I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12152

Have Econometric Analyses of Happiness Data Been Futile? A Simple Truth about Happiness Scales

Le-Yu Chen
Ekaterina Oparina
Nattavudh Powdthavee
Sorawoot Srisuma

FEBRUARY 2019

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12152

Have Econometric Analyses of Happiness Data Been Futile? A Simple Truth about Happiness Scales

Le-Yu Chen
Academia Sinica

Ekaterina Oparina
University of Surrey

Nattavudh Powdthavee
Warwick Business School and IZA

Sorawoot Srisuma
University of Surrey

FEBRUARY 2019

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Have Econometric Analyses of Happiness Data Been Futile? A Simple Truth about Happiness Scales*

Econometric analyses in the happiness literature typically use subjective well-being (SWB) data to compare the mean of observed or latent happiness across samples. Recent critiques show that com-paring the mean of ordinal data is only valid under strong assumptions that are usually rejected by SWB data. This leads to an open question whether much of the empirical studies in the economics of happiness literature have been futile. In order to salvage some of the prior results and avoid future issues, we suggest regression analysis of SWB (and other ordinal data) should focus on the median ra-ther than the mean. Median comparisons using parametric models such as the ordered probit and logit can be readily carried out using familiar statistical softwares like STATA. We also show a previously as-sumed impractical task of estimating a semiparametric median ordered-response model is also possi-ble by using a novel constrained mixed integer optimization technique. We use GSS data to show the famous Easterlin Paradox from the happiness literature holds for the US independent of any paramet-ric assumption.

JEL Classification: C24, C61, I31

Keywords: median regression, mixed-integer optimization, ordered-response model, subjective well-being

Corresponding author:
Nattavudh Powdthavee
Warwick Business School
Scarman Road
Coventry, CV4 7AL
United Kingdom

E-mail: nattavudh.powdthavee@wbs.ac.uk

* We thank Oliver Linton, João Santos Silva, Matt Shum and seminar participants at Academia Sinica for helpful comments and discussions.

1 Introduction

The study of human happiness has been cited as one of the fastest growing research fields in economics over the last two decades (Kahneman and Krueger (2006); Clark et al. (2008); Stutzer and Frey (2013)). By looking at what socioeconomic and other factors predict (or cause) people to report higher or lower scores on a subjective well-being (SWB) scale, researchers have been able to add new insights to what have become standard views in economics. For example, studies of job and life satisfaction have shown that people tend to care a great deal more about relative income than absolute income (Clark and Oswald (1996); Ferrer-i-Carbonell (2005)), while unemployment is likely to hurt less when there are more of it around (Clark (2003); Powdthavee (2007)). The use of SWB data has therefore enabled economists to directly test many of the assumptions made in conventional economic models that may have been previously untestable before the availability of proxy utility data (e.g., Di Tella et al. (2001); Stevenson and Wolfers (2013); Gruber and Mullainathan (2006); Boyce et al. (2013)). It has also led many policy makers to start redefining what it means to be successful as a community and as a nation (Kahneman et al. (2004); Stiglitz et al. (2009)).

A key component of happiness research is to understand the determinants of SWB and how they compare within and across populations. The field's rapid growth is partly due to the abundance of accessible SWB data and how they appear to be easy to analyze using familiar methods. In particular, a widely adopted approach to estimate SWB in the literature is either by linear regression (OLS) or ordinal methods (either ordered probit or logit). See, e.g., Ferrer-i-Carbonell and Frijters (2004). And, as is customary in many applied fields of economics, conclusions are then drawn based on conditional and unconditional mean comparisons using these estimates. However, happiness economists' *laissez-faire* attitude to performing mean comparisons on SWB while using seemingly harmless econometric techniques have come under recent scrutiny.

The problems originate from the facts that SWB is an *ordinal* measure and a *rank order* of any statistic between ordinal variables only makes sense if that order is stable for all increasing transformation of the ordinal variables. When the statistic of interest is the mean, the latter condition holds if and only if there is a first order stochastic dominance (FOSD). When the statistic of interest is an OLS estimate, a sufficient condition is related to there being a second order stochastic dominance. When the relevant stochastic dominance conditions do not hold, mean or conditional mean comparisons should not be used.

Take the 3-point ordinal happiness scale in the General Social Survey (GSS) for example. In the GSS, respondents are asked whether they are "1. not too happy", "2. pretty happy", or "3. happy". We know that a score of 3 is higher than a score of 2 or 1. But there is no other meaning to these numbers beyond their rankings. By applying monotonic increasing transformation of ordinal scales, which preserves the ranking order, Schröder and Yitzhaki (2017) provide conditions under which the mean ordering of SWB between groups or signs of OLS estimates can be arbitrarily changed. While it may be transparent that ordinal scales can cause problems if one treats them cardinally, similar issue exists for the latent models that are independent of ordinal scales.

Suppose we interpret observed SWB data through an ordered response model generated by latent threshold crossing conditions (cf. McFadden (1974)); noting that any monotone transformation of the latent variable and their thresholds are observationally equivalent. Bond and Lang (forthcoming in the *Journal of Political Economy*, BL hereafter) point out there can only be FOSD between two latent variables drawn from the same continuous two-parameter distribution, such as normal or logistic, if their means differ but their variances are the same. BL show the constant variance assumption is routinely rejected in the data. One implication of this is the sign of an ordered probit or logit estimate is not suitable for drawing conclusions based on the average marginal effect on the latent variable. The following example formally illustrates this point.

Example 1: Suppose we observe Y and D that respectively denote reported happiness from a 3-point scale and a female gender dummy so that $Y = \mathbf{1}[H < 0] + 2 \cdot \mathbf{1}[0 \leq H \leq 1] + 3 \cdot \mathbf{1}[H > 1]$ with latent happiness H satisfying $H|D \sim N(\beta_0 + \beta_1 D, \sigma^2(D))$. Suppose $\sigma^2(D)$ is known. Then we can identify $\beta_1 = E[H|D=1] - E[H|D=0]$ as long as $0 < \Pr[D=1] < 1$. When $\sigma^2(1) \neq \sigma^2(0)$, BL show it is possible to find an exponential function τ_1 such that if $\beta_1 > 0$ then $E[\tau_1(H)|D=1] < E[\tau_1(H)|D=0]$; if $\beta_1 < 0$ then another exponential function τ_2 can be found so that $E[\tau_2(H)|D=1] > E[\tau_2(H)|D=0]$. I.e., β_1 *can be identified but it cannot be used to rank the mean happiness between men and women.*

The results in Schröder and Yitzhaki (2017) and BL indicate that the stochastic order conditions needed to justify how SWB researchers currently go about analyzing SWB data are usually not satisfied. An open question then arises whether the econometric analyses routinely conducted on SWB data over the decades have been futile. In particular, using ordered probits, BL demonstrate this issue by taking on nine of the most well-known empirical results from the happiness literature. They show for each case that FOSD does not hold in the data by testing and rejecting the constant variance hypothesis of the latent happiness across groups and reverse their conclusions. Following this disconcerting development, currently, there is no practical suggestion on offer as to how SWB data should then be analyzed.

We propose a solution to the problems above by focusing on the *median* instead of the mean. We have three goals in this paper. The first is to restore credibility for some of the prior empirical results. The second is to propose a constructive way forward for analyzing SWB data. The third is to promote the use of median regression for categorical ordered data within economics and other social sciences. The last goal is important because the statistical reasonings behind the results in Schröder and Yitzhaki (2017) and BL apply to all categorical ordered data that can also be relevant to non-economists.

