Efficiency Analysis and the Measurement of Multidimensional Well-Being

Joseph Deutsch
Bar-Ilan University

Jacques Silber
Bar-Ilan University, LISER

Discussant: Flaviana Palmisano
University of Rome LUMSA and University of Luxembourg

34th IARIW General Conference, Dresden
August 21-27, 2016

Introduction

- Need to go 'beyond GDP' to measure economic performance and social progress. Give more prominence to the distribution of income, consumption and wealth; and broaden income measures to non-market activities (Stiglitz et al., 2009).

- Relevance of measuring well-being, a complex concept as already emphasized by Sen (1985):
  - There are many dimensions of well-being
  - Not possible to summarized them through a measure of income or wealth

Aims of the paper

- Derive measures of overall well-being for three West Asia countries (Armenia, Azerbaijan, Georgia) for which enough data are available
- Compare the findings concerning overall well-being with those based on its narrow view, one whose focus is only on material well-being
- Isolate the determinants of overall and material well-being.

Methods

- Correspondence analysis, to derive a synthetic index for each domain of well-being
- Efficiency analysis, the stochastic production frontier approach, to determine overall well-being at the individual level
- Regression analysis to explain overall and material well-being
- Shapley decomposition to determine the specific contributions of the various explanatory variables of the regression to its R-square

Correspondence analysis

- Special case of principal component analysis, to be applied in the case of categorical data
- Allows to analyse simple two-way (or multi-way) contingency tables where some measure of correspondence is assumed to exist between the rows and columns, where the two entries are qualitative characteristics of a population
- Reduces a multiplicity of variables to a smaller number of orthogonal linear combinations, preserving the original structure of the data
- Allows to obtain a graphical display of row and column points in biplots, which helps in discovering some structural relationships that may exist between the variables and the observations
- CA indicates how the variables are related and not simply whether there is such a link.

Correspondence analysis

- In this paper, first define different domains of well-being, in each domain several variables are assumed to characterize well-being in the domain.
- CA is applied separately to each domain and the first factor in each domain is then assumed to summarize the features of well-being in this domain and used in the second stage of the analysis.
- The amount of variance explained by each subsequent component decreases: the first component accounts for the largest proportion of the total variability in the data, the second component for the next largest amount not accounted for by the first component and so on for higher order components.
- Since we are looking for an indicator expressing as much as possible the latent information conveyed by the data, only the first principal component extracted as well-being indicator is retained

Stochastic production frontier approach

- On the basis of the inputs derived by CA in each domain of well-being an efficiency analysis is implemented and an output score (degree of well-being) attributed to each individual.
- The (first) factors derived separately from CA for each domain are considered as inputs in the production of a latent variable reflecting the overall degree of well-being of the individual.
- Capture the standard idea of the maximum possible output that can be obtained from a given set of inputs
- Test if individuals are different in the efficiency with which they convert personal resources into current levels of well being
- The SFA approach allows to distinguish between production inputs and efficiency/inefficiency factors and to disentangle distances from the efficient frontier between those due to systematic components and those due to noise.

Stochastic production frontier approach

- The SFA, the maximum output level for a given input set, is assumed to be stochastic in order to capture exogenous shocks beyond the control of individuals.

- Since all individuals, starting from the same set of inputs, are not able to produce the same frontier output, an additional error term is introduced, to represent technical inefficiency that, in turn, is under the control of individuals

Stochastic production frontier approach

$$Y_i = x_i\beta + (v_i - u_i),\ \forall i = 1, ...N \tag{1}$$

- Y_i: outcome of individual i;
- x_i: aggregated inputs (first factors) derived from the CA;
- β: vector of unknown parameters;
- $v_i - u_i$: unobserved random noise;
- v_i: random variables following the assumption of normally distributed error terms;
- u_i: non-negative random variables, assumed to capture the effects of technical inefficiency in production and to be independently distributed as truncations at zero.
- the mean of this truncated normal distribution is a function of systematic variables that can influence the efficiency of an individual: $m_i = z_i\delta + \epsilon_i$, where z_i is vector of variables which may have an effect on the production function of individuals and δ is a vector of parameters to be estimated.

