GDN Working Paper Series

Does Health Matter for Inequality in Transition Countries: The Case of Ukraine

Tamara Podvysotskaya, Elena Osinkina, Larysa Krasnikova and Yuriy Podvysotskiy

Working Paper No. 53

December 2011

About GDN

The **Global Development Network (GDN)** is a leading International Organization of developing and transition country researchers, policy and research institutes promoting the generation, sharing, and application to policy of multidisciplinary knowledge for the purpose of development. Founded in 1999, GDN is headquartered in New Delhi, with offices in Cairo and Washington.

This Working Paper has been prepared as part of the research which was conducted in the framework of a cooperation between The Vienna Institute for International Economic Studies (wiiw) and the Global Development Network, aiming to build research capacities in Southeast Europe, and financed by the Jubiläumsfonds of the Austrian National Bank (OeNB) and the Austrian Ministry of Finance. The views expressed in this publication are those of the author(s) alone.

Does Health Matter for Inequality in Transition Countries: The Case of Ukraine

Tamara Podvysotskaya[1], Elena Osinkina, Larysa Krasnikova and Yuriy Podvysotskiy[2]

Abstract

Significant deterioration in health and income inequality in Ukraine as well as in other CIS countries during the early transition motivated us to investigate the relationship between the two categories. Based on both macro- and micro-data analysis we focused on how different aspects of health affect income inequality in Ukraine. Significant impact of health has been proven on both micro- and macro levels. According to macro-level results, a 1% increase in life expectancy leads to a 2,1% decrease in income inequality as measured by index Gini. Micro-level research also evidenced significance of different health aspects for income inequality.

Keywords: Income Inequality, Health, Ukraine, Ukrainian Longitudinal Monitoring Survey

[1] Bureau of Economic and Social Technologies (BEST), Kyiv, Ukraine. E-mail: tpodvysotskaya@gmail.com

[2] Department of Finance, National university "Kyiv-Mohyla academy", Kyiv, Ukraine. E-mail: ypodvysotskiy@eerc.kyiv.ua

1. Introduction

Significant decrease in economic growth at the beginning of transition processes in Ukraine and other CIS countries was followed by a sharp decrease in all health aspects. Lack of public finance led to permanent underfunding of healthcare sector, growth of out-of pocket payments and decease of accessibility of medical services to the most vulnerable groups of population. This fact contributed to the growth of inequality which was observed at the beginning of transition period.

With our project we are going to fill this research gap and discuss the issues of inequality determined by different aspects of health on both macro and micro levels. We intend to study in detail how different aspects of health influence on income inequality and poverty in Ukraine.

The key motivation of our research proposal is based on the fact that health factor influence on inequality, especially in former USSR countries is heavily underestimated. Despite significant distortions in the economic development, social type of economy which dominated in Ukraine for 70 years of XX^{th} century managed to create quite efficient system of healthcare, which in 1978 at the Alma-Ata Conference was recognized as one of the most fair and accessible. For the last twenty years of transition most achievements of Soviet healthcare system were lost. In Ukraine for the period of transition were initiated 17 legislative draft bills, but none of them was adopted. Health care remains the only sector in Ukraine which wasn't reformed since time of independence. All this made considerable contribution towards deterioration of accessibility and quality of medical care in Ukrainian as well as in other CIS countries.

The research questions covered in our paper are the following. First, how much income inequality can be explained by inequality in health? Second, how much inequality can be explained by different aspects of health?

The paper is organized as follows. In Section 2 we provide a brief overview of existing literature on health-income relation, underlying the impact of health on income and income inequality. In Section 3 used data sources are presented, main variables described. Section 4 provides methodology used for our micro- and macro-level empirical investigations. Section 5 presents the results, while Section 6 draws conclusion.

2. Literature

Numerous studies have examined the relation between income and health, not many have related inequality in income to inequality of health within a given countries, particularly in Ukraine, as well as across countries.

This can be partially explained by the fact that an inequality measure can be constructed only for macro-level datasets. In turn, investigations of causality from health to income are also useful for understanding if differentials in health explain differentials in income.

Many papers evidence positive relation between health and income, although the question of causality leads to arguments. There are empirical papers that investigate the effect of income on health, and there are papers that investigate impact of health on income. Well-known theoretical framework explaining relation between health and income was introduced by Grossman (1972). His model of demand for health explains two-way causality: from income to health via decreasing substitution effect as well as from health to income through increased productivity.

Causality from income to health so far is more studied and more commonly recognized. For example, as Marmot et al. (2003) stated 'even in most affluent countries, people who are less well off have substantially shorter life expectancies and more illness than the rich.' Quite a number of empirical papers obtained empirical evidence of causality from income to health. In studies of Cutler, Deaton and Lleras Muney historical aspect of relationship between health and income is investigated. Marmot (2002) in his research paid attention to variation in mean income across different areas (countries and states) and related it to observed measures of health status.

Ross et al. (2000) related variations of inequality of income across areas to variation in the average level of health across areas.

Some papers concentrated on exogenous shocks to income, for example, Lindahl (2005) studied effect of lottery winnings on health, Frijters (2005) studied the effect upon health of income increase of East Germans after reunion of Germany. These researchers found rather significant effect of income on health; for example Lindahl (2005) found that a 10-percent income increase improves health by a twentieth of its standard deviation. Recent researches argue for prevailing importance for health of social, rather than pure economical, factors such as poverty, social support and inclusion, unemployment, working conditions and some other, see Marmot (2003).