The median is a centrality measure of a distribution. Unlike the mean, the median respects the ordinal property of SWB data because it is "equivariant" to all increasing transformations. I.e., denoting the median of a random variable Z by $Med(Z) := \inf\{z : \Pr[Z \leq z] \geq 1/2\}$, let τ be an increasing function then $\tau(Med(Z)) = Med(\tau(Z))$. Therefore the ranking of the medians cannot be reversed by a monotone transformation; freeing us from the burden of stochastic dominance conditions. While it has long been documented that the median should be the preferred summary statistic for describing central tendency of ordinal datasets (e.g. see Stevens (1946)), median regression is rarely ever used to study ordinal data. In this paper we will view categorical data through the lens of an ordered response model and aim to estimate the median of the latent variable.[1]

Our first contribution consists of a simple argument that implies a fair amount of credibility for some of prior empirical results can be instantly restored by simply re-interpreting them as medians. To see this, first note that the median of the latent variable from ordered probit and logit models are automatically obtained along with the mean whenever the latter is estimated. This is due to the fact that the median of a random variable that has a symmetric distribution[2] is identical to its mean when the latter exists. Normal and logistic distributions are examples of symmetric distributions. Secondly, BL's illustrations show their ordered probit estimates, which allow for a simple form of heteroskedasticity, provide qualitatively the same conclusions as prior results that do not account for heteroskedasticity. Therefore, by equivariance, the conclusions from nine well-known studies in the happiness literature selected by BL remain well intact if one re-interprets them as median rankings instead of mean rankings.

Our argument for using the median that doubles as the mean presents one pragmatic way forward for analyzing SWB data. If one is willing to assume that latent happiness has a known symmetric distribution, as long as heteroskedasticity[3] is accounted for, researchers can proceed to do the usual mean comparison but interpret it as the median. Leading examples of these models are ordered probit and ordered logit, both of which can be readily estimated using familiar statistical softwares such as STATA (Williams (2010)). However, we may be concerned whether empirical results are dependent on the symmetry and specific parametric assumptions. An alternative approach is to estimate the median regression of a semiparametric ordered response model directly.

The semiparametric model we shall consider imposes only a conditional median restriction and allows for a general form of heteroskedasticity. It does not a priori assume equality of the mean and median of the latent variable nor does it assume its distribution. One issue is that estimating this semiparametric median regression is a notoriously challenging task. The theory on identification and estimation of the median in our model has already been established by Lee (1992), who generalizes Manski (1985)'s binary choice framework to the multiple ordered choice setting. Lee's estimator is a

[1]It is possible to do a median comparison on observed SWB directly. The signs of the difference between medians do not depend on the ordinal scale and are therefore identified.

[2]A random variable Z is said to have a symmetric distribution if and only if there exists a value z_0 such that $\Pr[z_0 < Z \leq z_0 + t] = \Pr[z_0 - t < Z \leq z_0]$ for all $t \in \mathbb{R}$.

[3]The skedastic function can even be nonparametric, e.g. see Chen and Khan (2003).

generalization of Manski's maximum score estimator (MSE), which requires the solving of a difficult non-convex and non-smooth optimization problem in order to compute it. Indeed, only relatively recently Florios and Skouras (2008) have shown that it is practical to estimate Manski's MSE by reformulating it as solution to a constrained mixed integer linear programming (MILP) problem. We extend their insights and develop a novel MILP based estimator for an ordered choice model with any finite number of outcomes.

As an illustration, we revisit the Easterlin Paradox using the GSS data. The Paradox, named after Richard Easterlin, is an empirical observation that at any given point in time people's happiness correlates positively with income. Yet, people's happiness does not trend upwards as they become richer over time (e.g. as real income per capita grows). The Paradox is one of the most well-known findings in the literature of happiness. Its existence has come under question as BL show the average happiness of people can correlate positively or negatively with aggregate income depending on the distribution of the latent happiness. Our semiparametric estimate supports the Easterlin Paradox empirically and shows its existence does not depend on symmetry or parametric assumption.

Our advocacy for the median has implications for interdisciplinary subjects outside of economics. The econometric issues raised by Schröder and Yitzhaki (2017) and BL are relevant for all other ordered categorical data. Economists are certainly not the only researchers to analyze such variables. Discrete ordinal data, subjective or otherwise, are collected in surveys and experiments across a number of research fields (e.g. biometrics, medicine, politics and psychology to name a few) and are widely used for commercial purposes (e.g. in marketing for gauging consumer appetites and sentiments etc). At the same time, not all researchers misuse these data in the sense we have described above. For instance, applied researchers in the biomedical fields are more concerned with effects of proportional odds or hazard rates rather than parameter comparisons. In contrast, other social scientists often focus on interpreting parameters as well as comparing them across samples and take a similar approach to happiness economists working with SWB data.[4]

The remaining of the paper proceeds as follows. Section 2 gives an account on how statistical analysis for discrete ordinal data have been developed in economics and other disciplines, along with how median methods can contribute. Section 3 puts forward an empirical model of happiness that is suitable for comparing the median and suggests ways to estimate it. Section 4 presents the median estimator in a semiparametric ordered response model as a solution to a MILP problem. Section 5 revisits the Easterlin Paradox using GSS data. Section 6 concludes.

[4]A similar criticism to Schröder and Yitzhaki (2017) has been raised by Liddell and Kruschke (2018), who surveyed the 2016 volumes of 3 highly rated psychology journals and found that all papers that analyzed self-reported values from the (ordinal) Likert scale used OLS.

2 Analyzing discrete ordinal data: a brief review

We consider discrete ordinal data that are used for modelling categories arranged on a horizontal spectrum. The defining property of an ordinal variable is that there is a *rank order* over values it can take but the distances between these values are arbitrary and carry no information. A discrete ordinal variable can therefore, in constrast to cardinal variables, be put on a scale like $\{1, 2, .., J\}$ without any loss of generality. These measurements, either subjective or objective, are common in social and biomedical sciences. Examples include: individual happiness (unhappy, neither happy nor unhappy, happy), severity of injury in the accident (fatal injury, incapacitating injury, non-incapacitating, possible injury, and non-injury), lethality of an insecticide (unaffected, slightly affected, morbid, dead insects) and many other concepts.

The analysis of ordinal data in a regression framework is widely acknowledged to have been co-founded by two independent sources. The first can be credited to political scientists, McKelvey and Zavoina (1975) developed the now familiar ordered probit model to study Congressional voting on the 1965 Medicare Bill. They treated the reported scale as a censored version of the latent continuous variable, where the latent variable is modelled as linear function of the respondent's covariate vector and an additive normal error. This interpretation of the latent model can be seen as a threshold crossing counterpart to the random utility framework used by McFadden (1974) to study discrete unordered choices. At about the same time, Peter McCullagh, a well-known statistician, was writing his PhD thesis published in 1977 entitled "Analysis of Ordered Categorical Data". McCullagh (1980) focuses on modelling proportional odds and proportional hazards that become prominent in the biomedical fields. He also emphasizes the crucial role that latent variable plays in modelling ordinal variables under the framework of a generalized linear model. Huge theoretical and empirical literatures focusing on ordinal data across disciplines have since grown from these works with social scientists building on McKelvey and Zavoina (1975) and biomedical scientists following McCullagh (1980). We refer readers to Greene and Hensher (2010) and reference therein for the developments in social science; Agresti (1999) for the medical science; Ananth (1997) for epidemiology.

Researchers from different fields have different attitudes on how to analyze ordinal data. Applied researchers in biomedical fields pay a great deal of attention to choosing an appropriate model for their data (goodness of fit) but place less importance on the interpretation of individual parameters (e.g. coefficients in a generalized linear model). On the other hand, researchers in social sciences often focus on the model parameters. Many social scientists even have a tendency towards using OLS to facilitate interpretation of parameters despite there being no theoretical justification to analyze ordinal data in that way. Any efforts shown to justify this approach have been entirely empirical, for example, by finding evidence that linear regression and ordered probit/logit can generate comparable results.[5]

[5] In transport Gebers (1998) compared ordered data results of an OLS model and a logistic regression model to study driver injury severity for accidents involving trucks. He found that coefficients for both models were of the same sign and generally quite similar in magnitude and statistical significance. In international relations, Major

The wide spread and growing use of OLS to analyze ordinal dependent variables have received some critical attentions recently. In economics of well-being, Schröder and Yitzhaki (2017) highlight that though results from linear regression and ordinal models may appear similar, many implications of SWB analysis using linear regression framework can be overturned by changing the ordinal scale. Liddell and Kruschke (2018) provide a similar conclusion for the research in psychology. Both papers suggest researchers to use ordered response models instead of the OLS to avoid arbitrariness of the ordinal scale. But BL show that using ordered response models suffers from analogous problems when researchers are interested in mean comparisons.

