Measuring Health in HRS dataset using Bayesian Approach

Peyman Firouzi Naeim
Department of Economics
Andrew Young School of Policy Studies

This Version: April 2017
[Latest Version]

Abstract

This paper provides a measure of the physical health for the near retirement people based on the binary responses of individuals to the sets of questions, regarding the objective aspects of their health. Using three sets of questions: Activities of daily living (ADL), Lower Body Mobility (LBM) and Upper Body Agility (UBA), I employ the Item Response Theory and using Bayesian approach estimate the individual distribution of health for each individual in the Health and Retirement Survey (HRS) sample of 2008. The Analysis of 17,217 sets of individual responses shows that LBM and UBA affect the combined health index (Full Index) more strongly and ADL has a weak effect on the health index. While the correlation between the full measure of functional limitation and LBM and UBA is 0.92 and 0.84 respectively, the correlation between functional limitation and ADL is only 0.5.

1 Introduction

Measuring health, as simple as it seems that it is, can be a subtle task. The word *"Health,"* in contrast to its everyday life usage, does not have a solid definition. Relying on the self reported health status can lead to integrating different and perhaps contradictory dimensions of health. For example, while having cancer and difficulties in walking a short distances both may result in the poor health status, reported by individuals, they affect different aspects of person's life and aggregating these two into one measure of health may harm the research goals. On the other hand, self reported health is subjective and like every other subjective measure makes the task of inter-personal comparison, infeasible. This lack of a determined and solid definition is the main problem that should be tackled in every research that its focus is health outcome and more importantly those with the goal of constructing health index. In such studies, the measurement is often done by mapping the responses of individuals in a health questionnaire to the objective questions, to the uni or multidimensional measure of health.

In lacking the universal model of health that can provide such a transformation, and facing so many different dimensions of health, any effort in producing health index might be ad-hoc. Considering that different questions are targeting different subsets of health, and their relative importance, one instantly notices that weights that should be attached to the different questions are undefined. Furthermore, even if such weights exist, depending on the usage of the health index, the weights should be different. Think about two different questions such as:

1. Because of a health problem do you have any difficulty with walking one block?
2. Have you had surgery on your heart?

Both of these two questions are measuring health, but the first one is more accurate in measuring of the way health affects everyday activities (feeling of a person) while the second question tries to measure health as it affects the individual's length of life. It is accepted in the literature to divide the questions into two major subsets of "functional limitations" and "health conditions." Functional limitations set consists of the questions which are looking more for the aspects of health that affect the personal experience in everyday life. Health conditions, on the other hand, consists of the questions which are more relevant in measuring the length of the individual's life. By using this division, the researcher will be able to focus on the more narrow definition of health.

Poterba et al. (2013) use factor analysis to construct a general health index. They use responses to the different health questions containing both functional limitations and medical conditions and obtain the first principal component. The first principal component is the weighted average of the health indicators (questions), where the weights are chosen to maximize the proportion of the explained variance of the health indicators. The problem with the factor analysis approach is the linearity assumption that it imposes to the problem. Linearity requires the continuous response, however most of the responses in health related questions are designed to be binary/categorical.

Soldo et al. (2006) use the Item Response Theory (IRT) approach to solve the continuity issue. However, they only use "functional limitations". Item response models, fit a non-linear probability function (Probit, Logit) to each categorical health question. For each health question, these fitted functions are defined using two parameters: sensitivity (discrimination) and scale (difficulty). The health of an individual then governs the way he responds to each question.

Gustman and Steinmeier (2014) employ Item response theory to make the health index using only functional limitations and then incorporate the constructed health index in a health model, in which health index is a function of medical conditions. They also suggest that depending on the goal of the project and the way study utilizes the health measure researcher can choose which sets of questions are required to be used in the construction of the health measure. Although this is an ad-hoc approach, in the lack of the health model which can guide us in both weights and integration of life/death and good/bad feeling, this ad-hoc approach seems reasonable.

In this paper, I use functionality limitations (Total of 17 questions) to construct the health index. Functionality Limitations set contains three different sets of questions, namely, Activities of daily living (ADL), Lower Body Mobility (LBM) and Upper Body Agility (UBA) set.

Functionality limitations can be important in an individual's decisions toward retirement. To model the decision making process of individuals, researcher needs information about the distribution of his health as well as the first moment. By having full distributional information researcher can incorporate risk and uncertainty in an agent's decision-making process. Following Bayesian approach and estimating the posterior distribution for the health measure, I can provide such distributional information. None of the Soldo et al. (2006), Poterba et al. (2013) and Gustman and Steinmeier (2014) provide the distributional information of the constructed health index. I use the Bayesian approach in estimating the individual health index, and as a result, I will be able to provide full distributional knowledge of each person's health.

Another distinct feature of the current study is the way indices are constructed. Previous studies aggregate a collection of binary items and form a new categorical item which can be used in estimating the required parameters of a multinomial logistic regression. As an example, if the person's response to two binary items is yes, then a new categorical item which is produced using those two binary items has the domain of $\{0, 1, 2\}$, and in our example, the value will be $1 + 1 = 2$. However, this method has its deficiencies. As it is shown in Table 1, Some questions are a subset of other questions (Like G003 and G001). If an individual's response to one of those is yes (G001), then the other should be yes as well (G003). One might treat these variables as if they are ordered, but what is done in the literature is an aggregation of all binary items (Regardless of their dependencies) to get one unique value for each subgroup (LBM in this case). This approach puts the same weight on all the categories of health (LBM, ADL, and UBA) and disregards the importance of a category which might be incorporated into the questionnaire by design when for example they ask seven questions on LBM and five on ADL. Making categorical variables out of these subsets of items and then constructing the general health index imposes the same weight on all of the three subsets. In this paper, I incorporate the binary variables (items) directly into the process of constructing the health index. This method gives the model a higher degree of flexibility and does not impose any further assumption about the weights.

2 Estimation

This section explains the Item Response Theory and Bayesian estimation method required to reveal the unobserved health index.