There also exist dome empirical literature on causality from health to income; those papers come mainly from developed countries. The researchers, usually keeping in mind the reverse causality

problem, try to come up with exogenous health shocks or in other way to control for the reverse causality. Riphahn (1999) analyzed eleven rounds of German Socio-Economic Panel in a dynamic framework, and found that a health shock increases probability of leaving labor force and doubles probability of unemployment. Wu (2003) investigated the effect of exogenous health shocks on wealth of married couples aged between 51-61 and found strong effect. Dano (2005) used data on car accidents as exogenous shocks to health. She found that road injuries did not have significant overall effect on hourly wage rate, but it had significant effects on employment rates. Wagstaff (2007) uses changes in body mass index (weight in kilos divided by squared height in meters) as proxy for health. His results, obtained using Vietnamese individual-level dataset, suggest that adverse shock to health in fact were associated with reductions in earned income. He also found that non-medical consumption of better-off households decreased more than that of worse-off households, which was because worse-off households, unlike better-off ones, relied on dissaving and borrowing.

Previous evidence on poverty in Ukraine based on micro level largely comes from two studies by the World Bank (1996; 2005). These use different survey instruments and are therefore not strictly comparable. The former provides a static picture based on 1995 data while the latter titled "Determinants of Poverty during Transition covers the period from 1999 to 2003" but is not very informative about the developments in the 1990s, the period of the most intense transition.

Gustafsson and Nivorozhkina (2004) provide quite a good analysis of on the evolution of poverty and its determinants, also considering for health determinants, but they focus on one city only, which significantly restricts possibility for policy implications on the country level.

Bruck, Danzer and Murayev (2007) studied the incidence severity and determinants of household poverty in Ukraine over time using multiple measures of household welfare. They analyzed how and how much longterm transition process affects household welfare, particular attention in their research was paid to adjustments of the labour market. Despite they tried to investigate different determinants of poverty, health aspect was not covered at all.

Our study aims to fills this gap for Ukraine. We investigate causality from health to income inequality and from health to income in Ukraine and other transition economies. For this purpose we investigate on macro-level the effect of health on measure of income inequality, Gini index, and investigate on micro-level how much differentiation in income is explained by differentiation in health.

3. Data and Construction of Variables

3.1. Micro-level data

Here we introduce the ULMS individual-level database which was used for the micro-level empirical analysis in our paper. The ULMS panel data set is similar to Russian Longitudinal Monitoring Survey, and is conceived as a statistically representative sample of the population aged 15 to 72 years in Ukraine, which comprises approximately 8,500 individuals. We used two rounds of ULMS, namely for years 2004 and 2007 (similar number of observations in both). The ULMS data set covers a number of aspects relating to health aspects of individuals – mainly life styles, morbidity, chronic diseases, and self-assessment of their health by individuals.

Because 'health' is individual-specific parameter, we used *individual-level database* for our empirical investigation.

As dependant variable we used actual income for the last month, which was constructed as total of the following incomes:

- official and unofficial *salaries earned* in the workplace during the last month;
- market value of the *in-kind benefits* received in the workplace during the last month;
- *additional salary* (or salaries) for the last month (if person worked at more than one place);
- income from *household production* (agricultural and non-agricultural).

We did not include pensions, stipends and unemployment benefits, as these types of income do not seem to reflect available human capital of an individual.

Table 1. Description of variables used in regressions

Variable	Definition
Log of Income (for last 30 days)	Log of total: salary from all sources, household production at market prices
Drink (alcohol consumption)	*0*-"Less than once / month", …, *6*-"Every day"
Smoke	Number of cigarettes smoked per days
Sport (physical exercises out of work)	*0*-"No exercises", …, *4* "Daily exercises >30 min"
Height	In centimeters
Age	In full years
Male	*1* – male, *0* – female
Nowork (did not have a paid work during the last week)	*1*- did not have a paid work, *0* – had a paid work.
Arrears (wage arrears incidence)	*1*-arrears took place, *0* – otherwise

Variable	Definition
Education	*1*-vocational training, *2*-professional college, *3*-university and above; 0
Settlement (type of settlement dummy)	*1*-village, *2*-town, *3*-large city
Health (self-assessed)	*1* – bad, *2*-average, *3* – good
Morbidity (health aspect)	Constructed via PCA from a set of morbidity variables (more detailed info presented below)
Chronic morbidity (health aspect)	Constructed via PCA from a set of chronic morbidity variables (more detailed info below)
Life style (health aspect)	Constructed via PCA from a set of life style variables (more detailed info presented below)

For empirical analysis a limited sample was selected – including only strictly positive monthly income not exceeding 5500 UAH (USD 1000). Selection of the limited sample was performed based on analysis of income distribution within the total sample. We observed the following two features of the total sample:

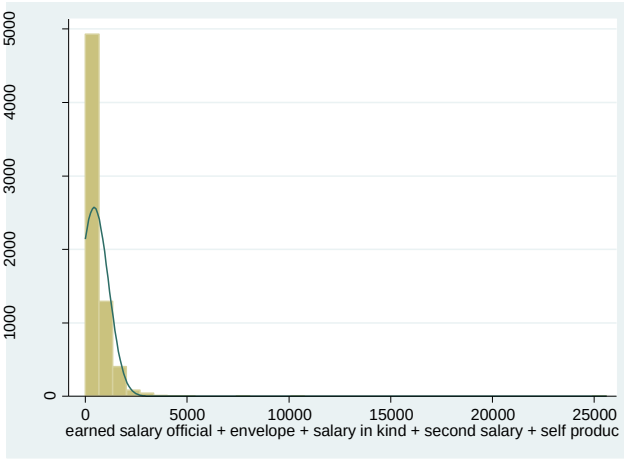
- high proportion (53%) of persons with zero income (relates to non-working aged people receiving pension, students and other dependants, and unemployed);
- few persons (only three observations) with incomes above 5500 UAH (USD 1000).