The notion of comparing the mean of ordinal variables in itself is not invalid. However, in order for the mean ranking between two ordinal variables to be identified, their ranking order must be the same for all increasing transformations. The necessary and sufficient condition for this is well-known: one variable has to *first-order stochastic dominate* the other. The first-order stochastic dominance (FOSD) condition can be tested empirically. Performing a FOSD test is straightforward for discrete distributions. E.g. with the observed ordinal scale but FOSD is typically not satisfied in SWB datasets. To test FOSD for the latent variable requires full specification of the latent distribution. For popular parametric models, such as ordered probit and ordered logit, if the latent variables have the same variance then FOSD follows from a difference in mean. However, if their variances differ then FOSD fails whenever their means differ. This knife-edge condition makes parametric identification of the mean rank order very fragile as BL have illustrated. One may ask if FOSD may hold if we remove parametric restrictions. In order to test this empirically we need to first identify and estimate the latent distribution semiparametrically or nonparametrically. There is also a large econometrics literature on this subject. E.g. see Carneiro et al. (2003), Cunha et al. (2007), Lewbel (1997, 2000), Lewbel and Schennach (2007), Honore and Lewbel (2002). However, it is not necessary to force ourselves to go through stochastic dominance conditions if we are prepared to move away from the mean.

In this paper we propose that a natural solution is to focus on the median instead of the mean. For commonly used ordered response models such as the ordered probit and logit, the median and mean are identical. Maximum likelihood estimation of the median in popular parametric models can therefore be obtained as readily as the mean using standard statistical softwares. But we do not have to limit ourselves to parametric results. There is an active econometric literature on semiparametric estimation for discrete response data that we believe can be very useful for SWB. This field of research originates from the work by Manski (1975, 1985) on maximum score estimation for binary choice and multiple unordered choice data and extended by Lee (1992) to multiple ordered choice data. Thus far, the development has been mostly theoretical with very limited real world applications,

(2012) compared the results of ordered probit and OLS models for the analysis of the effect of economics sanctions of countries with various regime types and also found that the models produce similar results. Ferrer-i-Carbonell and Frijters (2004) found the both models produce similar results for the analysis of the sources of individual well-being.

especially for non-binary discrete dependent variables. The main practical issue behind this is due to the fact that performing maximum score estimation for the discrete choice model is computationally difficult. More specifically, the ordered response median regression estimator proposed by Lee (1992) is a solution to a non-smooth and non-covex least absolute deviation (LAD) optimization problem. The corresponding LAD objective function, which is akin to Manski's maximum score objective function, is piecewise constant with numerous local solutions, leading many to consider finding a global optimizer an impossible task. The computational difficulty for solving the maximum score estimation problem is well-noted in the econometrics literature (e.g. see Manski and Thompson (1986), Pinkse (1993), Skouras (2003)).

Recent advances in numerical methods have made the scope of computing maximum score type estimators possible. Florios and Skouras (2008) propose a mixed integer optimization (MIO) based approach for the computation of maximum score estimators of Manski (1985), who consider the simpler two-choice special case of our problem. In particular, they show that Manski's binary choice maximum score estimation problem can be equivalently reformulated as a mixed integer linear programming problem (MILP). Thanks to the developments in MIO solution algorithms and fast computing environments, this reformulation enables exact computation of global solutions through modern efficient MIO solvers. Well-known numerical solvers such as CPLEX and Gurobi can be used to effectively solve the MIO problems. See e.g., Bertsimas and Weismantel (2005) and Conforti et al. (2014) for recent and comprehensive texts on the MIO methodology and applications.

There are also theoretical reservations for using the maximum score type estimator worth mentioning. While the estimators of Manski (1985) and Lee (1992) have been shown to be consistent under very weak conditions, they converge at a cube-root rate and have non-standard asymptotic distrubition. See Kim and Pollard (1990) and Seo and Otsu (2018). Inference for this type of estimators can be done by subsampling (Delgado et al. (2001)) or a model-based smoothed bootstrap (Patra et al. (2018)). As an alternative, Horowitz (1992) proposes a smoothed maximum score (SMS) estimation approach that employs a smooth approximation of the original maximum score objective function. The SMS estimator can converge at a rate arbitrarily close to the parametric rate yet that rate improvement requires additional assumptions on the smoothness of the density of underlying latent unobservables. In practice users also have to choose smoothing tuning parameters to implement the SMS estimator. Relatedly, smoothing induces additional bias that is difficult to correct especially for models with heteroskedasticity (see Kotlyarova and Zinde-Walsh (2009)). It is therefore unclear whether one would prefer SMS over MS estimators in finite samples even if one is willing to make more assumptions on the data generating process.

3 An empirical model and parameter of interest

Happiness economists are interested in using SWB data taken from two groups, say A and B, to draw conclusions on whether people in group A are happier than those in group B. Examples of groups include, gender, martial status, country, time etc. Previous analyses use the mean as a statistic to compare happiness across groups. We will use the median. We consider comparisons between only two groups for brevity. The arguments below are general and can be straightforwardly extended to any finite number of groups.

Suppose we observe $\left(Y^l, X^l\right)$ for $l = A, B$, where Y^l denotes reported happiness level taking values from $\mathcal{Y} = \{1, \ldots, J\}$ and X^l denotes a vector of observed covariates. We assume Y^l is derived from a threshold crossing model based on latent continuous happiness variable, H^l, s.t.:

$$\begin{aligned} Y^l &= j \times \mathbf{1}\left[\gamma^l_{j-1} < H^l \leq \gamma^l_j\right] \quad \text{for} \quad j = 1, \ldots, J, \\ H^l &= X^{l\top}\theta^l + U^l, \end{aligned} \tag{1}$$

for some strictly increasing real thresholds $\left\{\gamma^l_j\right\}_{j=1}^{J-1}$ with $\gamma^l_0 = -\infty, \gamma^l_J = +\infty$, θ is a vector of parameters, and U^l is an unobserved scalar accounting for other factors.

When we are dealing with subjective variables, such as happiness, $\left\{\gamma^l_j\right\}_{j=1}^{J-1}$ are unknown parameters that have to be estimated in the model along with θ^l. Some location and scale normalizations on θ^l and $\left\{\gamma^l_j\right\}_{j=1}^{J-1}$ will be necessary within each group. We will discuss more on this in Section 4. Since ordinal variables are scale free, we will also need to make some assumptions in order to compare latent happiness across groups. We suggest a pragmatic way to put happiness across groups on the same scale is to assume the thresholds in (1) are the same for $l = A, B$.[6,7] For notational simplicity suppose both groups contain the same set of covariates. Then we suppress the group index and pool the two threshold crossing models together in one unified framework:

$$\begin{aligned} Y &= j \times \mathbf{1}\left[\gamma_{j-1} < H \leq \gamma_j\right] \quad \text{for} \quad j = 1, \ldots, J, \\ H &= X^\top\pi^A + D \cdot X^\top\pi^B + U, \end{aligned} \tag{2}$$

where D is a dummy variable taking value 1 for group B and 0 otherwise, $\pi^A = \theta^A$ and $\pi^B = \theta^B - \theta^A$.[8] It is straightforward to allow covariates to differ across groups or to impose restrictions on the coefficients (e.g. equality of some parameters across groups to be the same). We emphasize here that using a common set of thresholds for both groups put H^A and H^B does not resolve the general non-identification of the mean ranking. Indeed, BL adopt this same framework in their empirical studies in order to reverse happiness results.