Item Response Theory (IRT)

Item Response theory assumes there is an underlying individual specific parameter that explains the responses of each individual to a group of related questions. This underlying parameter (θ) which is assumed to be distributed normally in the population, can be person's knowledge in a field of study which will be reviled by an exam which is designed in a multiple choice question format or can be person's health which will be uncovered by responses of individual to questions related to individual's health. For the observed health measure i (item) probability for falling in the response group j is (Lord (1952)):

$$\Pr(Z_i = j) = F\left(\beta_{i,j} - \alpha_i \theta\right) - F(\beta_{i,j-1} - \alpha_i \theta)$$

where, α_i measures the sensitivity (discrimination) of an item i and β_i measures the scale (difficulty) of item i. Letting $Y = [y_{m,i}]_{n \times k}$ to be the vector of the binary responses for all individuals and all the items. Then:

$$y_{m,i} = \begin{cases} 1, \; person\ m\ response\ to\ item\ i\ is\ positive\ (yes, feel\ pain) \\ 0, person\ m\ response\ to\ item\ i\ is\ negative\ (no,\ no\ pain) \end{cases}$$

Probability of observing $y_{m,i} = 1$ can be modeled using Normal CDF:

$$P\left(y_{m,i} = 1 \,\middle|\, \theta_m, \alpha_i, \beta_i\right) = g\left(\alpha_i\theta_m - \beta_i\right) = \int_{-\infty}^{\alpha_i\theta_m-\beta_i} \frac{1}{(2\pi)^{\frac{1}{2}}} e^{-\frac{t^2}{2}}.dt$$

Bayesian Approach

Following J Albert (1992) Assume $y_{m,i}$ is independent Bernoulli random variable with probability of success (Lord (1953))

$$p_{m,i} = \Phi\left(\alpha_i\theta_m - \beta_i\right)$$

Where Φ is Standard Normal CDF. Assuming that $\alpha_i < 0$ shows a person with better health has a smaller probability of answering yes to item i. Defining augmented continuous variable $Z_{m,i}$ to be:

$Z_{m,i} \sim N(\alpha_i\theta_m - \beta_i{,}1)$ and $y_{m,i} = \begin{cases} 1\ if\ Z_{m,i} > 0 \\ 0\ if\ Z_{m,i} \leq 0 \end{cases}$

The joint distribution is: $p\left(\theta_m, \alpha_i, \beta_i, Z_{m,i} \,\middle|\, y_{m,i}\right) \propto f\left(y_{m,i} \,\middle|\, Z_{m,i}, \theta_m, \alpha_i, \beta_i\right) p\left(Z_{m,i}, \theta_m, \alpha_i, \beta_i\right) = f\left(y_{m,i} \,\middle|\, Z_{m,i}\right) p\left(Z_{m,i} \,\middle|\, \theta_m, \alpha_i, \beta_i\right) p\left(\theta_m\right) p(\alpha_i, \beta_i)$

This model then has $n + 2k$ parameters to estimate. I employ Gibbs sampling procedure to estimate parameters.

Gibbs Sampling

I am interested in simulating from the joint posterior distribution of $Z_{m,i}$, θ_m and item characteristics α_i, β_i. Given the parameters θ, α, β the augmented variables Z is independent variable with distribution a such as:

$$Z_{m,i}|\theta, \alpha, \beta, y \;\; \sim \begin{cases} N_{0,\infty}\left(\alpha_i\theta_m - \beta_i, 1\right) & if\ y_{m,i} = 1 \\ N_{-\infty,0}\left(\alpha_i\theta_m - \beta_i, 1\right) & if\ y_{m,i} = 0 \end{cases}$$

Conditional on Z, α, β, and assuming the inter-personal independency of the health measure, θ has independent posterior density such as:

$$\pi\left(\theta_m \,\middle|\, Z, \alpha, \beta, y\right) = C.\ \phi(\theta_m; 0, 1) \prod_{i=1}^{k} \phi\left(Z_{m,i}; \alpha_i \theta_m - \beta_i, 1\right)$$

where C is constant and using the assumption that $\theta_m \sim N(0,1)$. It follows that

$$L(\theta_m) \propto N\left(\frac{\sum_i \alpha_i \left(Z_{m,i} + \beta_i\right)}{\sum_i \alpha_i^2}, \frac{1}{\sum_i \alpha_i^2}\right)$$

Knowing that θ_m has standard normal prior:

$$\theta_m | Z, \alpha, \beta, y \sim\ N\left(\frac{\sum_i \left(Z_{m,i} + \beta_i\right) \alpha_i + \frac{\mu}{\sigma^2}}{\frac{1}{\sigma^2} + \Sigma_i \alpha_i^2} \quad , \quad \frac{1}{\frac{1}{\sigma^2} + \Sigma_i \alpha_i^2}\right),$$

where $\sigma = 1, \;\; \mu = 0$ by our prior (Standard Normal distribution of health for each individual). The conditional posterior distribution of (α_i, β_i), assuming that items are uncorrelated, can be written as:

$$\pi\left(\alpha_i, \beta_i \,\middle|\, \theta, Z, y\right) = C \prod_{i=1}^{n} \phi\left(Z_{m,i}; \alpha_i \theta_m - \beta_i, 1\right) I\left(\alpha_j > 0\right)$$

Next, knowing $Z_{m,i} = \alpha_i \theta_m - \beta_i + \epsilon_i$, let $X = [\overrightarrow{\theta}\ \ -\overrightarrow{1}]$. The $Z_i = [\theta - 1]\left[\alpha_i\ \beta_i\right]' + \epsilon_i$ is $n \times 1$ vector and it follows that conditional probability of the last two sets of parameters can be written as:

$$(\alpha_i, \beta_i) | \theta, Z_i, Y_i \ \sim N_2\left((X'X)^{-1} X' Z_i \ \ , \ (X'X)^{-1} I\left(\alpha_i < 0\right)\right)$$

Finally, knaowing the joint distribution of α_i and β_i, I use the first moments as the estimated values of the parameters.