The measure of individual well-being

The measure of individual well-being is obtained calculating the technical efficiency of the above production function:

$$d_i = e^{(z_i\delta - \epsilon_i)} = e^{(-u_i)} \tag{2}$$

Explaining individual well-being

$$d_i = \alpha + \omega_i \beta + u_i \tag{3}$$

- ω_i is a vector of determinants of individuals well-being
- u_i is the normally distributed error term

Data

- The Caucasus Barometer Survey: reliable, comparable data on household composition, knowledge, social and political attitudes, and practices across the South Caucasus
- Armenia, Azerbaijan, and Georgia
- Domains of well-being: participation and trust in the public sphere, social relations, health and material well-being
- Which period? Last survey year?

Descriptive statistics

	Armenia	Azerbaijan	Georgia	Three countries together
Urban	0.3676	0.3682	0.2411	0.3235
Rural	0.3594	0.3182	0.4222	0.3687
Male	0.3347	0.4621	0.4061	0.3991
Age	46.1646	42.1826	47.9633	45.5649
Square of age	2427.9973	2024.1265	2617.1242	2369.5209
Education	12.0302	11.4432	13.1617	12.2451
Islam	0.0069	0.9970	0.0688	0.3343
Married	0.2586	0.6576	0.2912	0.3932
Household size	3.8464	4.3136	3.5872	3.8999
Square of household size	18.3690	21.7576	16.2057	18.6578
Knowledge of Russian	0.8402	0.3795	0.7281	0.6587
Knowledge of English	0.1742	0.0886	0.2151	0.1621
Knowledge of computer	0.2888	0.1735	0.3594	0.2779
Azerbaijan				0.3088
Georgia				0.3502

Regression analysis: Armenia

	Dependent variable: Overall well-being		Dependent variable: material well-being	
Variable	**Coefficients**	**t-values**	**coefficients**	**t-values**
Constant	1.3878345	47.20	0.7975123	38.75
Urban	-0.0193421	-2.52	-0.0203804	-3.80
Rural	-0.0133349	-1.65	-0.0102882	-1.82
Male	-0.0362753	-5.59	0.0037819	0.83
Age	-0.0025229	-2.44	-0.0020426	-2.82
Square of age	0.0000349	3.36	0.0000174	2.39
Education	0.0046908	3.90	0.0042913	5.10
Islam	0.0225894	0.62	-0.0029065	-0.11
Married	0.0043465	0.62	-0.0031629	-0.64
Household size	0.0167576	3.12	0.0152953	4.07
Square of household size	-0.0012400	-2.19	-0.0010819	-2.73
Knowledge of Russian	0.0187395	2.07	0.0187275	2.95
Knowledge of English	0.0120687	1.28	0.0061636	0.93
Knowledge of computers	0.0054163	0.65	0.0130104	2.24

Number of observations: 1458
Overall well-being: R-square = 0.0707; Adjusted R-square = 0.0623. F-value for the regression: 8.45
Material well-being: R-square=0.1205; Adjusted R-square=0.1126; F-value for the regression: 15

Regression analysis: Azerbaijan

	Dependent variable: Overall well-being		Dependent variable: material well-being	
Variable	**Coefficients**	**t-values**	**coefficients**	**t-values**
Constant	11.8316226	26.51	0.4824276	4.87
Urban	-0.1362290	-2.57	0.0249705	2.13
Rural	-0.2341676	-4.24	-0.0194569	-1.59
Male	0.0851656	1.93	0.0216113	2.21
Age	-0.0038454	-0.49	-0.0026678	-1.52
Square of age	0.0000828	1.00	0.0000278	1.52
Education	0.0282320	2.91	0.0151228	7.03
Islam	-0.1374738	-0.35	-0.1614601	-1.88
Married	-0.0304890	-0.61	0.0220557	2.00
Household size	0.0685128	1.60	0.0307573	3.25
Square of household size	-0.0053741	-1.33	-0.0020082	-2.25
Knowledge of Russian	0.0085835	0.17	0.0299646	2.69
Knowledge of English	-0.0136726	-0.15	0.0731208	3.63
Knowledge of computers	0.1444519	2.08	0.0620521	4.02

Note: Number of observations: 1320
Overall well-being: R-square = 0.0412; Adjusted R-square = 0.0316. F-value for the regression: 4.31
Material well-being: R-square=0.1847; Adjusted R-square=0.1766; F-value for the regression: 22.76