Therefore, limiting our sample in the way described above, we concentrate on income of persons using their human capital, and also avoid obtaining biased results due to few outliers. Distribution of income (year 2007) for the total sample and for the limited sample is presented in Figure 1.

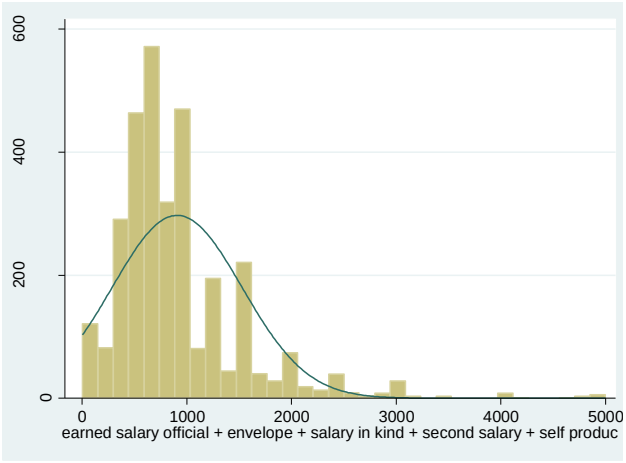
Dataset for year 2004 was limited to non-negative income only, as we did not identify outliers identified for this dataset.

Figure 1. Distribution of income, year 2007

A. Total sample



B. Limited sample (UAH 1 – 5500)

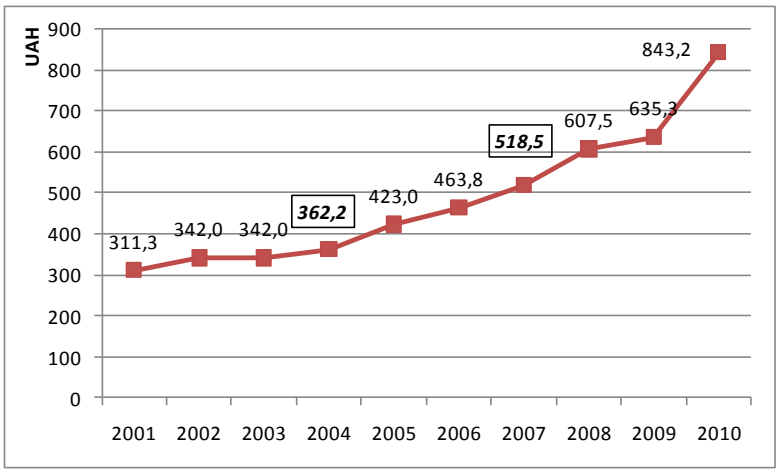


Identification of 'poor' category

It is usual to look at sub-samples within a total sample to identify some specific relationships. Usually, within selected sample 'poor' and 'rich' categories are identified. In our analysis only 'poor' category is identified. We do not identify 'rich' category, as thinness of the upper tail of income distribution problem, which is inherent to all questionnaire-based datasets (only three observations of income above USD 1000).

As an identifier for the 'poor' category we used official subsistence level, published by government of Ukraine (Figure 2). Official subsistence level includes minimum needs of food, clothes and social goods; official subsistence level is calculated by Ukrainian health-care authority, based on WHO norms of food and nutrition needs.

Figure 2. Subsistence level in Ukraine in 2001-2010



For year 2004 monthly subsistence level was UAH 362.2 (USD 72), while for year 2007 monthly subsistence level was UAH 518.5 (USD 102).

Summary statistics of variables used for empirical analysis for both years is presented in Table 2.

Table 2. Summary statistics of micro-level data

	2004		2007	
Variable	**Mean**	**Std. dev.**	**Mean**	**Std. dev.**
Ln of Income	6.533	0.871	6.548	0.894
Drink (alcohol consumption)	2.999	1.524	2.242	1.368
Smoke	0.836	0.371	16.258	9.916
Sport (physical exercises out of work)	0.569	1.122	0.565	1.110
Height	169.45	8.75	169.77	8.727
Age	40.97	12.17	37.59	12.60
Male	0.474	0.499	0.484	0.500
Nowork	0.057	0.231	0.065	0.246
Education	2.786	0.902	2.637	1.053
Settlement (type of settlement dummy)	2.153	0.848	2.127	0.858
Poor	1.265	0.441	1.275	0.446
Health (self-assessed)	2.182	0.591	2.262	0.617
Morbidity (health aspect)	92.558	5.898	92.992	6.413
Chronic morbidity (health aspect)	91.563	13.391	91.890	13.086
Life style (health aspect)	33.347	17.939	73.445	7.914

** - for year 2004 there was no info on number of cigarettes, instead "yes/no" data on smoking used*

3.2. Macro-level data

Major sources of data for the empirical research on macro-level include "OECD Health DATA 2009", WIDER databases, World Bank WDI 2009, and TransMONEE 2009 database of UNICEF. Combined macro-level panel database for estimation of macro-model includes information across 51 European countries (including Commonwealth of Independent States, Central and Eastern Europe, as well as Western, Northern and Southern Europe), for years 1991-2006. Major variables used are presented in Table 3.