[6] This appears more appealing than estimating the median of H^l separately for each group and comparing them.

[7] In some other applications of categorical ordered data there may be possible information available on the thresholds. For example, they may be known to be equally spaced or even perfectly known (e.g. income reported in brackets).

[8] In practice, especially if the number of groups is large, one can a priori set some of the coefficients of the same covariates across groups to be the same.

We denote the parameter of interest for median comparison by,

$$\lambda(X) := Med(H|X, D=1) - Med(H|X, D=0). \tag{3}$$

By equivariance of the median, the sign of $\lambda(X)$ is the same for any increasing transformation of H.

Proposition 1. If $Med(H|X, D)$ is identified, then the sign of $\lambda(X)$ is identified.

In practice, we need to impose some assumptions on the distribution of $U|X, D$ in order to identify and estimate $\lambda(X)$.

Example 2: For a median comparison based on an ordered probit model, we would assume U in (2) satisfies $U|X, D \sim N(0, \sigma^2(X, D))$. Then $H|X, D \sim N\left(X^\top\pi^A + D \cdot X^\top\pi^B, \sigma^2(X, D)\right)$ and $\lambda(X) = X^\top\pi^B$.

If we set the support of Y in Example 2 to $\{1, 2, 3\}$, let γ_1 and γ_2 be 0 and 1 respectively, and reduce X to a constant then we have the same setup as Example 1. Using the interpretation of Example 1, in this case the sign π^B determines the median happiness ranking between men and women.

In practice $\lambda(X)$ can be identified and estimated under different assumptions. The simplest case would be to assume $U|X, D$ has a normal or logistic distribution. Then the conditional mean and median of H coincide, and we can use standard statistical softwares for estimating the mean of an ordered probit or logit to estimate the median. We state the relation between mean and median of the latent variable under symmetry as a proposition.

Proposition 2. If $U|X, D$ has a symmetric distribution at zero, then $E[H|X, D] = Med(H|X, D)$.

If the distribution of $U|X, D$ is not symmetric then we have to estimate the median directly. In the next section we will focus on estimating $Med(H|X, D)$ when the only assumption made on U is $Med(U|X, D) = 0$.

4 Estimating ordered response models using MILP

We consider the problem of estimating a general semiparametric median regression. There is no need to distinguish different types of covariates (denoting group) in this section. We revert to (1) and drop the group index, leading to:

$$Y = j \times \mathbf{1}[\gamma_{j-1} < H \leq \gamma_j] \quad \text{for} \quad j = 1, \ldots, J, \tag{4}$$

$$H = X^\top\theta + U. \tag{5}$$

We make the following assumptions.

Assumption I

(i) X is a $(P+1)-$dimensional vector that does not contain a constant term.

(ii) The first element of θ is 1.

(iii) $Med(U|X) = 0$.

Since we only observe outcomes from the events $\{[\gamma_{j-1} < H \leq \gamma_j] : j = 1, \ldots, J\}$, it is well-known that (θ, γ) can only be identified upto location and scale normalizations. I(i) imposes a location normalization as the regression intercept cannot be separately identified from the threshold parameters. I(ii) is a scale normalization, requiring that we know the effect X_1 has on the median to be positive (it can also be negative by setting θ to be -1). There are alternative ways to normalize the location and scale. For instance, one may set $\gamma_1 = 0$ and $\gamma_2 = 1$ as done in Example 1. In some applications, as we will show in Section 4, prior assumption on the sign in I(ii) may be reasonable due to economic intuition or other institutional knowledge. I(iii) imposes the zero median restriction. In particular, it follows from I(iii) that

$$\Pr[Y \leq j|X] > 0.5 \Longleftrightarrow \gamma_j > X^\top\theta, \tag{6}$$
$$\Pr[Y \leq j|X] = 0.5 \Longleftrightarrow \gamma_j = X^\top\theta, \tag{7}$$
$$\Pr[Y \leq j|X] < 0.5 \Longleftrightarrow \gamma_j < X^\top\theta. \tag{8}$$

The sign-matching relations above yield the conditional median for Y:

$$Med(Y|X) = \sum_{j=1}^{J} j \times \mathbf{1}\left[\gamma_{j-1} < X^\top\theta \leq \gamma_j\right].$$

Let $X = (X_1, \widetilde{X})$, where X_1 is a scalar random variable and $\widetilde{X}$ is the subvector of X that excludes the covariate X_1, and let $\theta = (1, \beta)$. We can then write $X^\top\theta = X_1 + \widetilde{X}^\top\beta$. Let $\Theta \subset \mathbb{R}^{P+J}$ denote the parameter space containing (β, γ), where P and J are the dimensions of β and γ respectively. We would like to estimate (β, γ). The estimator for β can then be used to estimate $Med(H|X)$.

We use (b, c) to denote a generic point of Θ. Given a random sample $(Y_i, X_i)_{i=1}^n$, our estimator of β and γ, $\widehat{\beta}$ and $\widehat{\gamma}$, solves the following LAD estimation problem:

$$\min_{(b,c)\in\Theta} \sum_{i=1}^{n} \left| Y_i - \sum_{j=1}^{J} j \times \mathbf{1}[c_{j-1} < X_{1i} + \widetilde{X}_i^\top b \leq c_j] \right|. \tag{9}$$

Because $\left|Y - \sum_{j=1}^{J} j \times \mathbf{1}\left[\gamma_{j-1} < X^\top\theta \leq \gamma_j\right]\right|$ can be equivalently expressed in the following form $|Y - J| + \sum_{j=1}^{J-1} [|Y - j| - |Y - j - 1|] \times \mathbf{1}\left[X^\top\theta \leq \gamma_j\right]$, the problem (9) above is therefore equivalent to the following minimization problem:

$$\min_{(b,c)\in\Theta} \sum_{i=1}^{n} \sum_{j=1}^{J-1} [|Y_i - j| - |Y_i - j - 1|] \times \mathbf{1}\left[X_{1i} + \widetilde{X}_i^\top b \leq c_j\right]. \tag{10}$$

As a result, we can reformulate the LAD problem in (9) into the following mixed integer linear programming (MILP) problem:

$$\min_{(b,c)\in\Theta,(d_{i,1},...,d_{i,J-1})_{i=1}^{n}} \sum\nolimits_{i=1}^{n} \sum\nolimits_{j=1}^{J-1} [|Y_i - j| - |Y_i - j - 1|] \times d_{i,j} \tag{11}$$

subject to

$$(d_{i,j} - 1)\, M_{i,j} \leq c_j - X_{1i} - \widetilde{X}_i^{\top} b < d_{i,j}(M_{i,j} + \delta),\ \ (i,j) \in \{1,...,n\} \times \{1,...,J-1\}, \tag{12}$$

$$c_j < c_{j+1},\ \ j \in \{1,...,J-2\}, \tag{13}$$

$$d_{i,j} \leq d_{i,j+1},\ \ (i,j) \in \{1,...,n\} \times \{1,...,J-2\}, \tag{14}$$

$$d_{i,j} \in \{0,1\},\ \ (i,j) \in \{1,...,n\} \times \{1,...,J-1\}, \tag{15}$$

where $\delta > 0$ is a user-chosen tolerance level (we use $\delta = 10^{-6}$ as in our numerical study), and

$$M_{i,j} \equiv \max_{(b,c)\in\Theta} \left| c_j - X_{1i} - \widetilde{X}_i^{\top} b \right|,\ \ (i,j) \in \{1,...n\} \times \{1,...,J-1\}. \tag{16}$$

Solving the constrained MILP problem (11) is equivalent to solving the minimization problem (10) and hence the LAD problem (9). To see this, take any $(b,c) \in \Theta$, the sign constraints (12) and the dichotomization constraints (15) ensure that $d_{i,j} = \mathbf{1}\left[X_{1i} + \widetilde{X}_i^{\top} b \leq c_j\right]$ for $(i,j) \in \{1,...n\} \times \{1,...,J-1\}$. We then enforce monotonicity of the threshold parameters through inequality constraints (13). Note that (12) and (13) together imply (14), which we explicitly impose so as to further tighten the MILP problem.