3 Simulation

I simulate three different data sets using the various discrimination and scale parameters from Soldo et al. (2006). Defining four different items and 500 individuals for each data set. The simulation process can be summarized as:

1. For each individual, draw the health (θ_i) from standard normal distribution;

2. Using $\alpha and,\ \beta$ calculate the index for each individual.

3. Calculate the $Prob(x < Index_i) = f(Index_i)$

4. Generate random number $\{0,1\}$ using binomial random generator, having probability of success of $f(Index_i)$

Three simulated data sets differed in $f(Index_i)$ function. The function f governs the distribution of people in the society with respect to their health index. By changing the health ranking, I analyze the sensitivity of the estimates to the different distributions of health. The estimation method I employ is based on the assumption that, the health distribution is standard normal. As a result when I use the same distribution in simulating the data, I have correctly specified model while it is not the case when I do not use standard normal distribution in simulation. To see how prone is the analysis to the misspecification, I use a slightly skewed and highly skewed generalized extreme value distribution in second and third simulations. The parameters for Generalized Extreme Value distributions are shown in Table 2. Figure 2 also shows the probability distribution functions for all three distributions of health.

3.1 Correctly Specified Model

Assuming the normality of the health distribution, I can estimate α and β parameters for each item. The estimation results are shown in table 3.
Simultaneously I estimate the health index for each individual. Figure 3 shows the relationship between estimated and true value of individual health index.

3.2 Misspecified Model

For the Misspecified model, the estimation results are provided in Table 4. The model is unable to recover the true discrimination parameters. However, it is performing better in estimating the slopes.
Figure 4 shows the relationship between true and estimated individual health index (rankings). Comparing to the correctly specified model, the health measures estimates are more categorized.

3.3 Highly Misspecified Model

As it is shown in Table 5, none of the parameters can be estimated in the highly misspecified model. The model is also weak in estimating the health indices. The standard errors are big, and the model cannot recover the true parameters.

Table 6 shows the comparison between Three estimations (correctly specified, misspecified and Highly Misspecified), by providing ratios of health indices which are in the one standard deviation of the true individual health index. I call this ratio the success ratio. The second row in Table 6 shows the average standard error of the estimated health indices. While the first column gives us success ratio, it is not enough for comparison purpose, since bigger standard errors produce higher success ratio, and the trade offs between these two should be taken into account.

As it appears from Table 6, even knowing the misspecified model does not provide a good estimation of item parameters, the index measure which is produced using both models, gives the almost same success ratio.

4 Data

I employ 2008 wave of the Health and Retirement Study data set, containing 17,217 observations, after dropping observations with the missing items. Seventeen binary items, separated into the three groups of Lower Body Mobility (LBM), Upper Body Agility (UBA) and Activities of Daily Life (ADL). While LBM is important in how mobile a person is, UBA is important for job-related activities. ADL contains cognitive abilities necessary for everyday life. Table 7 provides the summary of each question and the subgroup they belong to.

5 Results

I estimate the health index using three separated item subgroups(LBM, UBA, and ADL). I use items in each subgroup (LBM, UBA, and ADL), and construct health index only using them. The correlation between the estimated individual health indices using these three subgroups, can be used as a guideline for combining these three groups of items in order to construct a global health index.

LBM, ADL, and UBA respectively contain seven, five and five binary response items. After estimating each of the three groups, the correlation between ADL and other two groups is weak as we expected[1]. Table 8 provides the correlation structure between different health Indices. By the survey design, lower values of the health index means a

[1]Cognitive difficulty is only remotely related to the physical difficulties and hence the correlation between ADL which contains cognitive difficulties and the other two models which are more related to the physical difficulties is weak.

healthier individual. I leave this order to be consistent with the literature.

The distribution of health among individuals is shown for all three subgroups in figure 6.

The health index for ADL is more concentrated around -0.75, which means there are more people who recognized as "healthy" using ADL measure. The two other physical measures (UBA and LBM) are more alike.

Next, to be able to compare the results with other works, I need to merge all the different subgroups. Note that the current analysis does not suggest it since three different health Indices do not match closely which means ADL is not measuring the same aspect of health as UBA and LBM and cannot be merged into one measure.

Table 9 shows that, LBM and UBA affect the full index, far more than ADL. While correlation between ADL and full index is 0.5, the correlation between UBA and full index is considerably higher at 0.84. Correlation is even stronger between LBM and full index (0.92). The distribution of the index which is constructed using full dataset (Full Index) is shown alongside other three Indices. The Full Index has a bimodal distribution which confirms the other studies of health measurement[2]. Figure 7 shows the densities of different health Indices. The full index has a bi-modal distribution, but the mode on the right-hand side is rather flat.

Next, I compare the results with other works in the literature. We chose Gustman and Steinmeier (2014) for this purpose. Results are shown in Figure 8. While Gustman and Steinmeier (2014) use almost five times bigger data set that is used in current study ,they use ordered probit maximum likelihood estimator instead of Probit Bayesian estimate which is employed in this paper. The items sets are different as well. Gustman and Steinmeier (2014) use pain and self-reported health as well as ADL, UBA, and LBM to construct the full health index.

Gustman and Steinmeier (2014) predict a relatively less healthy population in Lower Body Mobility and Upper Body Mobility while their full index estimate of the health measure has a uni-modal distribution(Except a less apparent mode in very unhealthy part of the population) instead of bi-modal distribution I estimate in this paper.

Distribution of Individual Health

One of the advantages of the Bayesian analysis is the ability to provide the individual health distribution. It might be a concern that assuming a standard normal distribution for individual health, the estimated distribution of the individual health might be

[2] Gustman and Steinmeier (2014) report that health index in the U.S. is bi-modal.

the same "standard normal" distribution that is fed into the Bayesian algorithm as the prior distribution. Figure 9 shows three different individual health distributions which are drawn from the pool of the individuals health distributions which were estimated in the simulation section, under the correctly specified model. Using a correctly specified model that assumes the standard normal distribution for individual health and distribution of health in the society, three distribution of individual health, are neither same nor standard normal. It stresses out the importance of providing the full distribution of health when it is required for health investment decisions under risk, in which skewness in the distribution of health can play an important role. The importance of the full distributional knowledge is even more important in the presence of the non-linear models where Jensen's inequality distorts the expected value of the non-linear function from the given non-linear transformation of the expected value.

6 Conclusion

This paper provides a measure of the physical health for the near retirement people based on the binary responses of individuals to the sets of questions, regarding the objective aspects of their health. find that LBM and UBA are behaving more consistently compared with the ADL. While the correlation between the full measure of functional limitation and LBM and UBA is .92 and .84 respectively, the correlation between functional limitation and ADL is only .5. This leads to the conclusion that ADL targets different dimension of the physical health rather than LBM and UBA and it does not contribute to the physical health as LBM and UBA do. In particular, utilizing the variables of the LBM suffices to construct an objective measure of physical health in situations where there are not enough information about the objective responses to the functional limitation questions.

