

## Human Development Report **2005**

Human Development Report Office OCCASIONAL PAPER

## Trends in Global Income Distribution, 1970-2000, and Scenarios for 2015

Dikhanov, Yuri. 2005.



### TRENDS IN GLOBAL INCOME DISTRIBUTION, 1970-2000, AND SCENARIOS FOR 2015

Yuri Dikhanov WORLD BANK ydikhanov@worldbank.org

### **SUMMARY**

The paper builds on the author's prior research in areas of evolution of the global income distribution and of the "quasi-exact" polynomial interpolation of density functions. The 1970-2000 estimates are augmented with two 2015 scenarios: (a) distribution-neural growth (national distributions kept constant) and (b) pro-poor growth (the poor's income grows at twice the average rate until 2015). The scenarios are based on historical 1990-2002 trends in GDP growth and UN population projections for 2015. Compared to 2000, the distribution-neutral growth scenario for 2015 shows a decline in the Gini – 0.300, Theil 1 and 2 – 0.114 and 0.082, respectively, and a decline in absolute poverty from 1,172 mln. in 2000 to 689 mln. in 2015. These changes are explained to a large degree by the projected fast growth in India and China. The pro-poor growth scenario resulted in additional 253 mln. people rescued from poverty. Two more simulations are presented: (1) transfers being made to the poor in 2000, and (2) distribution-neutral growth occurring during 1970-2000. An annex discusses advantages of the "quasi-exact" polynomial interpolation of income distributions.

The views expressed in the paper are solely the author's and should not be attributed to the World Bank or its Board of Directors.

### **TABLE OF CONTENTS**

INTRODUCTION	
Methodological Issues	
World Inequality Accounting	4
SIMULATIONS	5
SCENARIO I. DISTRIBUTION-NEUTRAL GROWTH, 2000-2015	6
SCENARIO II. PRO-POOR GROWTH	7
SCENARIO III. TRANSFER TO THE POOR	9
SCENARIO IV. DISTRIBUTION-NEUTRAL GROWTH, 1970-2000	11
HALVING POVERTY	13
BOX I. WHEN INCOME GROWTH CANNOT CROWD OUT POverty	14
ANNEX I. Quasi-Exact POLYNOMIAL INTERPOLATION of Distribution FUNCTIONS	15
Precision of the procedure	15
"Good" case for interpolation	16
"Bad" case for interpolation	17
COMPARISON TO UNIMODAL FITTING PROCEDURES	18
Comparison of General Quadratic [GQ] interpolation to actual distribution	
Comparison of "Quasi-Exact" Polynomial interpolation to actual distribution	20
ANNEX II. DECOMPOSITION OF INEQUALITY MEASURES	
1. Decomposition of GINI - coefficient	21
2. Decomposition of entropy (Theil) indexes	24
3. Decomposition of normalized variance	25
4. Decomposition of decile ratio	25
5. Lorenz curve	25
6. Some properties of log-normal distribution	
7. Two alternative representations of the Gini coefficient	
References	
ANNEX III. World Income Distribution, 1970-2015	
ANNEX IV. HOW TO READ THE DISTRIBUTION DIAGRAM	

### INTRODUCTION<sup>1</sup>

The paper studies recent trends in the global income distribution. The 1970-2000 estimates are supplemented with two 2015 scenarios: (a) distribution-neural growth (national distributions kept constant) and (b) pro-poor growth (the poor's income grows at twice the average rate until 2015). The scenarios are based on historical 1990-2002 trends in GDP growth and UN population projections for 2015. In addition, two more simulations are presented: (c) transfers to the poor in 2000, and (d) distribution-neutral growth during 1970-2000. Scenario (c) examines the alleviation of poverty through targeted transfers occurring at one time, and scenario (d) studies effects of changes in national distributions on global poverty. Annex II discusses various inequality measures and shows difficulties of using the Gini coefficient for decomposing inequality.

### **METHODOLOGICAL ISSUES**

Three major statistical concerns are to be addressed in the paper<sup>2</sup>: (1) the relevance and quality of the basic data; (2) the choice of methodology for international comparisons; (3) the need to distinguish the separate 'between'' country and "within" country differences affecting inequality.