Regression analysis: Georgia

	Dependent variable: Overall well-being		Dependent variable: material well-being	
Variable	**Coefficients**	**t-values**	**coefficients**	**t-values**
Constant	1.7960513	28.81	0.6083504	21.63
Urban	0.0213191	1.28	-0.0094580	-1.26
Rural	0.0281276	1.74	-0.0147864	-2.03
Male	-0.0298425	-2.36	0.0095116	1.67
Age	-0.0046623	-2.18	-0.0010999	-1.14
Square of age	0.0000640	3.05	0.0000137	1.45
Education	0.0100649	4.14	0.0059003	5.39
Islam	0.0654692	2.62	0.0212132	1.88
Married	-0.0203037	-1.45	-0.0053207	-0.84
Household size	0.0633456	5.66	0.0363038	7.19
Square of household size	-0.0057162	-4.70	-0.0031963	-5.83
Knowledge of Russian	0.0397967	2.52	0.0335963	4.71
Knowledge of English	0.0240622	1.28	0.0160914	1.90
Knowledge of computers	0.0530750	3.15	0.0424203	5.58

Note: Number of observations: 1497
Overall well-being: R-square = 0.0737; Adjusted R-square = 0.0656; F-value for the regression: 9.07
Material well-being: R-square=0.1944; Adjusted R-square=0.1874 F-value for the regression: 27.53

Shapley decomposition of overall well-being

Determinants	Armenia	Azerbaijan	Georgia	Three countries together
AREA OF RESIDENCE (URBAN OR NOT)	11.04	33.55	2.38	14.40
GENDER	28.12	7.85	4.39	4.68
AGE	18.15	9.47	12.63	9.52
EDUCATION	20.76	24.09	22.18	10.50
RELIGION	0.16	0.32	3.67	3.76
MARITAL STATUS	0.33	0.81	1.58	0.30
HOUSEHOLD SIZE	9.03	2.89	24.97	6.41
KNOWLEDGE (of Russian, English, Computers)	12.42	21.01	28.19	24.39
COUNTRY				26.04
TOTAL	100.00	100.00	100.00	100.00

Shapley decomposition of material well-being

Determinants	Armenia	Azerbaijan	Georgia	Three countries together
AREA OF RESIDENCE (URBAN OR NOT)	11.15	7.34	8.87	4.64
GENDER	0.97	3.53	1.01	1.89
AGE	16.39	2.74	6.16	7.92
EDUCATION	23.01	36.07	21.89	25.60
RELIGION	0.23	1.16	0.61	0.96
MARITAL STATUS	0.21	1.02	0.65	0.67
HOUSEHOLD SIZE	19.70	6.04	18.41	15.66
KNOWLEDGE (of Russian, English, Computers)	28.33	42.10	42.41	37.76
COUNTRY				4.92
TOTAL	100.00	100.00	100.00	100.00

Conclusions

- The paper derives measures of material and overall well-being for three countries in West Asia: Armenia, Azerbaijan and Georgia.
- Two-step procedure: correspondence analysis to aggregate the different variables within each of the four domains of well-being; stochastic frontier approach to estimate overall individual well-being
- Regression analysis of individual overall well-being on a set of explanatory variables (area of residence, gender, age, education, religion, marital status, size of household and familiarity with the Russian and English languages and the computer)
- Shapley decomposition to determine the impacts of the explanatory variables on the R-square of these regressions.
- In all cases education and knowledge (familiarity with Russian, English and the computer) had important contributions.
- Depending on the country the area of residence, age and the size household had also significant contributions.

Conclusions

- Analysis of the determinants of material well-being
- In all the three countries the most important Shapley contributions to the R-square of these regressions were education and knowledge (familiarity with Russian, English and the computer)
- The size of the household played generally also a role as well as age and the area of residence.
- It is then clear that the determinants of the overall level of individual well-being are not the same as those of individual material well-being, education and knowledge (as defined previously) being, as expected, more important determinants of material than overall well-being.

Comments

- Very interesting paper, which explores data on countries that we do not know much about
- Nice methodological approach
- But you need to stress the contribution (empirical/ methodological), the relationship with existing works on same/different data?
- Justify the choice of the variables for the OLS analysis
- Do these methods account for the associations between the different dimensions?
- Two-step aggregation procedure: within dimensions and between dimensions. Is there an assumption of independence between the two steps?
- Is it possible to find a connection between these statistical techniques and the normative features that are at the base of the more traditional aggregation procedures?

Comments

- How is it different from nonlinear principal component analysis (NLPCA) used to measure well-being with categorical variable?
- What about using more survey years?
- I would introduce a table with summary statistics on well-being in these countries, al least to understand in which country well-being is higher/lower. Also some information on the results of the correspondence analysis and efficiency analysis in the empirical part.