A Bayesian approach which is employed in this paper provides a whole distribution of individual's health. From the researcher's point of view, this means the ability to analyze the behaviors which involve decision making under uncertainty. The retirement decision is one of the decisions which requires such distributional information and using Bayesian approach for estimating health index one can analyze the health investment decisions more precisely.

The analysis of the different indices and the correlation structure among them suggests multi-dimensional full health index. This guides us to the more elaborated Bayesian approach in which the optimal number of dimensions (Factors) that is needed for analyzing the health index can itself be estimated.

References

Albert, J. (1992). Bayesian estimation of normal ogive item response curves using Gibbs sampling. *Journal of educational statistics*, 17(3):251–269.

Gustman, A. and T. Steinmeier (2014). Integrating retirement models: Understanding household retirement decisions. In Factors Affecting Worker Well-being: The Impact of Change in the Labor Market. *Emerald Group Publishing Limited*, Pp. 79–112.

Lord, F. (1952). The relation of the reliability of multiple-choice tests to the distribution of item difficulties. *Psychometrika*, 17(2):181–194.

Lord, F. (1953). On the statistical treatment of football numbers. *Scaling*, P. 402.

Poterba, J., S. Venti, and D. Wise (2013). Health, education, and the postretirement evolution of household assets. *Journal of Human Capital*, 7(4):297–339.

Soldo, B., O. Mitchell, R. Tfaily, and J. McCabe (2006). Cross-cohort differences in health on the verge of retirement. *National Bureau of Economic Research*, w12762.

Table 1: Binary Responses

Variable Name	Question	Response	
		Yes	No
G001	Because of a health problem do you have any difficulty with walking several blocks?	X	
G002	(Because of a health problem do you have any difficulty) with running or jogging about a mile?		X
G003	(Because of a health problem do you have any difficulty) with walking one block?	X	
G004	(Because of a health problem do you have any difficulty) with sitting for about two hours?		X
G005	(Because of a health problem do you have any difficulty) with getting up from a chair after sitting for long periods?	X	
G006	(Because of a health problem do you have any difficulty) with climbing several flights of stairs without resting?	X	
G007	(Because of a health problem do you have any difficulty) with climbing one flight of stairs without resting?		X
Categorical Item		4	

Example of Binary Response to Lower Body Mobility (LBM) sets of items and constructing a categorical item using them.

Table 2: Simulation Parameters

Parameter	Misspecified	Highly Misspecified
k: tail index (shape) parameter :	0	1
σ: scale parameter :	1	5
μ: threshold (location) parameter :	1	5

Characteristics of distribution of individual health index in the three synthetic state of the world.

Table 3: Simulation Results (Correctly Specified)

Health Index	α	$\widehat{\alpha}$	$SE(\widehat{\alpha})$	β	$\widehat{\beta}$	$SE(\widehat{\beta})$
Item = 1	1.67	1.38	0.19	-1.12	-0.97	0.13
Item = 2	2.98	2.53	0.32	-1.68	-1.41	0.20
Item = 3	3.17	2.89	0.34	-1.83	-1.53	0.16
Item = 4	2.03	2.10	0.21	-1.23	-1.20	0.14

α (Slope) and β (discrimination) parameters estimation result for Correctly model

Table 4: Simulation Results (Misspecified)

Health Index	α	$\widehat{\alpha}$	$SE(\widehat{\alpha})$	β	$\widehat{\beta}$	$SE(\widehat{\beta})$
Item = 1	1.67	1.69	0.19	-1.12	0.41	0.13
Item = 2	2.98	1.71	0.20	-1.68	-0.10	0.12
Item = 3	3.17	2.30	0.33	-1.83	-0.06	0.15
Item = 4	2.03	2.37	0.29	-1.23	0.16	0.14

α (Slope) and β (discrimination) parameters estimation result for misspecified model

Table 5: Simulation Results (Highly Misspecified)

Health Index	α	$\widehat{\alpha}$	$SE(\widehat{\alpha})$	β	$\widehat{\beta}$	$SE(\widehat{\beta})$
Item = 1	1.67	0.89	0.21	-1.12	1.98	0.22
Item = 2	2.98	0.96	0.20	-1.68	1.63	0.19
Item = 3	3.17	0.66	0.15	-1.83	1.41	0.13
Item = 4	2.03	0.88	0.19	-1.23	1.88	0.20

α (Slope) and β (discrimination) parameters estimation result for highly misspecified model

Table 6: Sensitivity Analysis

	Correctly Specified Model	Misspecified Model	Highly Misspecified Model
Success Ratio	0.838	0.80	0.8220
Standard Error	0.4949	0.47	0.74

Ratio of Health Indices which lay in one standard error distance to true health index.

Table 7: HRS Questions for the objective evaluation of the physical health

Variable	Question
	Lower Body Mobility
G001	Because of a health problem do you have any difficulty with walking several blocks?
G002	(Because of a health problem do you have any difficulty) with running or jogging about a mile?
G003	(Because of a health problem do you have any difficulty) with walking one block?
G004	(Because of a health problem do you have any difficulty) with sitting for about two hours?
G005	(Because of a health problem do you have any difficulty) with getting up from a chair after sitting for long periods?
G006	(Because of a health problem do you have any difficulty) with climbing several flights of stairs without resting?
G007	(Because of a health problem do you have any difficulty) with climbing one flight of stairs without resting?
	Upper Body Agility
G008	(Because of a health problem do you have any difficulty) with stooping, kneeling, or crouching?
G009	(Because of a health problem do you have any difficulty) with reaching or extending your arms above shoulder level?
G010	(Because of a health problem do you have any difficulty) with pulling or pushing large objects like a living room chair?
G011	(Because of a health problem do you have any difficulty) with lifting or carrying weights over 10 pounds, like a heavy bag of groceries?
G012	(Because of a health problem do you have any difficulty) with picking up a dime from a table?
	Activities of daily living (ADL)
G021	(Because of a health or memory problem do you have any difficulty with) bathing or showering?
G023	(Because of a health or memory problem do you have any difficulty with) eating, such as cutting up your food?
G025	(Because of a health or memory problem do you have any difficulty with) getting in or out of bed?
G040	Because of a health or memory problem, do you have any difficulty using a map to figure out how to get around in a strange place?
G041	(Because of a health or memory problem, do you have) any difficulty preparing a hot meal?

