp. 809–827.

Diener, E. and Biswas-Diener, R., "Will money increase subjective well-being?" *Social Indicators Research*, 2002, 57(2), pp. 119–169.

Diener, E., Diener, M. and Diener, C.L., "Factors predicting the subjective well-being of nations," *Journal of Personality and Social Psychology*, 1995, 69, pp. 851–864.

Diener, E., Gohm, C.L., Suh, E.M. and Oishi, S., "Similarity of the relations between marital status and subjective well-being across cultures," *Journal of Cross Cultural Psychology*, 2000, 31(4), pp. 419–436.

Diener, E., Inglehart, R. and Tay, L., "Theory and validity of life satisfaction scales," *Social Indicators Research*, 2013, 112, pp. 497–527.

Diener, E. and Lucas, R.E., "Personality and subjective well-being," in "Well-being: The foundations of hedonic psychology," , D. Kahneman, E. Diener and N. Schwarz, eds., Russell Sage Foundation, New York, 1999, pp. 213–229.

Diener, E., Oishi, S. and Tay, L., *Handbook of Well-Being*, Salt Lake City, UT: DEF Publishers, 2018.

Dittmann, J. and Goebel, J., "Your house, your car, your education: The socioeconomic situation of the neighborhood and its impact on life satisfaction in Germany," *Social Indicators Research*, 2010, 96, pp. 497–513.

Dolan, P., Peasgood, T. and White, M., "Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being," *Journal of Economic Psychology*, 2008, 29(1), pp. 94–122.

Easterlin, R.A., "Will raising the incomes of all increase the happiness of all?" *Journal of Economic Behavior and Organization*, 1995, 27, pp. 35–47.

Exton, C., Smith, C. and Vandendriessche, D., "Comparing Happiness across the World," *OECD Statistics Working Paper*, 2015.

Farrell, M.J., "The measurement of productive efficiency," *Journal of the Royal Statistical Society: Series A (General)*, 1957, 120(3), pp. 253–281.

Frey, B. and Stutzer, A., *Happiness and economics: How the economy and institutions affect human well-being*, New Jersey: Princeton University Press. Princeton, 2010.

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

Gerdham, U. and Johannesson, M., "The Relationship between Happiness, Health, and Socio-economic Factors: Results Based on Swedish Microdata," *Journal of Socio-Economics*, 2001, 30, pp. 553–7.

Gratton, C. and Taylor, P., “The economics of work and leisure,” in “Work and Leisure,” , J.T. Haworth and A.J. Veal, eds., London: Routledge, 2004, pp. 85–106.

Griffiths, William, Zhang, Xiaohui and Zhao, Xueyan, “Estimation and efficiency measurement in stochastic production frontiers with ordinal outcomes,” *Journal of Productivity Analysis*, Aug 2014, 42(1), pp. 67–84.

Haller, M. and Hadler, M., “How Social Relations and Structures Can Produce Happiness and Unhappiness: An International Comparative Analysis,” *Social Indicators Research*, 2006, 75, pp. 169–216.

Hartog, J. and Oosterbeek, H., “Health, wealth and happiness: why pursue a higher education?” *Economics of Education Review*, 1998, 17(3), pp. 245–256.

Harvey, A.S. and Mukhopadhyay, A.K., “When twenty-four hours is not enough: Time poverty of working parents,” *Social Indicators Research*, 2007, 82, pp. 57–77.

Haworth, J. and Lewis, S., “Work, leisure and well-being,” *British Journal of Guidance & Counselling*, 2005, 33(1), pp. 67–79.

Helliwell, J. and Huang, H.F., “How’s your government? International evidence linking good government and well-being,” *British Journal of Political Science*, 2008, 38, pp. 595–619.

Helliwell, J.F., “How’s life? Combining individual and national variables to explain subjective well-being,” *Economic Modelling*, 2003, 20(2), pp. 331–360.

Helliwell, J.F. and Barrington-Leigh, C.P., “Viewpoint: Measuring and Understanding Subjective Well-Being,” *Canadian Journal of Economics*, 2010, 43(3), pp. 729–753.

Helliwell, J.F., Huang, H. and Wang, S., “The Distribution of World Happiness,” in “World Happiness Report 2016,” , J.F. Helliwell, Layard J.F., R. and J. Sachs, eds., New York, 2016, pp. 8–48.

Helliwell, J.F., Layard, R., Sachs, J.D. and De Neve, J.E., “World Happiness Report 2020,” Tech. rep., New York, 2020.

Hsieh, C.M., “Income and financial satisfaction among adults in the United States,” *Social Indicators Research*, 2004, 66, pp. 249–266.

Inglehart, R., Haerpfer, C., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano, J. and Lagos, *World Values Survey: All Rounds - Country-Pooled Datafile*, Madrid: JD Systems Institute, 2014.