Table 3. Summary statistics of macro-level data

Variable	Obs	Mean	Std. dev.	Min.	Max
Gini	556	32.049	7.123	18.000	62.500
Year	816	1998.5	4.613	1991	2006
Lifexp	660	73.677	4.561	62.386	81.515
Gdp	751	$236*10^{11}$	$4.81*10^{11}$	$7.09*10^{8}$	$2.90*10^{12}$
Gdppc	541	13115	14748	155	89778
gdp_def	739	119.1	688.5	-15.7	15442.3
Export	737	$7.97*10^{10}$	$1.47*10^{11}$	$8.15*10^{7}$	$1.31*10^{12}$
Import	737	$7.66*10^{10}$	$1.40*10^{11}$	$2.40*10^{8}$	$1.15*10^{12}$
Open	737	91.875	40.769	22.229	326.598
Reinvest	734	21.3	5.4	2.6	53.2
Cis	816	0.216	0.412	0	1
Cee	816	0.392	0.489	0	1

Gini coefficient ('gini') – represents level of inequality for a given economy in a certain year. Gini ranges from 0 (complete equality) to 100 (absolute inequality). Sample max represents Armenia in 1996, min represents Slovak Republic in 1991.

Public health ('lifexp') – is included to capture the joint effect of various aspects of health on inequality. Application of life expectancy as a appropriate proxy for public health is suggested and advocated by WHO, and is used in large number of researches. During further stages of our research we plan to define a number of aspects of health and analyze the effect of each of the aspects of health separately. Maximum Life expectancy in the sample refers to Switzerland in 2006, while minimum to Turkmenistan in year 2002.

GDP per capita ('gdppc') – considered as an explanatory variable for Gini. Measured in current US Dollars. Minimum – Tajikistan, 1996, maximum - Luxembourg, 2006.

Inflation (‘gdp_def’) – deflator of GDP, considered as an explanatory variable for inequality, as it is probably more comparable between countries, than CPI. Minimum observed value– Bosnia and Herzegovina in 1997, maximum – Georgia in 1996.

Openness of the economy (‘open’) – calculated as the ratio of exports plus imports to GDP, multiplied by 100. Minimum – Tajikistan in 1992, maximum – Luxembourg in 2006.

Dummy variables for CIS and CEE countries (‘cis’ and ‘cee’) – these dummy variables (0 and 1) were constructed so that to facilitate analysis of peculiarities for these groups of countries.

Preliminary results of empirical analysis of the dataset are presented in paragraph “Preliminary results and preliminary policy implications”

4. Empirical methodology

4.1. Micro-level methodology

In order to estimate hypothesis about influence of different aspects of health on income inequality, we estimated the inequality equation with each of the following health aspects:

- **Health** – as self-defined.
- **Chronic morbidity** – diseases of heart, lungs, liver, kidneys, digestive tract, back problems, and other chronic diseases. Estimated by principal-component analysis method of data-reduction.
- **Morbidity** – incidences of diabetes, myocardial infarction, high blood pressure, insult, anemia, tuberculosis. Estimated by principal-component analysis method of data-reduction.
- **Style of life** – smoking, drinking and doing exercises. Estimated by principal-component analysis method of data-reduction.

For each of the two separate data sets – ULMS 2004 and ULMS 2007 databases – we estimated the following model:

$$\ln(Income_i) = \alpha + \beta_1 \cdot Male_i + \beta_2 \cdot Nowork_i + \beta_3 \cdot Arrears_i + \beta_4 \cdot Pensioner_i + \\ + \beta_5 \cdot Educ_i + \sum_{j=2}^{3} \delta_j Settl_j + \gamma \cdot Health_i + \varepsilon_i \qquad (1)$$

Where, *Health* – is proxy for health and aspects of health of i^{th} individual. All other variables are as described in the previous paragraph.

We make use of the Principal Component Analysis (PCA) methodology in order to overcome possible multicollinearity problem, achieve data-reductions and facilitate intuitive interpretation of the results.

Morbidity was constructed based on data on incidence of serious diseases like diabetes, myocardial infarction, high blood pressure, stroke, anemia and tuberculosis. Scoring coefficients are presented in Table 4.

Table 4. Scoring coefficients for Morbidity

Variable	2004	2007
Diabetes	0.378	0.365
Myocardial infarction	0.550	0.526
High blood pressure	0.470	0.506
Stroke	0.574	0.547
Anemia	-0.003	0.184
Tuberculosis	-0.066	-0.034

Practically all diseases have positive signs within the aspect of 'Morbidity' (but for anemia and tuberculosis which have very low absolute values), therefore 'Morbidity' is *negatively correlated* with health (or decreasing in health, in other words).

Chronic morbidity was constructed based on data on availability of chronic diseases like heart disease, illness of the lungs, liver disease, kidney disease, gastrointestinal disease, spinal problems. Scoring coefficients are presented in Table 5.

Table 5. Scoring coefficients for Chronic Morbidity

Variable	2004	2007
Heart disease	0.438	0.424
Illness of the lungs	0.166	0.219
Liver disease	0.509	0.469
Kidney disease	0.450	0.450
Gastrointestinal disease	0.413	0.374
Spinal problems	0.365	0.415
Other chronic illnesses	0.122	0.196

All chronic diseases have positive signs within the aspect of 'Chronic morbidity', therefore 'Chronic morbidity' is *negatively correlated* with health.

Life style was constructed based on data on smoking intensity, alcohol drinking intensity and frequency of doing out-of-work exercises. Scoring coefficients are presented in Table 6.