This equivalence enables us to employ the modern MIO solvers to exactly compute $(\widehat{\beta}, \widehat{\gamma})$. For the numerical implementation, note that the values $(M_{i,1}, ..., M_{i,J-1})_{i=1}^{n}$ in the inequality constraints (12) can be computed by formulating the maximization problem in (16) as linear programming problems, which can be efficiently solved by modern optimization solvers. Hence these values can be easily computed and stored as the input to the MILP problem (11).

5 Revisiting the Easterlin Paradox

Easterlin (1974, 1995, 2005) examined the relationship between happiness and income for many countries. He finds that despite the economic growth reported well-being stays stable over time. In contrast, there are strong evidence that within a given time period those with high income are happier than those with low income (see Dolan et al. (2008), for review). The Paradox leads to a widespread idea that increasing the income of all does not improve well-being of all, and it is the relative, not absolute, income that is important for individual well-being (see, e.g., Layard (2005)). This conclusion makes economic growth a debatable aim of public policies. The Paradox is, however, not without disagreements. Particularly, Stevenson and Wolfers (2008) suggest there is a strong positive dependence between GDP and well-being within countries and across time in all countries they consider apart from for the US. It is worth noting that Easterlin (1995) did not estimate happiness in a regression framework. He simply plotted the shares of the very happy people (out of very happy, pretty happy, not too happy) against time but did not find a positive trend despite rapid GDP growth. Easterlin (1995) used the GSS data from 1972 to 1991. On the other hand Stevenson and Wolfers (2008) estimated the ordered probit of happiness on year fixed effects and plotted them against income also finding a negative trend, using the GSS data from 1972 to 2006.

We use a subset of the GSS (from the years 1972-2006) data from Stevenson and Wolfers (2008); the latter is the same dataset that BL use in their illustration. The GSS has been regularly collecting information on attitudes and behaviors of American people since 1972. Among other questions, the survey asks about the respondent's happiness: *Taken all together, how would you say things are these days – would you say that you are very happy, pretty happy, or not too happy?* The survey was administered to a nationally representative sample of about 1,500 respondents each year between 1972 and 1993 (except 1979, 1981, 1995); around 3,000 respondents each second year between 1994 and 2004, and around 4,500 respondents in 2006[9]. We prepare the data using the same process described in Stevenson and Wolfers (2008), which involves using the same sample weights, omitting oversamples, and making adjustments for the change in question order in several waves.

We estimate the semiparametric ordered response model using MILP as described in Section 4. Solving an MILP problem remains computationally expensive for large-scale problems. To mitigate the computational cost, we select a random sample of 500 responders who answer the happiness question in the GSS bi-annually from the years 1974 to 1990, 1991, 1993, and bi-annually again from 1994 to 2006. The total size of our sample is 9500 from 19 different years. We standardize the income variable to have zero mean and unity variance. We then perform the median regression of reported happiness on this covariate along with age, age-squared, degree dummy, female dummy, marriage status dummy, and time fixed effects. There are 25 parameters[10] in the semiparametric model to be estimated including the two latent thresholds.

We estimate the parametric medians from ordered probit and logit models that allow for heteroskedasticity using the oglm command in STATA created by Williams (2010). The parametric models use an exponential function with a linear index for the skedastic function and either a normal or logistic error variable with mean zero and variance one. In contrast the semiparametric model allows for a general form of heteroskedasticity under Assumption I. For I(ii), we set the coefficient parameter associated with *income* to one. Note that the appropriate covariate that corresponds to I(ii) should have a known a priori effect on happiness as well having a rich support (see Horowitz

[9]Due to the ballot scheme design only half of the respondents were allocated the happiness question in years 1994-2004; in 2006 two-thirds of the respondents were inquired about their happiness.

[10]We have 23 parameters in $Med\,(H|X)$, see equation (5). After we have normalized the income coefficient to one, there are 5 coefficients corresponding to *age*, *age-squared*, *degree*, *female*, *marriage* and 18 time effects.

(2009, Chapter 4)). The covariate *income* is a suitable choice because it has quite a rich support[11,12] and it is a consensus in the happiness literature that it affects happiness positively once time (which is highly correlated with GDP growth) is controlled for. Table 1 gives the point estimates for all models. We divide the parametric estimates by their respective estimates of the coefficient on *income* in order to facilitate their comparison with the semiparametric estimates.

Table 1 shows that parametric and semiparametric estimates associated with socioeconomic factors are qualitatitvely the same and similar in magnitude. The semiparametric model assumes a priori that the income effect is positive. This coincides with the signs of the income effect estimated in the ordered probit and logit models. The other effects conform with the convention in the literature that a person with a degree and/or being a female and/or being married is happier than their respective counterparts. The effect on age suggests some convexity, which has been suggested in the happiness literature reflecting a midlife nadir (Cheng et al. (2017)). The time effects and the thresholds from the parametric and semiparametric models, however, differ. The parametric time effects are mixed in signs whereas the semiparametric ones are mostly positive and always higher than their parametric counterparts. This simply means the (time effect) intercept for the median regression of the semiparametric latent happiness is higher than the parametric counterparts in each of the years we consider. Relatedly, while there is an overlapping region including the origin, the range between the parametric thresholds cover larger negative region and less of the positive region compared to the semiparametric one. We attribute the estimated non-socioeconomic differences to the difference in the shape of the distribution of the additive error term. Nevertheless, the individual difference in time effects between models within a given year is not particularly relevant. We are more interested in how they compare across time in order to investigate the Easterlin Paradox.

There are two components to the Easterlin Paradox. One is that income is positively correlated with people's happiness within any time period. The other is that people's happiness does not increase as the nation becomes more prosperous over time. Figure 1 plots the time fixed effects from the semiparametric and ordered probit models against time, we include fitted lines through them and super-impose in the graph of how log-real per capita GDP grows with time. Figure 2 plots the time fixed effects from the semiparametric and ordered probit models against log-real per capita GDP directly with fitted lines. Figures 3 and 4 provide analogous plots of the semiparametric estimate against the ordered logit.

The figures above unambiguously support the empirical story of Easterlin Paradox along the

[11]The GSS first collects income data in 12 bands, computes their mid-points and debases them. So *income* in our sample can take 228 possible values.

[12]The standard sufficient conditions for identification of maximum score type estimators require one of the covariates has support on $\mathbb{R}$ conditional on the explanatory variables. The full support condition is not necessary. It can be reduced to bounded or even finite support as long as the support is rich enough to guarantee that the sign-matching conditions in equations (6) to (8) only hold at the data generating parameter value. See Manski (1988) and Horowitz (2009, Chapter 4).