Kahneman, D. and Deaton, A., “High income improves evaluation of life but not emotional well-being,” *Proceedings of the National Academy of Sciences*, 2010, 107(38), pp. 16489–16493.

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

Krueger, A.B. and Schkade, D.A., "The reliability of subjective well-being measures," *Journal of Public Economics*, 2008, 92(8), pp. 1833–1845.

Kumbhakar, Subal C., "Specification and estimation of multiple output technologies: A primal approach," *European Journal of Operational Research*, 2013, 231(2), pp. 465 – 473.

Kumbhakar, Subal C. and Sun, Kai, "Derivation of marginal effects of determinants of technical inefficiency," *Economics Letters*, 2013, 120(2), pp. 249–253.

Kuosmanen, T. and Kortelainen, M., "Stochastic non-smooth envelopment of data: semi-parametric frontier estimation subject to shape constraints," *Journal of Productivity Analysis*, 2012, 38(1), p. 11–28.

Langbein, L. and Knack, S., "The Worldwide Governance Indicators: Six, One, or None?" *Journal of Development Studies*, 2010, 46, pp. 350–370.

Lovell, C.A.K., Richardson, S., Travers, P. and Wood, L., "Resources and functionings: A new view of inequality in Australia," in "Models and Measurement of Welfare and Inequality," , Wolfgang Eichhorn, ed., Heidelberg: Springer-Verlag, 1994, pp. 787–807.

Lucas, Richard E and Diener, Ed, "Personality and subjective well-being: Current issues and controversies." , 2015.

MacKerron, G., "Happiness economics from 35 000 feet," *Journal of Economic Surveys*, 2012, 26(4), pp. 705–735.

Mamatzakis, E.C. and Tsionas, M.G., "Making inference of British household's happiness efficiency: a Bayesian latent model," *European Journal of Operational Research*, 2021, 294(1), pp. 312–326.

McBride, M., "Relative-income effects on subjective well-being in the cross-section," *Journal of Economic Behavior & Organization*, 2001, 45, pp. 251–278.

Mizobuchi, H., "Measuring socio-economic factors and sensitivity of happiness," *Journal of Happiness Studies*, 2017, 18(2), pp. 463–504.

Narayan, D., Chambers, R., Shah, M.K. and Petesch, P., *Voices of the Poor. Crying out for Change*, New York: Oxford University Press for the World Bank, 2000.

Nettle, D., *Happiness: The science behind your smile*, Oxford: Oxford University Press, 2005.

Nikolova, M. and Popova, O., "Sometimes Your Best Just Ain't Good Enough: The Worldwide Evidence on Well-Being Efficiency," *The BE Journal of Economic Analysis & Policy*, 2021, 21(1), pp. 83–114.

Oswald, A.J., "Happiness and economic performance," *The Economic Journal*, 1997, 107(445), pp. 1815–1831.

Ott, J.C., "Good governance and happiness in nations: Technical quality precedes democracy and quality beats size," *Journal of Happiness Studies*, 2010, 11, pp. 353–368.

——, "Government and happiness in 130 nations: Good governance fosters higher level and more equality of happiness," *Social Indicators Research*, 2011, 102(1), pp. 3–22.

Owen, Donald B., "Tables for Computing Bivariate Normal Probabilities," *The Annals of Mathematical Statistics*, 1956, 27(4), pp. 1075–1090.

O'Connor, K.J., "Happiness and welfare state policy around the world," *Review of Behavioral Economics*, 2017, 4(4), pp. 397–420.

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

Peiro, A., "Happiness, satisfaction and socio-economic conditions: Some international evidence," *The Journal of Socio-Economics*, 2006, 35(2), pp. 348–365.

Powdthavee, N., "How much does money really matter? Estimating the causal effects of income on happiness," *Empirical Economics*, 2010, 39(1), pp. 77–92.

Rayo, L. and Becker, G.S., "Evolutionary efficiency and happiness," *Journal of Political Economy*, 2007, 115(2), pp. 302–337.

Roberts, K., *Leisure in contemporary society*, Wallingford: CAB International, 1999.

Salinas-Jiménez, M., Artés, J. and Salinas-Jiménez, J., "How Do Educational Attainment and Occupational and Wage-Earner Statuses Affect Life Satisfaction? A Gender Perspective Study," *Journal of Happiness Studies*, 2013, 14(2), pp. 367–388.

Schneider, S.M., "Income inequality and subjective wellbeing: Trends, challenges, and research directions," *Journal of Happiness Studies*, 2016, 17(4), pp. 1719–1739.

Schyns, P., "Cross-national differences in happiness: Economic and cultural factors explored," *Social Indicators Research*, 1998, 43(1-2), pp. 3–26.

Sen, Amartya, *Commodities and capabilities*, Amsterdam: North-Holland, 1985.