Table 6. Scoring coefficients for Life Style

Variable	**2004**	**2007**
Smoke	0.619	0.607
Drink	0.615	0.592
Sports	-0.489	- 0.531

Negative variables 'smoking' and 'drinking' enter aspect of 'Life Style' with positive signs, while positive variable 'sports' enters with a negative sign, therefore obtained 'Life style' is *negatively correlated* with health.

New variables (the first components), obtained from application of PCA method, are usually not measured in some economically sensible units. Therefore, we implemented normalization of these variables, mapping them into 0-100 scale, to achieve better interpretation and comparability. Besides, in order that each aspect of health to be positively correlated with health, we applied two different normalization approaches, depending whether a health aspect is positively (equation 2) or negatively (equation 3) correlated with health:

$$x_i^{NORM} = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \cdot 100 \qquad \text{normalization of health aspects positively correlated with health} \qquad (2)$$

$$x_i^{NORM} = \frac{x_{\max} - x_i}{x_{\max} - x_{\min}} \cdot 100 \qquad \text{normalization of health aspects negatively correlated with health} \qquad (3)$$

After application of the normalization method all three aspects of health became increasing in health (that is positively 'correlated' with health) and scaled on 0-100 scale.

Descriptive statistics of the obtained components of health-aspects is provided below:

Table 7. Descriptive statistics for Aspects of Health

Variables *(normalized 0-100)*	**2004** Average	Sd.dev	Obs.	**2007** Average	Sd.dev	Obs.	Effect on health
Morbidity	92.558	5.898	3298	92.992	6.413	3147	*Negative*
Chronic morbidity	91.563	13.391	3296	91.890	13.086	3147	*Negative*
Life styles	33.347	17.939	1098	73.445	7.914	912	*Negative*

The data set under analysis is a cross-section, which is exposed to the heterogeneity problem. In order to overcome the heterogeneity problem, we apply the OLS with generalized residuals.

4.2. Macro-level methodology

As for dynamic methodology, we intend using Arellano-Bower (1995)/ Blundell-Bond (1998) linear generalized method of moments, which is specially designed for data-sets with small "T" (time periods) and large "N" (number of individuals). This methodology allows to deal with autocorrelation and heteroscedasticity, as well as cushion the endogeneity / reverse-causality problem. An important advantage of dynamic set-up over cross-section one is that it is less vulnerable to the problems of unobserved heterogeneity in individuals, as it includes fixed-effects in the panel data (Roodman, 2006).

The regression equation would include lagged dependent variable in its right-hand side, as well as full set of time dummies, besides additional lags of practically all regressors form the matrix of instruments, according to the Arellano-Bower/ Blundell-Bond methodology. The equation of relationship between income inequality and health factor, as introduced in the cross-section paragraph, under dynamic set-up is arranged as follows (variables are logged):

$$\Delta G_{i,t} = \alpha_0 + (\alpha_1 - 1)G_{i,t-1} + \alpha_2 y_{i,t} + \alpha_3 H_{i,t} + \gamma Z_{i,t} + \lambda N_{i,t} + \varepsilon_{it}, \quad \text{or alternatively}$$

$$G_{i,t} = \alpha_0 + \alpha_1 G_{i,t-1} + \alpha_2 y_{i,t} + \alpha_3 H_{i,t} + \gamma Z_{i,t} + \lambda N_{i,t} + \varepsilon_{it}, \quad \text{where}$$

$G_{i,t}$ - Gini index is a proxy for income inequality in country i, time period t;

$G_{i,t-1}$ - is prior realization of Gini index, α_1 is expected to be in range (0; 1);

$y_{i,t}$ - per capita GDP in country i, time period t;

$H_{i,t}$ - health factor variable in country i, time period t, calculated by principal component analysis methodology from a set of relevant health variables;

The other variables include years of schooling attainment ($Z_{i,t}$) and number of other variables as in cross-section set-up – inflation, degree of international openness, etc. ($N_{i,t}$), and time-dummies.

5. Results

5.1. Micro-level estimation results

Based on the methodology described above, we estimated equation (1) for the datasets of years 2004 and 2007.

Several specifications were estimated for each year:

- OLS model including health as self-defined (model 1a);
- IV model including health as self-assessed, health instrumented by height and mother education (model 1b);
- Model including health as self-defined, including dummy 'poor' and its interaction terms (model 1c);
- Model including 'Morbidity' aspect of health (model 2);
- Model including 'Chronic morbidity' aspect of health (model 3);
- Model including 'Life style' aspect of health (model 4a);
- Model including components of 'Life style' explicitly – drinking, smoking and doing exercises (model 4b).

Actual estimation results obtained are presented in Table 8 and Table 9.

The parameters of the effect of 'health' and its aspects on personal income for most models are positive and significant, which aligns with expected outcome. Here we detail estimation results regarding each aspect of health.

Aspect of *self-assessed* health is significant under each specification. The model with instrumental variables for health was estimated to verify robustness of estimated parameters of the other models, as there could be inverse relation from income to health. Estimated parameters

of the IV model are comparable to that of the OLS model, therefore we do not instrument health in other models.

According to estimates of Model 1a, an increase in *health* by 1 standard deviation leads to an increase in income by 6,7% (2004: by 3,3%); by multiplying coefficients by mean of *health* we obtain elasticity coefficient, showing percentage change in income in relation to percentage change in health – 0,24 (2004: 0,12).

Coefficients of health aspects of *Morbidity* and *Chronic morbidity* mainly do not show statistical significance, although signs of the coefficients match expectations. We suggest that these factors do in fact determine income differential, but their effect should be analyzed in dynamics. According to estimates of Model 2 (year 2004), an increase in *Morbidity* aspect by 1 standard deviation leads to as increase in income by 4,1%; coefficient of elasticity equals 0,65.