	Parametric estimates				Semiparametric
	Probit	(scaled)	Logit	(scaled)	estimates
income	0.118	1.000	0.192	1.000	1.000
age	-0.027	-0.226	-0.045	-0.236	-0.227
age_sq	0.0003	0.0023	0.0005	0.0024	0.0024
degree	0.144	1.227	0.241	1.255	1.446
female	0.113	0.965	0.199	1.034	1.223
married	0.505	4.296	0.881	4.581	5.541
year dummies					
1974	-0.027	-0.230	-0.045	-0.233	1.923
1976	0.003	0.029	0.004	0.021	2.018
1978	0.001	0.007	0.007	0.037	1.410
1980	-0.035	-0.301	-0.061	-0.318	1.660
1982	-0.040	-0.344	-0.073	-0.378	1.304
1984	-0.002	-0.014	0.000	-0.002	1.611
1986	0.049	0.417	0.081	0.420	1.109
1988	0.112	0.955	0.202	1.049	2.894
1990	0.053	0.448	0.093	0.486	1.965
1991	-0.092	-0.785	-0.173	-0.900	0.553
1993	-0.103	-0.872	-0.181	-0.940	0.739
1994	-0.136	-1.158	-0.241	-1.256	0.517
1996	-0.086	-0.729	-0.149	-0.776	0.901
1998	-0.113	-0.963	-0.210	-1.091	1.253
2000	0.084	0.711	0.135	0.700	1.482
2002	-0.074	-0.632	-0.143	-0.745	0.893
2004	-0.035	-0.296	-0.063	-0.326	1.812
2006	-0.091	-0.771	-0.157	-0.815	-0.334
cut1	-1.470	-12.505	-2.556	-13.300	-6.329
cut2	0.266	2.261	0.397	2.066	4.578
N obs.	9500	9500	9500	9500	9500

income is a variable for family income in constant dollars (base = 1986), the variable is normalized by subtracting the mean and dividing by standard error; *age* is the respondent's age; *age_sq* stands for age squared; *degree*, *female*, *married* are dummy variables that take the value of 1 respectively if the respondent obtained bachelor or graduate degree, is female, is currently married.

Table 1: Parametric and semiparametric estimates

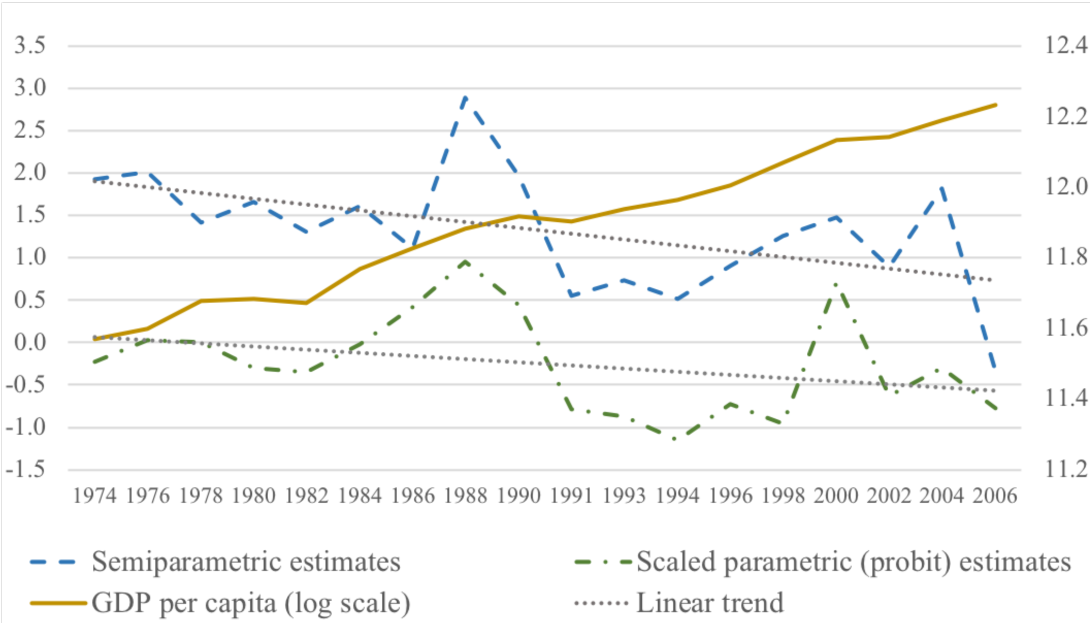

Figure 1: Parametric (probit) and semiparametric estimates. Median happiness (left axis) and Logarithm of GDP per capita (right axis)

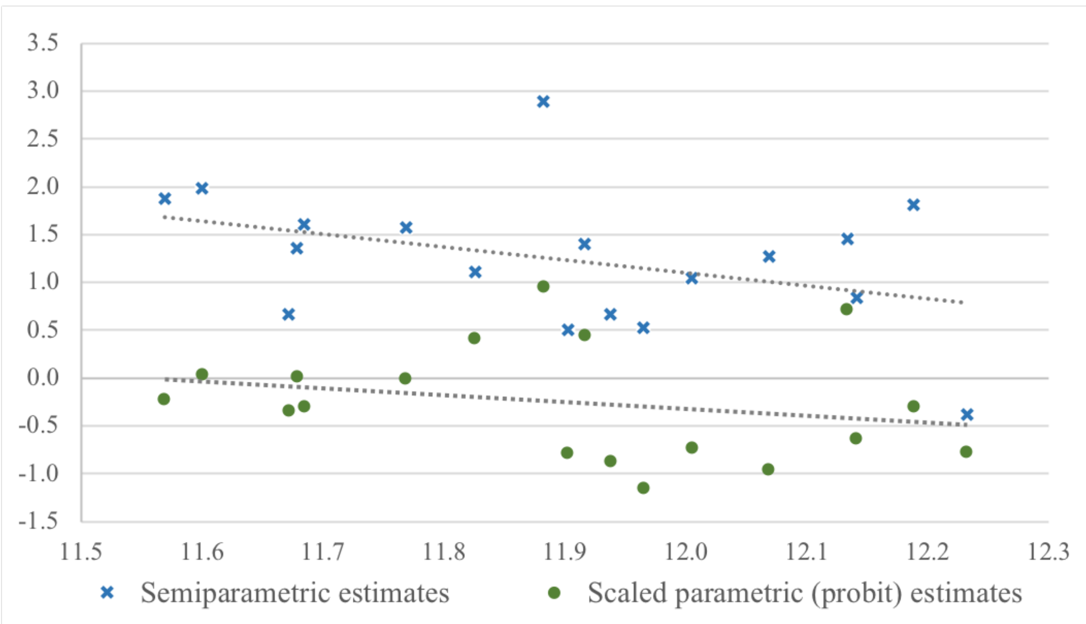

Figure 2: Parametric (probit) and semiparametric estimates. Median happiness (vertical axis) and Logarithm of GDP per capita (horizontal axis)

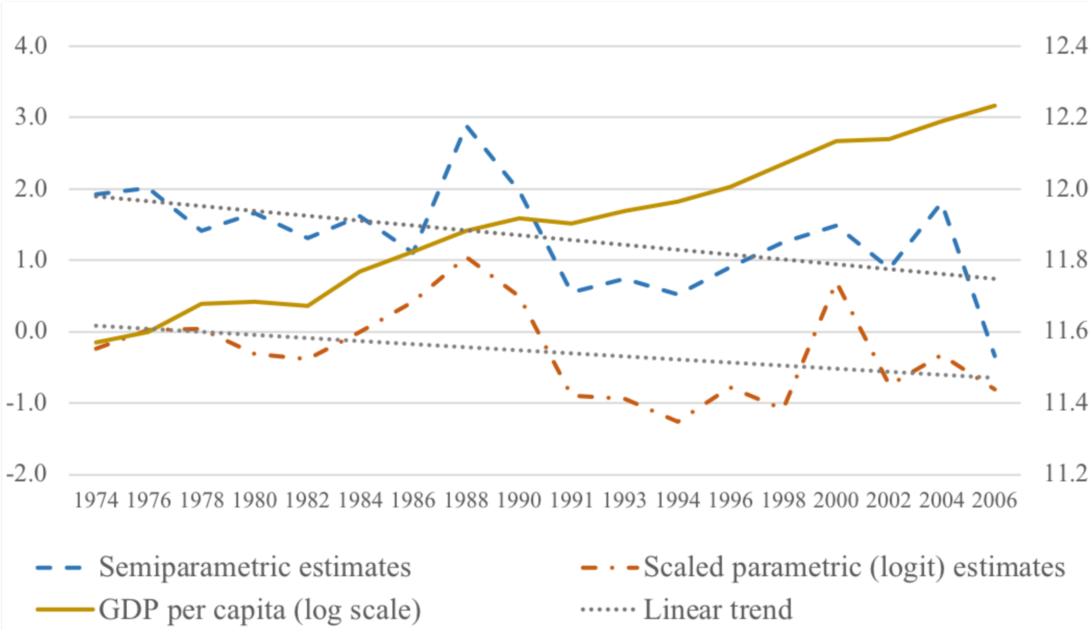

Figure 3: Parametric (logit) and semiparametric estimates. Median happiness (left axis) and Logarithm of GDP per capita (right axis)