Variables which are used in estimation and their definitions.

Table 8: Inter-Models correlation of the health measure

	ADL	**LBM**	**UBA**
ADL	1.00		
LBM	0.36	1.00	
UBA	0.37	0.65	1.00

Correlation Between Different Sub-Models.

Table 9: Prediction Power Comparison

	ADL	**LBM**	**UBA**	**Full**
Full	.5	0.92	0.84	1

Correlation Between The index that is Constructed using full dataset and each of sub- Indices.

Slope and Scale in IRF

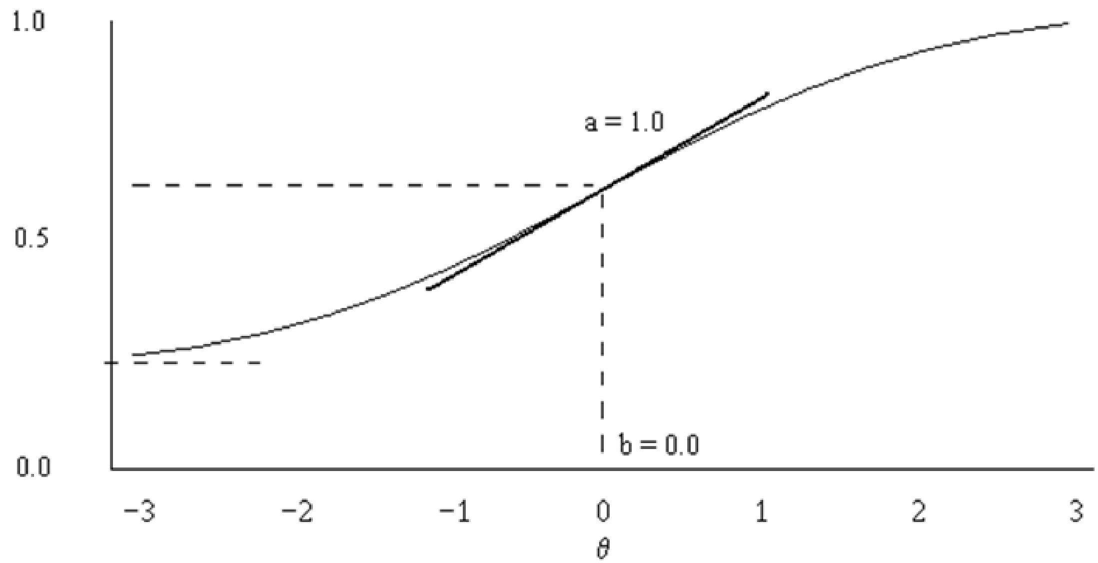

Figure 1: Item Response Theory and the fitted function. A is slope parameter, and b is scale.

Distribution of the three specifications of the model

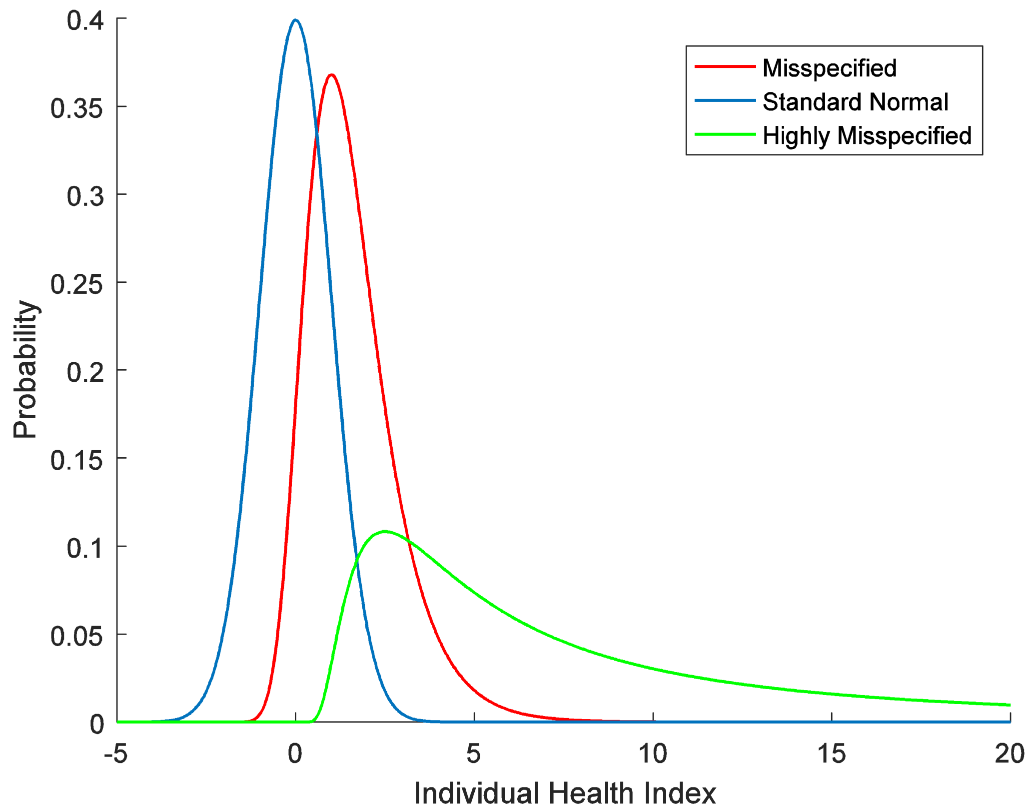

Figure 2: Probability Density Function for Correctly Specified, Misspecified and Highly Misspecified models.