Health aspect of *Life-style* is statistically significant for year 2007. According to outcomes of estimation of Model 4a, an increase in Life-style by 1 standard deviation leads to 3,2% increase in income, coefficient of elasticity equals 0,29.

Besides, the effect of individual components of Life-style was estimated (Model 4b). The effect of *Sports* on income has positive, while *Smoking* has negative signs in either year.

Estimated coefficients of the other variables of equation (1) provide interesting observations upon income differentials and income inequality.

On average, men have by 26-32% higher income than women, which is consistent over all model specifications and periods.

Persons not currently employed in a paid workplace (*'Nowork'* = 1), naturally, earn less. Positive income of such persons relate to their household production of agricultural and non-agricultural goods.

Another proxy of human capital, *Education*, also significantly explains differentials in income. Each additional level of education (secondary, technical school, university, graduate school) seem to add, on average, 6-7% (2004: 5%) to personal income. Significance of both '*Health'* and '*Education'* provides evidence for the need for stimulating both health and education infrastructure as a means of income inequality reduction.

Coefficients of dummies for type of *Settlement* are significant. Estimation outcomes evidence that persons living in towns earn, on average, by 5-15% more, while persons living in cities earn by 15-50% more than persons from countryside.

Inclusion of dummy for people below subsistence level ('*Poor*') adds significant explanatory power to the model, although, it does not seem to provide valuable inferences. Probably, a dynamical analysis might lead to interesting conclusions about sub-category of poor people.

Table 8. Micro-level estimation results (year 2007)

Log of Income	**(1a)**	**(1b)**	**(1c)**	**(2)**	**(3)**	**(4a)**	**(4b)**
Male	0.320***	0.307***	0.254***	0.262***	0.262***	0.233***	0.322***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Poor*Male			-0.419***	-0.429***	-0.434***	-0.258	
			(0.000)	(0.000)	(0.000)	(0.216)	
Nowork	-1.255***	-1.224***	-0.860***	-0.863***	-0.864***	-0.546***	-0.904
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
Educ	0.072***	0.073***	0.061***	0.062***	0.062***	0.071***	0.033
	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.210)
Poor*Educ			-0.179***	-0.180***	-0.179***	-0.304***	
			(0.000)	(0.000)	(0.000)	(0.001)	
Settlement=2	0.063*	0.056***	-0.009	-0.006	-0.008	0.062	0.179**
	(0.096)	(0.000)	(0.769)	(0.833)	(0.802)	(0.282)	(0.015)
Settlement=3	0.139***	0.138	-0.006	-0.007	-0.007	0.024	0.355***
	(0.000)	(0.000)	(0.834)	(0.818)	(0.815)	(0.664)	(0.000)
Health	0.108***	0.196**	0.044***				
	(0.000)	(0.041)	(0.004)				
Poor**Health*[(1)]			-0.096*	-0.009**	-0.003	-0.003	
			(0.054)	(0.011)	(0.120)	(0.838)	
NMorbid				0.003			
				(0.118)			
Nmchronic					0.000		
					(0.309)		
Nlstyle						0.004**	
						(0.045)	
Drink							-0.048**
							(0.021)
Smoke							-0.002
							(0.045)
Sport							0.011
							(0.721)
Constant	5.966***	5.776*	6.533***	6.386***	6.559***	6.292***	6.376
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Poor			-0.350**	0.287	-0.255	-0.352	
			(0.011)	(0.398)	(0.230)	(0.726)	
Observations	3136	2967	3136	3146	3146	912	912
R-squared	0.195	0.187	0.506	0.505	0.505	0.488	0.083
Prob > F	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Robust p values in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
(1) – interaction variables for each aspect of health

Table 9. Micro-level estimation results (year 2004)

Log of Income	**(1a)**	**(1b)**	**(1c)**	**(2)**	**(3)**	**(4a)**	**(4b)**
Male	0.266***	0.038	0.031*	0.273***	0.272***	0.279***	0.243***
	(0.000)	(0.783)	(0.067)	(0.000)	(0.000)	(0.000)	(0.000)
Poor*Male			-0.020				
			(0.643)				
Nowork	-0.183***	0.024	-0.107***	-0.195***	-0.191***	-0.140	-0.154
	(0.007)	(0.860)	(0.005)	(0.004)	(0.005)	(0.218)	(0.165)
Educ	0.0515***	-0.002	-0.021**	0.050***	0.051***	0.007	-0.000
	(0.002)	(0.860)	(0.022)	(0.002)	(0.002)	(0.783)	(0.993)
Poor*Educ			0.086***				
			(0.000)				
Settlement=2	0.408***	0.395***	0.151***	0.413***	0.412***	0.391***	0.363***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Settlement=3	0.642***	0.635	0.203***	0.650***	0.650***	0.573***	0.558
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Health	0.055**	1.400*	-0.016				
	(0.018)	(0.076)	(0.252)				
Poor*$Health^{(1)}$			0.077***				
			(0.005)				
NMorbid				0.007***			
				(0.003)			
Nmchronic					0.001		
					(0.208)		
Nlstyle						0.0003	
						(0.830)	
Drink							0.043***
							(0.008)
Smoke							-0.002
							(0.978)
Sport							0.073***
							(0.000)
Constant	5.760***	3.074*	6.917***	5.183***	5.756***	6.017***	5.902***
	(0.000)	(0.076)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Poor			-2.007***				
			(0.000)				
Observations	3273	3184	3273	3294	3292	1097	1097
R-squared	0.135	n/a	0.739	0.137	0.135	0.088	0.104
Prob > F	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Robust p values in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
(1) – interaction variables for each aspect of health