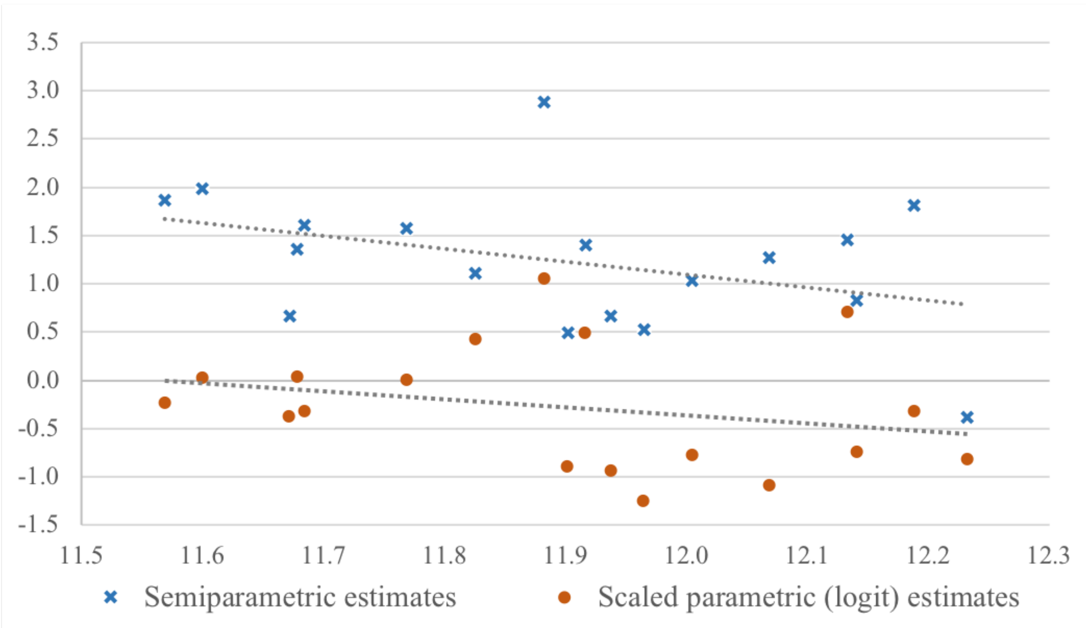

Figure 4: Parametric (logit) and semiparametric estimates. Median happiness (vertical axis) and Logarithm of GDP per capita (horizontal axis)

timespan dimension. We can look at the median regression for each of the 19 years separately for evidence of the intra-year component of the Paradox. We do not estimate the income effect of the semiparametric model as it is assumed to be positive a priori under Assumption I(ii). But we find the coefficients for *income* are positive for all years for both ordered probit and logit models. In addition, while we caution against directly comparing individual parameters across regressions, we find the effects of the parametric and semiparametric median estimates on *age*, *age-squared*, *degree*, *female*, *marriage status* uniformly agree for most of the years under consideration. We therefore conclude that all models concur on supporting Easterlin Paradox for the US with GSS data.

6 Conclusion

Categorical ordered data, such as happiness scales or other life satisfaction measures, are prevalent and can be potentially very useful in helping us to understand subjective concepts in social sciences. Empirical studies in some research fields, such as happiness economics, rely exclusively on this type of discrete ordinal data. From a theoretical point of view, it is clear that ordinal data should not be treated like cardinal variables. A perhaps more subtle point is that ranking of means between ordinal variables is not identified even when an ordered response model is used unless there is first order stochastic dominance. There are serious practical consequences when the ordinal nature of data is not respected. Schröder and Yitzhaki (2017) and Bond and Lang (forthcoming, BL) emphatically illustrate this point by showing the necessary stochastic dominance conditions typically do not hold in subjective well-being data used in the happiness literature and conclusions drawn in huge body of empirical work can be systematically overturned by applying simple transformations.

Without any rebuttal, much of the mean comparison results that form the basis of empirical studies in the happiness literature appear to have no empirical content. By exploiting the fact that the mean and median of symmetric distributions are identical, we show prior results from parametric ordered response models can still be useful if we focus on the median instead of the mean. In particular, making use of BL's empirical illustrations, it appears the leading studies in the happiness literature remain valid simply by changing the notion of mean ranking to median ranking.

Generally we suggest a median comparison of latent happiness between groups can be studied systematically in a single equation by pooling the relevant groups together. This approach puts the latent happiness across groups on the same scale to facilitate comparison in a transparent manner. The median can be estimated in different ways depending on what is assumed on the distribution of the latent variable. If one is willing to impose symmetry of the distribution and specify the shape parametrically, such as normal or logistic distributions, then the median (together with the mean) can be readily obtained from standard statistical softwares. Semiparametric median regression can also be estimated directly without relying on any distributional assumption beyond imposing a median restriction based on the estimator of Lee (1992).

Estimating the semiparametric median under such a weak condition is a numerically challenging task. In this paper we re-interpret Lee's estimator as a solution to an MIO problem. A suitable MIO solver can then be used to estimate the median. In our empirical study, we use the Gurobi Optimizer[13] to investigate the Easterlin Paradox for the US using the GSS data. We find the Paradox exists empirically in both parametric models, the ordered probit and order logit, and the semiparametric model. In particular we find the median happiness does not trend upwards with time or as GDP rises in all three cases suggesting that the well-known Paradox is not an artefact from symmetry or other parametric distributional assumptions.

Finally, we want to emphasize that our advocacy of the median regression analysis is not exclusive for analyzing subjective well-being data. The arguments and results in our papers are useful for all social scientists and other researchers who wish to analyze categorical ordered data. We also expect inference on semiparametric median regression and larger-scale estimation will become feasible as computing hardwares and solver algorithms for mixed integer optimization problems continue to progress.

References

Agresti, A. (1999). Modelling ordered categorical data: recent advances and future challenges. *Statistics in Medicine*, 18(17-18):2191–2207.

Ananth, C. (1997). Regression models for ordinal responses: a review of methods and applications. *International Journal of Epidemiology*, 26(6):1323–1333.

Bertsimas, D. and Weismantel, R. (2005). *Optimization Over Integers.* Dynamic Ideas, Belmont, Mass.

Bond, T. N. and Lang, K. (forthcoming). The Sad Truth About Happiness Scales. *Journal of Political Economy.*

Boyce, C. J., Wood, A. M., Banks, J., Clark, A. E., and Brown, G. D. A. (2013). Money, well-being, and loss aversion: does an income loss have a greater effect on well-being than an equivalent income gain? *Psychological Science*, 24(12):2557–2562.

Carneiro, P., Hansen, K. T., and Heckman, J. J. (2003). Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice. *International Economic Review*, 44(2):361–422.

[13]The Gurobi Optimizer is a numerical solver for various mathematical programs including mixed integer linear programming problems. It is freely available for academic purposes. See http://www.gurobi.com/.

Chen, S. and Khan, S. (2003). Rates of convergence for estimating regression coefficients in heteroskedastic discrete response models. *Journal of Econometrics*, 117(2):245–278.

Cheng, T. C., Powdthavee, N., and Oswald, A. J. (2017). Longitudinal Evidence for a Midlife Nadir in Human Well-being: Results from Four Data Sets. *The Economic Journal*, 127(599):126–142.

Clark, A. E. (2003). Unemployment as a Social Norm: Psychological Evidence from Panel Data. *Journal of Labor Economics*, 21(2):289–322.