Senik, C., "What can we learn from subjective data? The case of income and well-being," *Journal ofEconomic Surveys*, 2004, 19(1), pp. 43–63.

Stevenson, B. and Wolfers, J., *Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox, Brookings Papers on Economic Activity*, Washington, DC: Spring, Brookings Institution Press, 2008.

Stutzer, A. and Frey, B.S., "Does marriage make people happy, or do happy people get married?" *The Journal of Socio-Economics*, 2006, 35(2), pp. 326–347.

——, "Recent developments in the economics of happiness: A selective overview. Recent Developments in the Economics of Happiness: A Selective Overview," , 2012, no 7078.

Van Praag, B. and Ferrer-i Carbonell, A., *Happiness Quantified: A Satisfaction Calculus Approach*, Oxford University Press, 2004.

Veenhoven, R., "Is happiness relative?" *Social Indicators Research*, 1991, 24(1), pp. 1–34.

——, "Inequality of happiness in nations: Introduction," *Journal of Happiness Studies*, 2005, 6, pp. 351–355.

——, *Bibliography of Happiness. (Section F 'Happiness and Society') World database of happiness*, Netherlands: Erasmus University Rotterdam, 2012.

Verme, Paolo, "LIFE SATISFACTION AND INCOME INEQUALITY," *Review of Income and Wealth*, 2011, 57(1), pp. 111–127.

Winkelmann, R., "Unemployment, social capital, and subjective well-being," *Journal of Happiness Studies*, 2009, 10(4), pp. 421–430.

A Sample

A.1 Observations per country

Table 7: Composition of the pooled dataset by country and wave of the survey

Country	**Wave 1 1981-84**	**Wave 2 1990-94**	**Wave 3 1994-1998**	**Wave 4 1999-2004**	**Wave 5 2005-2009**	**Wave 6 2010-2014**	**Total**
Albania			X	X			1,926
Algeria				X		X	1,842
Argentina	X	X	X	X	X	X	3,071
Armenia			X			X	2,938
Australia	X				X	X	2,261
Azerbaijan			X			X	2,705
Bangladesh			X	X			2,641
Belarus		X	X			X	3,437
Brazil		X			X	X	4,484
Bulgaria			X		X		1,730
Burkina Faso					X		701
Canada				X	X		3,479
Chile		X	X	X	X	X	3,854
China		X	X	X	X	X	5,447
Colombia			X		X	X	4,415
Cyprus					X	X	1,970
Ecuador						X	1,189
Egypt				X	X	X	5,284
Ethiopia					X		1,291
Estonia			X			X	1,509
Finland	X		X		X		1,808
France					X		875
Georgia			X		X	X	4,583
Germany			X		X	X	5,364
Ghana					X	X	2,603
Guatemala					X		924
Hungary	X		X		X		963
India		X	X	X	X	X	6,813
Indonesia				X	X		2,526
Iran				X	X		4,292
Iraq				X	X	X	5,029
Italy					X		654
Japan	X	X	X	X	X	X	2,796
Jordan				X	X	X	3,217
Kazakhstan						X	1,497
Kyrgyzstan				X		X	2,442
Lebanon						X	1,101
Macedonia			X	X			1,625
Malaysia					X	X	2,377
Mali					X		551
Mexico	X	X	X	X	X	X	6,225
Moldova			X	X	X		2,811
Morocco				X	X	X	1,222
Netherlands					X	X	2,329

(*continued on next page*)

Table 7 (*Continued*)

Country	Wave 1 1981-84	Wave 2 1990-94	Wave 3 1994-1998	Wave 4 1999-2004	Wave 5 2005-2009	Wave 6 2010-2014	Total
Nigeria		X	X	X		X	3,286
Norway			X		X		1,970
Pakistan			X	X		X	2,057
Peru			X	X	X	X	4,929
Philippines			X	X		X	2,349
Poland		X	X		X	X	2,924
Romania			X		X	X	4,053
Russia			X		X	X	5,851
Rwanda					X	X	2,478
Serbia			X	X	X		2,166
Slovenia			X		X	X	1,966
South Africa	X		X	X	X	X	10,865
South Korea	X		X	X	X	X	3,485
Spain			X	X	X	X	3,775
Sweden	X		X	X	X	X	2,986
Switzerland			X		X		2,019
Tanzania				X			962
Thailand					X	X	2,493
Trinidad & Tobago					X	X	1,927
Tunisia						X	902
Turkey			X	X	X	X	6,893
Uganda				X			542
Ukraine			X		X	X	4,592
United States	X		X	X	X	X	4,610
Uruguay			X		X	X	2,794
Uzbekistan						X	1,448
Venezuela			X	X			1,117
Yemen						X	569
Zambia					X		1,133
Zimbabwe				X		X	2,276
Total							210,218

A.2 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