5.1. Macro-level estimation results

Considering the macroeconomic panel-data analysis we estimated the following equation:

$G_{t,i} = \beta_0 + \alpha G_{i,t-1} + \beta_1 Lifex_{i,t} + \beta_2 gdppc_{i,t} + \beta_3 gdp\_def_{i,t} + \beta_4 open_{i,t} + \delta Year_{i,t} + \varepsilon_{i,t}$, where

$G_{t,i}$ – index Gini at time *t* for country *i*;

$Lifex_{t,i}$ – life expectancy (proxy for health) at time *t* for country *i*;

$gdppc_{t,i}$ – GDP per capita at time *t* for country *i*;

$gdp\_def_{t,i}$ – GDP deflator at time *t* for country *i*;

$open_{t,i}$ – openness of the economy - (*Exports*+*Imports*)/*GDP* - at time *t* for country *i*;

$Year_{i,t}$ – complete vector of year-dummies (except for years 1991 and 1994);

$\varepsilon_{i,t}$ – error term.

We estimated a number of specifications of the mentioned equation – ordinary OLS, fixed- and random-effects, fixed- and random-effects with autoregressive residual, system GMM Arellano-Bover / Blundell-Bond. Results of all specifications in summarized in the table, and Stata output is provided in the Appendix. The equation with estimated coefficients using the system GMM methodology is provided (t-values in parentheses):

$$\hat{G}_{t,i} = \underset{(11.15)}{92.698} + \underset{(3.22)}{0.206} G_{i,t-1} - \underset{(-9.92)}{0.898} Lifex_{i,t} + \underset{(3.82)}{0.00007} gdppc_{i,t} + \underset{(3.15)}{0.0025} gdp\_def_{i,t} - \underset{(-7.85)}{0.0339} open_{i,t}$$

The direction of coefficients obtained coincides with underlying intuition. There is certain 'path dependency' as Gini coefficient is dependent on its *lagged value*.

GDP per capita does not have an effect on inequality, as richer countries are not usually more equal. Theoretically, inequality should not be related to the level of average per capita income, as income growth rate of the richest quintiles could outpace that of the poorest.

GDP deflator is positive and statistically significant. The result is consistent with the mainstream economic theory, which defines a number of ways in which inflation in a given economy 'hits' the low-income part of population stronger than the high-income part of population.

Coefficient of *openness of the economy* is negative and statistically significant, meaning that higher openness may lead to a decrease in income inequality. This is, probably, also coincides with expectations.

Table 10. Macro-level estimation results

	(1) Simple OLS		(2) Fixed effects		(3) Autoreg FE		(4) Arellano-Bover	
	coeff.	*p-value*	*coeff.*	*p-value*	*coeff.*	*p-value*	*coeff.*	*p-value*
$Gini_{(t-1)}$	0.704***	0.000	0.284***	0.000	-0.157***	0.002	0.206***	0.001
$Lifexp_{(t)}$	-0.375***	0.000	0.179	0.581	0.573***	0.000	-0.898***	0.000
GDP p.c. $_{(t)}$	0.000	0.162	-0.000	0.538	-0.0001	0.109	0.0001***	0.000
$GDP\ deflator_{(t)}$	0.003***	0.001	0.0003	0.755	0.003**	0.019	0.0025***	0.002
$Openness_{(t)}$	-0.015***	0.002	0.0004	0.977	0.0029	0.791	-0.0339***	0.000
Constant	37.876***	0.000	10.405	0.658	-5.341***	0.001	92.698***	0.000
R^2 *(adjusted)*	0.7044		0.5415		0.4096		0.5109	
No obs.	417		417		376		417	

* significant at 10%; ** significant at 5%; *** significant at 1%

Simple OLS may not be optimal with panel data. The estimates of coefficients derived from regression may suffer from omitted variable bias – which arises when there is some unknown variable or variables that cannot be controlled for that affect the dependent variable. While using specific panel-data methodology we can control for omitted variables even without observing them. Hence, fixed- and random- effects methodology would be superior to simple OLS.

The decision on fixed versus random-effects estimation is usually based on Hausman test. If the null hypothesis is rejected then the fixed effects estimation should be used. For our model fixed effects were chosen based on this criterion.

In case of autocorrelation problems the estimators remain unbiased and consistent despite no longer efficient (we include lag of Gini coefficient among explanatory variables). The inefficiency of estimators means that t statistics and F statistics tests cannot be trusted any more. With this in view, we applied 'xtregar' operator in Stata, which offers a within estimator for a fixed effect model and the Baltagi-Wu GLS estimator of the random effects model. Although random effects Baltagi Wu GLS model provided us with coefficients closer to expectations, based on Hausman test we were forced to pick fixed-effects model with autocorrelated residuals. We relate the 'illogical' signs in both fixed-effects models to the endogeneity problem in our model.

It is important to mention that our proxy for health, *Life expectancy*, could be itself a function of inequality in current and/or previous period – inequality relates to unequal medical treatment and

higher stress in the society. Endogeneity problem leads to biased estimates, and unreliable predictions. To deal with the endogeneity problem, we estimate Arellano-Bover/ Blundell-Bond linear dynamic panel-data model. Briefly, benefits of this method is that it corrects for unobserved effects by taking differences, applies instrumental variable procedure to correct for endogeneity problem. According to results of estimating parameters of Model 4, an increase in life expectancy by 1 year, leads to a decrease in index Gini by 0,9. In terms of average elasticity, an increase in health (life expectancy) by 1% leads to decrease in inequality (index Gini) by 2,1%.