Data versus Prediction: Correctly Specified Model

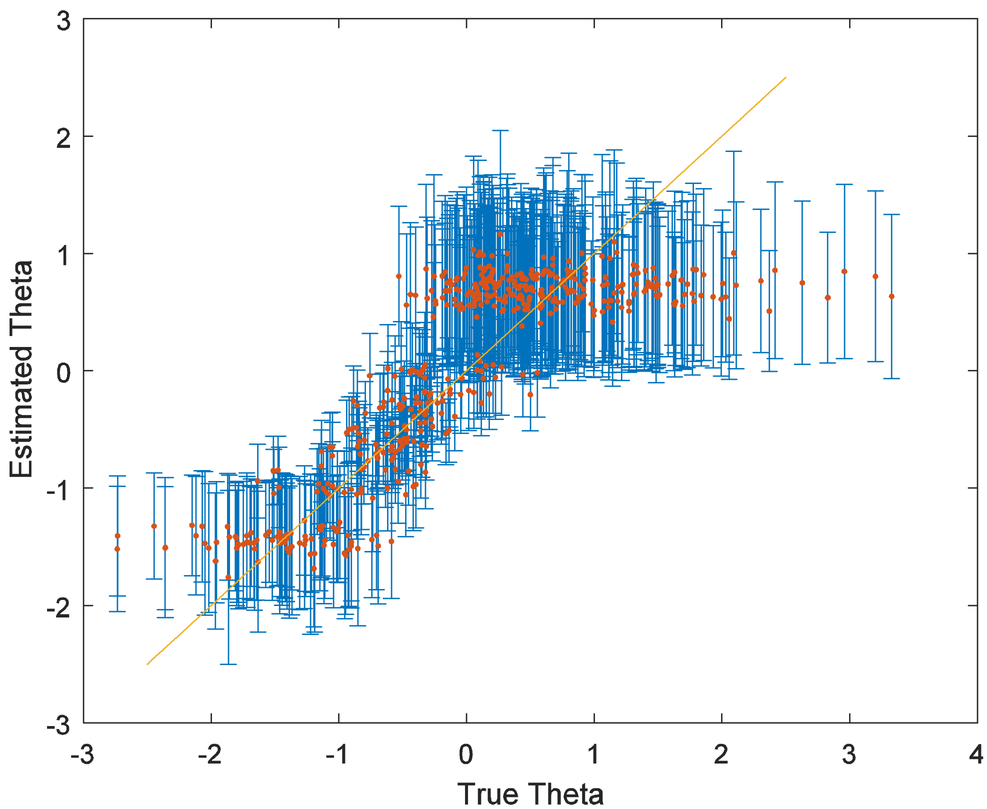

Figure 3: Relationship between True and estimated individual health index for correctly specified model.

Data versus Prediction: Misspecified Model

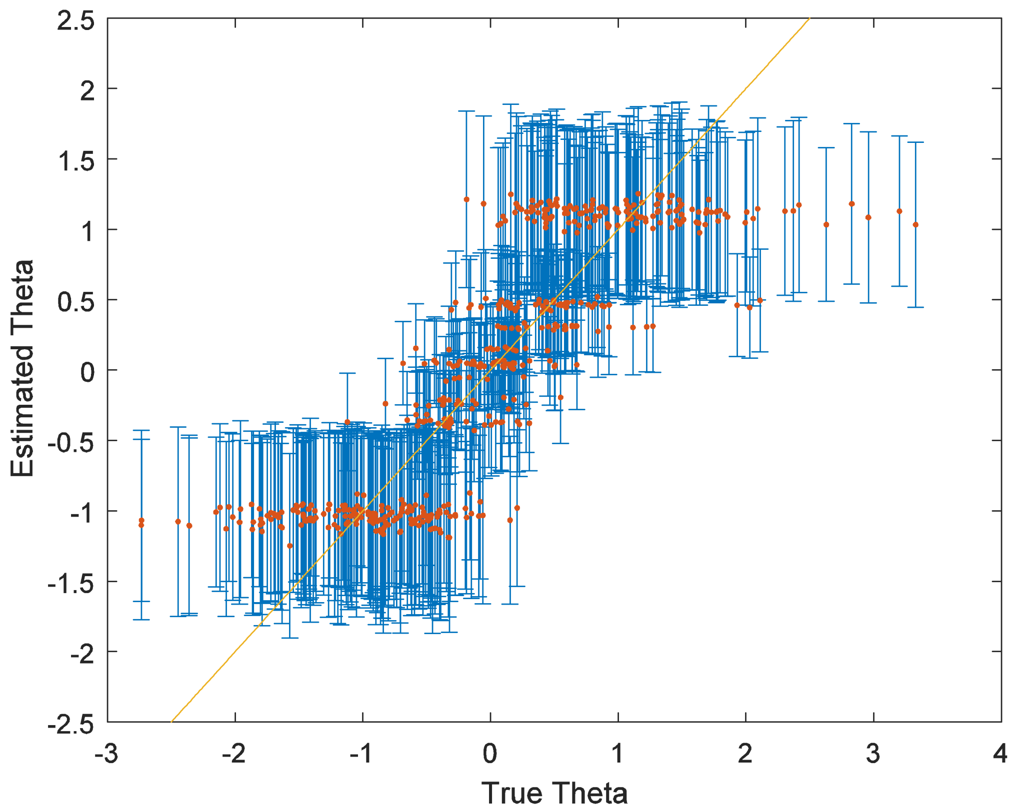

Figure 4: Relationship between True and estimated individual health index for Misspecified Model.