Clark, A. E., Frijters, P., and Shields, M. A. (2008). Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles. *Journal of Economic Literature*, 46(1):95–144.

Clark, A. E. and Oswald, A. J. (1996). Satisfaction and comparison income. *Journal of Public Economics*, 61(3):359–381.

Conforti, M., Cornuejols, G., and Zambelli, G. (2014). *Integer Programming*. Graduate Texts in Mathematics. Springer International Publishing.

Cunha, F., Heckman, J. J., and Navarro, S. (2007). The Identification And Economic Content Of Ordered Choice Models With Stochastic Thresholds. *International Economic Review*, 48(4):1273–1309.

Delgado, M. A., Rodriguez-Poo, J. M., and Wolf, M. (2001). Subsampling inference in cube root asymptotics with an application to Manski's maximum score estimator. *Economics Letters*, 73(2):241–250.

Di Tella, R., MacCulloch, R. J., and Oswald, A. J. (2001). Preferences over Inflation and Unemployment: Evidence from Surveys of Happiness. *American Economic Review*, 91(1):335–341.

Dolan, P., Peasgood, T., and White, M. (2008). 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*, 29(1):94–122.

Easterlin, R. A. (1974). Does Economic Growth Improve the Human Lot? Some Empirical Evidence. In David, P. A. and Reder, M. W., editors, *Nations and Households in Economic Growth*, pages 89–125. Academic Press.

Easterlin, R. A. (1995). Will raising the incomes of all increase the happiness of all? *Journal of Economic Behavior & Organization*, 27(1):35–47.

Easterlin, R. A. (2005). Feeding the Illusion of Growth and Happiness: A Reply to Hagerty and Veenhoven. *Social Indicators Research*, 74(3):429–443.

Ferrer-i-Carbonell, A. (2005). Income and well-being: an empirical analysis of the comparison income effect. *Journal of Public Economics*, 89(5-6):997–1019.

Ferrer-i-Carbonell, A. and Frijters, P. (2004). How Important is Methodology for the estimates of the determinants of Happiness? *The Economic Journal*, 114(497):641–659.

Florios, K. and Skouras, S. (2008). Exact computation of max weighted score estimators. *Journal of Econometrics*, 146(1):86–91.

Gebers, M. (1998). Exploratory Multivariable Analyses of California Driver Record Accident Rates. *Transportation Research Record: Journal of the Transportation Research Board*, 1635:72–80.

Greene, W. H. and Hensher, D. A. (2010). *Modeling Ordered Choices: A Primer*. Cambridge University Press.

Gruber, J. and Mullainathan, S. (2006). Do Cigarette Taxes Make Smokers Happier? In Ng, Y. K. and Ho, L. S., editors, *Happiness and Public Policy: Theory, Case Studies and Implications*, pages 109–146. Palgrave Macmillan UK, London.

Honore, B. E. and Lewbel, A. (2002). Semiparametric Binary Choice Panel Data Models Without Strictly Exogeneous Regressors. *Econometrica*, 70(5):2053–2063.

Horowitz, J. L. (1992). A Smoothed Maximum Score Estimator for the Binary Response Model. *Econometrica*, 60(3):505–531.

Horowitz, J. L. (2009). *Semiparametric and Nonparametric Methods in Econometrics*. Springer Series in Statistics. Springer-Verlag, New York.

Kahneman, D. and Krueger, A. B. (2006). Developments in the Measurement of Subjective Well-Being. *Journal of Economic Perspectives*, 20(1):3–24.

Kahneman, D., Krueger, A. B., Schkade, D., Schwarz, N., and Stone, A. (2004). Toward National Well-Being Accounts. *American Economic Review*, 94(2):429–434.

Kim, J. and Pollard, D. (1990). Cube Root Asymptotics. *The Annals of Statistics*, 18(1):191–219.

Kotlyarova, Y. and Zinde-Walsh, V. (2009). Robust Estimation in Binary Choice Models. *Communications in Statistics - Theory and Methods*, 39(2):266–279.

Layard, R. (2005). *Happiness: Lessons from a New Science*. Penguin Press, London/New York.

Lee, M. J. (1992). Median regression for ordered discrete response. *Journal of Econometrics*, 51(1-2):59–77.

Lewbel, A. (1997). Constructing Instruments for Regressions with Measurement Error when no Additional Data are Available, with an Application to Patents and R&D. *Econometrica*, 65(5):1201–1214.

Lewbel, A. (2000). Semiparametric qualitative response model estimation with unknown heteroscedasticity or instrumental variables. *Journal of Econometrics*, 97(1):145–177.

Lewbel, A. and Schennach, S. M. (2007). A simple ordered data estimator for inverse density weighted functions. *Journal of Econometrics*, 136:189–211.

Liddell, T. M. and Kruschke, J. K. (2018). Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology*, 79:328–348.

Major, S. (2012). Timing Is Everything: Economic Sanctions, Regime Type, and Domestic Instability. *International Interactions*, 38(1):79–110.

Manski, C. F. (1975). Maximum score estimation of the stochastic utility model of choice. *Journal of Econometrics*, 3(3):205–228.

Manski, C. F. (1985). Semiparametric analysis of discrete response: Asymptotic properties of the maximum score estimator. *Journal of Econometrics*, 27(3):313–333.

Manski, C. F. (1988). Identification of Binary Response Models. *Journal of the American Statistical Association*, 83(403):729–738.

Manski, C. F. and Thompson, T. S. (1986). Operational characteristics of maximum score estimation. *Journal of Econometrics*, 32(1):85–108.

McCullagh, P. (1980). Regression Models for Ordinal Data. *Journal of the Royal Statistical Society. Series B (Methodological)*, 42(2):109–142.

McFadden, D. L. (1974). Conditional Logit Analysis of Qualitative Choice Behavior. In *Frontiers in Econometrics*. Wiley, New York.

McKelvey, R. D. and Zavoina, W. (1975). A statistical model for the analysis of ordinal level dependent variables. *The Journal of Mathematical Sociology*, 4(1):103–120.

Patra, R. K., Seijo, E., and Sen, B. (2018). A consistent bootstrap procedure for the maximum score estimator. *Journal of Econometrics*, 205(2):488–507.

Pinkse, C. A. P. (1993). On the computation of semiparametric estimates in limited dependent variable models. *Journal of Econometrics*, 58(1):185–205.

Powdthavee, N. (2007). Are there Geographical Variations in the Psychological Cost of Unemployment in South Africa? *Social Indicators Research: An International and Interdisciplinary Journal for Quality-of-Life Measurement*, 80(3):629–652.

Schröder, C. and Yitzhaki, S. (2017). Revisiting the evidence for cardinal treatment of ordinal variables. *European Economic Review*, 92:337–358.

Seo, M. H. and Otsu, T. (2018). Local M-estimation with discontinuous criterion for dependent and limited observations. *The Annals of Statistics*, 46(1):344–369.

Skouras, S. (2003). An algorithm for computing estimators that optimize step functions. *Computational Statistics & Data Analysis*, 42(3):349–361.

Stevens, S. S. (1946). On the Theory of Scales of Measurement. *Science*, 103(2684):677–680.

Stevenson, B. and Wolfers, J. (2008). Happiness Inequality in the United States. *The Journal of Legal Studies*, 37(S2):S33–S79.

Stevenson, B. and Wolfers, J. (2013). Subjective Well-Being and Income: Is There Any Evidence of Satiation? *American Economic Review*, 103(3):598–604.

Stiglitz, J., Sen, A. K., and Fitoussi, J. P. (2009). The measurement of economic performance and social progress revisited: Reflections and Overview. Technical Report 2009-33, Sciences Po.

Stutzer, A. and Frey, B. S. (2013). Recent Developments in the Economics of Happiness: A Selective Overview. In *Recent Developments in the Economics of Happiness*. Cheltenham.

Williams, R. (2010). Fitting heterogeneous choice models with oglm. *The Stata Journal*, 10(4):540–567.