Conclusions

The focus of present analysis was to explore how different aspects of health affect the relation to income inequality. According to the theoretical and empirical debate, it was assumed that in the case of Ukraine and other transition countries influence on income inequality of health factors was heavily underestimated.

Our research was motivated by data on sharp deterioration in health, combined by significant increase in income inequality, during the early transition period in Ukraine and other CIS countries.

In order to address this research question, we implemented empirical analysis on both micro- and macro levels. This allowed us to verify robustness of the results and capture specific inference from either level of data.

Significant impact of health has been proven on both micro- and macro levels. Specifically, according to macro-level results, a 1% increase in life expectancy leads to a 2,1% decrease in income inequality as measured by index Gini. Micro-level research evidenced significance of health as proxied by self-assessed level of health, and by life-styles aspects of health.

Interestingly, other important variables explaining income differentials were level of education, gender and size of settlement. This hints to a policy implications about equal importance of improvements in health-care and education for combating income inequality.

Out findings are consistent with previous international studies that have indicated evidence about health-income inequality causality. Our findings expand current debates about health-income

inequality association by considering for different aspects of health in income equation in case of Ukraine.

It is necessary that future empirical investigations should deeper consider for health-care sector, besides it should utilize subsequent rounds of ULMS in order to apply dynamic analysis. The meaning of specific health aspects that influence income inequality is important in order to get a deeper understanding of the health-income inequality relationship.

References

Beck, Th., Demirguc-Kunt, A., Levine R. (2004). Finance, Inequality and Poverty: Cross-Country Evidence. – NBER Working Paper 10979, 47 p. Working Paper at: http://papers.nber.org/papers/w10979.pdf?new_window=1

Becker C., Bloom D. (1998). The Demographic Crisis in the Former Soviet Union: Introduction. – World Development, No.11, pp. 1913 – 1919, at: http://econpapers.repec.org/article/eeewdevel/v_3A26_3Ay_3A1998_3Ai_3A11_3Ap_3A1913-1919.htm

Brainerd, E. (2000). Women in transition: Changes in gender wage differentials in Eastern Europe and the former Soviet Union. – Industrial & Labor Relations Review 54(1), pp 138-162.

Bruck T., Danzer A., Muravyev A. (2007). Determinants of poverty during transition: Household Survey Evidence from Ukraine. – IZA Discussion Paper 3228.

Buddelmeyer H., Cai L. (2009). Interrelated Dynamics of Health and Poverty in Australia http://www.iza.org/index_html?lang=en&mainframe=http%3A//www.iza.org/en/webcontent/publications/papers/viewAbstract%3Fdp_id%3D4602&topSelect=publications&subSelect=papers

Dano A.M. (2003). Road Injuries and Long-Run Effects on Disposable Income, Wages, and Employment. – AKF, Institute of Local Government Studies, Discussion paper.

Dasgupta P.S., Sen A.K., Starrett D.A. (1973). Notes on the measurement of inequality. – Journal of Economic Theory 6, pp. 180-187.

Frijters P., Haisken-DeNew J. and Shields M. (2005). The Casual Effect of Income on Health: Evidence from German Reunification. – Journal of Health Economics 24 (5), 997-1017.

Garvey G. (1952). Inequality of Income: Causes and Measurement. – Part II, NBER, Chapter p. 25-48, at http://papers.nber.org/chapters/c9764.pdf?new_window=1

Grossman M. (1972). On the Concept of Health Capital and the Demand for Health. – Journal of Political Economy, No 80, pp. 223 –255.

Jaques Silber (1999). Handbook on Income Inequality Measurement.

Klugman J.,Micklewright J. and G. Redmond (2002). Poverty in the transition: social expenditures and the working-age poor. – CEPR Discussion Paper 3389.

Lindahl M. (2005). Estimating the Effect of Income on Health and Mortality Using Lottery Prizes as Exogenous Source of Variation in Income. – Journal of Human Resources 40(1), pp 144-168.

Marmot M., Wilkinson R. (2003). Social Determinants of Health. The Solid Facts. Second Edition. – International Center for Health and Society, World Health Organization.

Peichl A., Schaefer T., Scheicher C. (2008). Measuring Richness and Poverty: A Micro Data Application to Europe and Germany

http://www.iza.org/index_html?lang=en&mainframe=http%3A//www.iza.org/en/webcontent/publications/papers&topSelect=publications&subSelect=papers

Riphahn R. (1999). Income and Employment Effects of Health Shocks. A Test case for the German Welfare State. – Journal of Population Economics 12(3), pp 363-389.

Wagstaff A. (2007). The Economic Consequences of Health Shocks: Evidence from Vietnam. – Journal of Health Economics 26(1), pp. 82-100.

Wu S. (2003). The Effects of Health Events on the Economic Status of Married Couples. – Journal of Human Resources 38(1), pp. 219-230.

World Bank (2005). Ukraine: Poverty Assessment. Poverty and Inequality in a Growing Economy. – Washington DC, The World Bank.

World Bank (2007). Ukraine: Poverty Update. – Washington DC, The World Bank.

Zimmerman F. J. and M. R. Carter (2003). Asset smoothing, consumption smoothing and the reproduction of inequality under risk and subsistence constraints. – Journal of Development Economics 71(2), pp. 233-260.