Data versus Prediction: Highly Misspecified Model

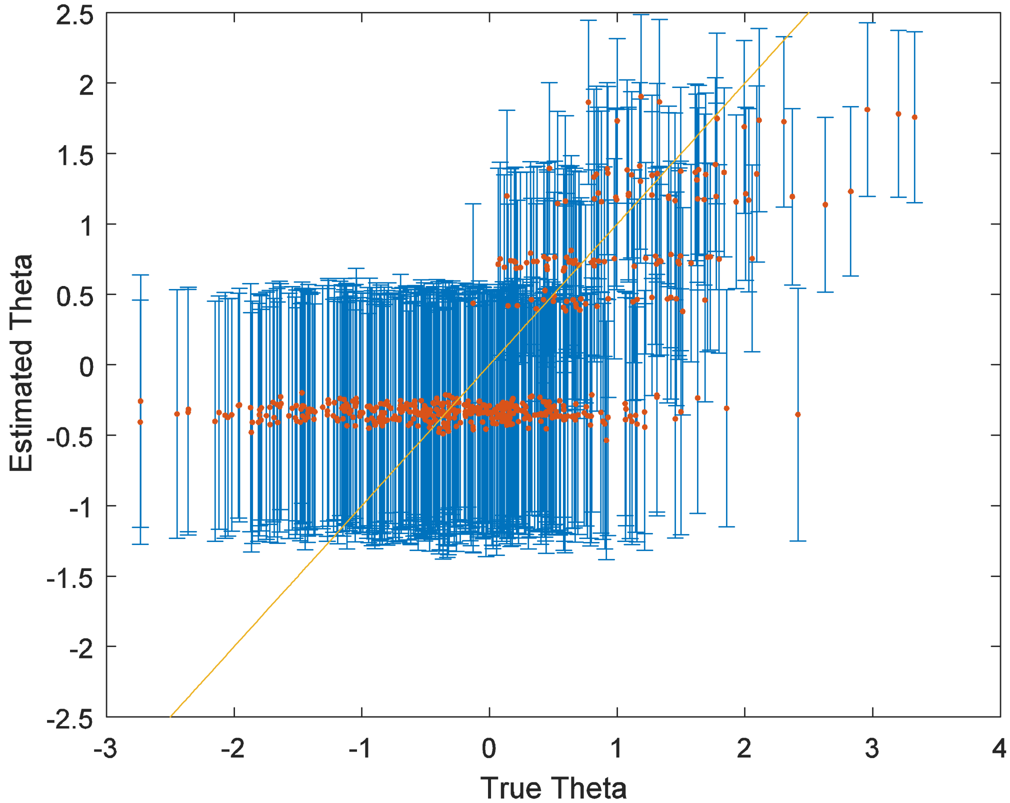

Figure 5: Relationship between True and estimated individual health index for the Highly Misspecified Model.

Probability Density for the Constructed Measures of Health

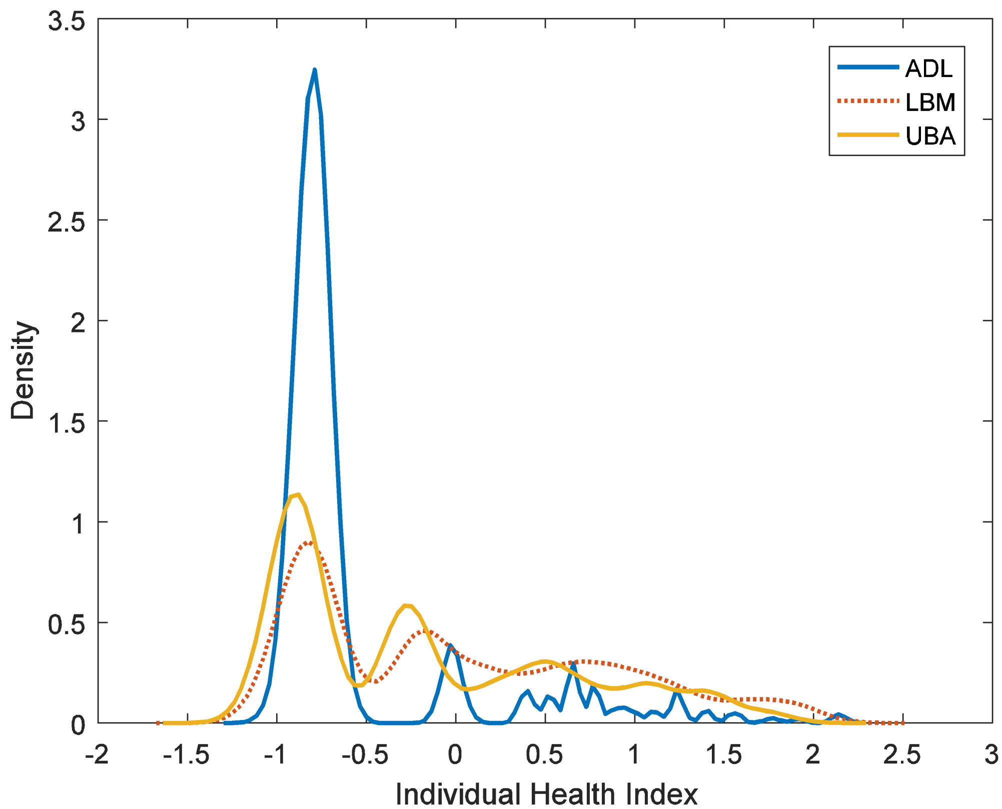

Figure 6: Distribution of Health using three different health measure which is constructed using different subsets of items. (Lower values mean Healthier)

Comparison with Gustman and Steinmeier (2014)

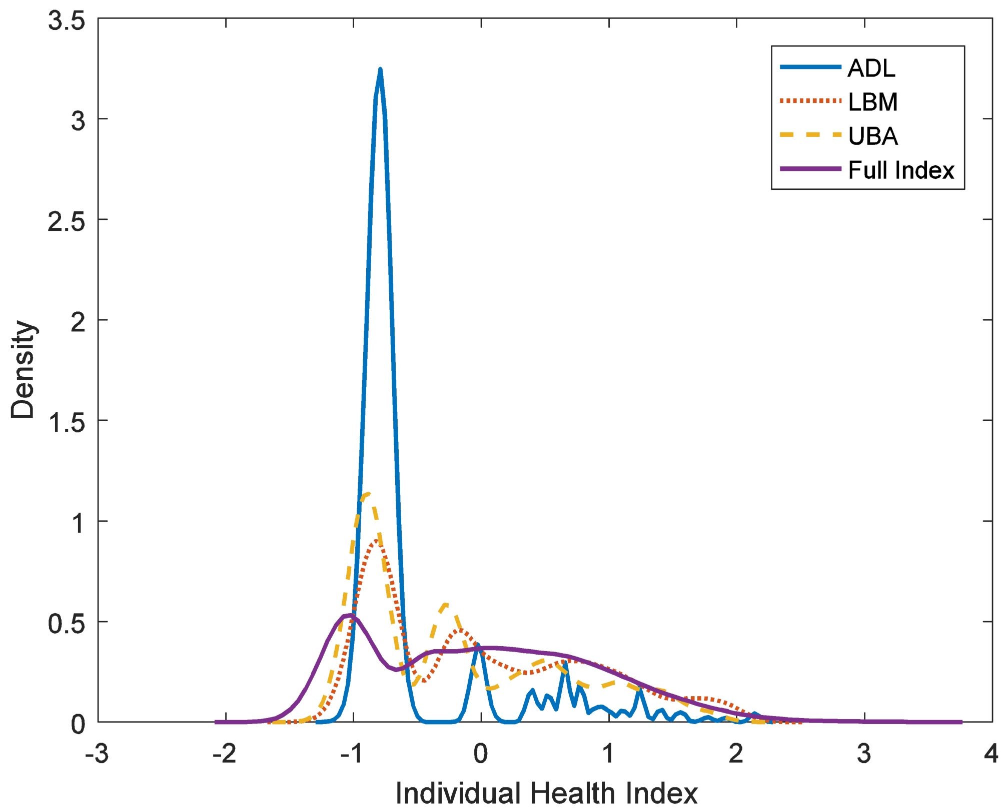

Figure 7: Comparison between the distribution of health using all the information set and three subgroups. (Lower values mean Healthier)

Individual distribution of the constructed health status.

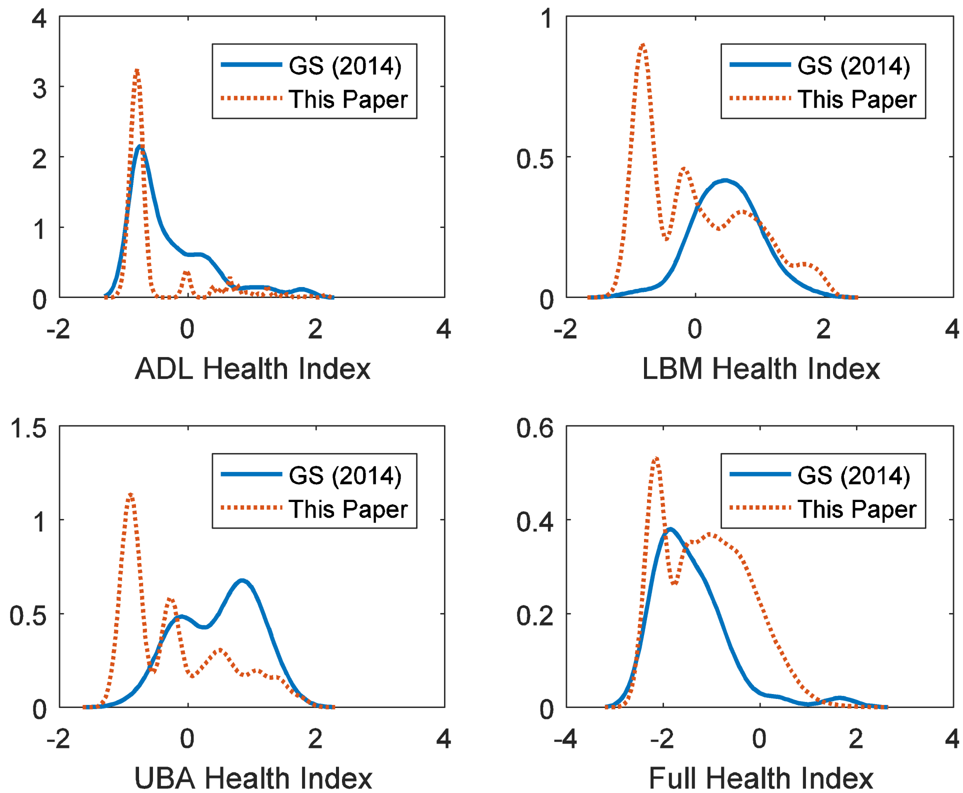

Figure 8: Comparing Gustman and Steinmeier (2014) and this work in distribution of constructed Health Indices. (Lower values mean Healthier)

Data versus Simulation: Selected Variables

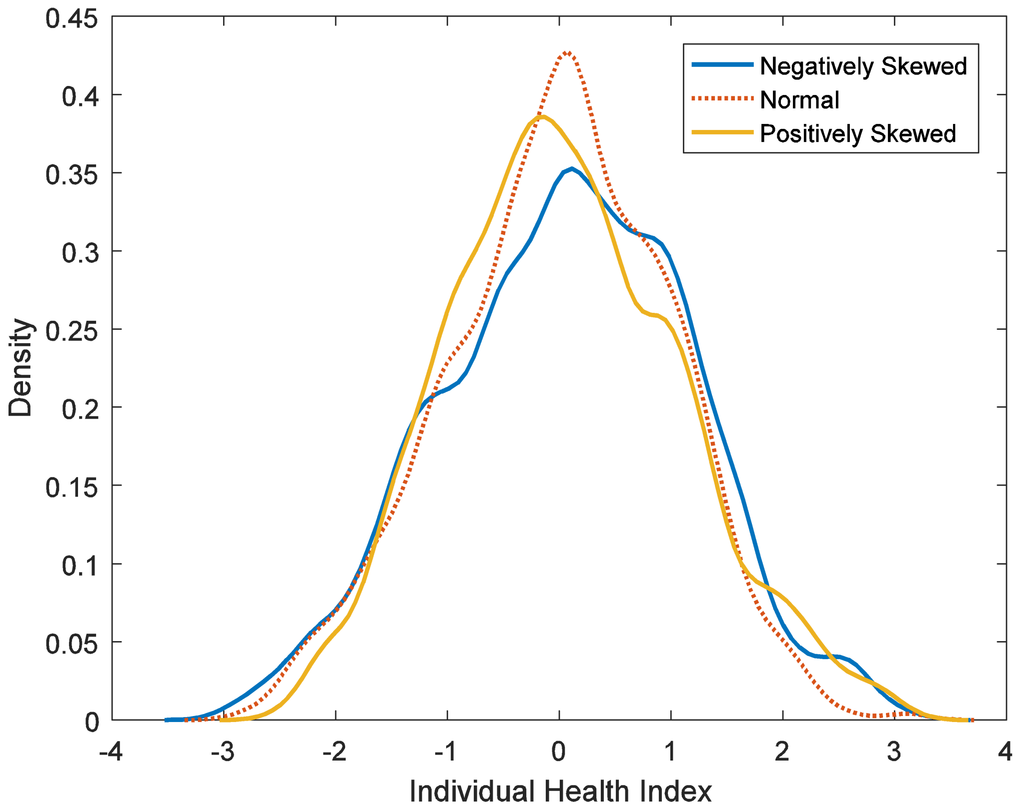

Figure 9: Health Index constructed for three different individuals using correctly specified model and simulated data.