In the first case, the data are not perfect and there is a continuing conceptual and practical problem of mapping micro household information (not necessarily representative) onto the comprehensive national accounts aggregates. The re-estimation of observed national benchmark numbers by applying official price indices and other indicators to specific measures is required to adjust everything to a common reference year basis. That was one of the reasons bihind recasting means from the household surveys into PCE terms.

As to the second issue, the study uses PPPs rather than currently reported official exchange rates to convert data expressed in national currencies into a common international unit of account (the dollar). While some econometric techniques are required to extend the PPPs reported for 118 benchmark countries to other countries for which such data are not available, the use of PPPs to equalize underlying price level differences obviates the problems caused by the volatility of exchange rates both over time and between countries. This enables the analysis to include the regional composition of inequality.

Third, the "within" country profiles are based on national household surveys while the "between" differences are calculated from PPP converted World Bank GNP per capita estimates.

The preferred methodology utilizes a formalized continuous distribution function, rather than discreet distribution derived from quintile and decile data groups. National level household income distribution data generally relate to consumption rather than income. When drawn from household survey results, the outlay data usually refer to actual expenditures rather than true consumption including that occurring as a result of receipts in kind or imputed income from own account production, specifically from garden plots, for own

3

<sup>&</sup>lt;sup>1</sup> The paper was presented at the *Third Forum on Human Development*, "Cultural Identity, Democracy, and Global Equity", organized by the UNDP and the French Ministry of Foreign Affairs, held in Paris, France during January 17-19, 2005.

<sup>&</sup>lt;sup>2</sup> See *Dikhanov*, *Ward* for detailed methodology.

consumption. These elements are estimated separately. Whichever variable is chosen, this will have a particular impact on the shape of the curve. The use of reported expenditure figures, for example, tends to condense the shape of the actual income distribution curve and compress the apparent range of inequality. But such expenditures, and particularly overall consumption estimates, are closely related to disposable income, particularly at lower income levels, and they tend to have a more relevant bearing on household wellbeing. The real distribution of individual welfare, however, raises separate issues and has to be considered alongside the public provision of collective amenities and the allocation by governments and non-government organizations of non-market goods and services to households and individuals.

### WORLD INEQUALITY ACCOUNTING

SCOPE OF THE ARGUMENT: At the outset, let's define the scope of the current argument. It is about global *income* inequality. It is not concerned with a more holistic perception of global inequality that would need to incorporate some notional value of non-market goods and services delivered by governments and NGOs to individual recipients across the world. This would require a far more extensive and sophisticated set of impact measures than is currently available in the existing range of collective social indicators.

It is also not about the distribution of wealth and ownership of productive assets which, as the primary sources of economic power, are undoubtedly very important in *explaining* the extent of global inequality. The distribution characteristics are drawn primarily from household surveys, where the concept of income used in this analysis refers to disposable income rather than to a national accounts definition of gross personal income that includes unrealized capital gains and taxes, both of which can be significant.

For the purposes of this exercise, global income is taken to be the sum of the reported as well as estimated and imputed *Personal Consumption Expenditures* (PCE) of all countries as presented in the World Bank databases. Thus, national accounting categories were used in defining national and global incomes instead of averages from household surveys. This has been done for the following reasons: (1) to be consistent with national accounts and to arrive at global income computed using consistent methodology across countries; (2) to attempt to extend national accounts to income groups; and (3) to keep consistency between national GDP/PCE growth rates and incomes used in inequality measurements<sup>3</sup>.

The PCE values in local currencies are converted to "international" dollars using 1999 purchasing power parities (PPPs). These PPPs are in turn derived from the 1993-96 ICP exercise and in some cases (including such an important case as India) from 1980 and 1985 exercises. In the case of example China, a growth rate different from the official one was employed<sup>4</sup> [from Maddison (2001)]. Also China-India GDP per capita ratio was kept in line with that publication. In addition, the PPP in Geary-Khamis terms were converted to

<sup>3</sup> For instance, during the 90s, the Indian household sample survey means were growing slower than the PCE, which may indicate a bias in the national household surveys.

<sup>4</sup> As it turns out, the extent to which global income inequality has grown over the last decade is significantly influenced by the real economic growth of China. In particular, if the reported official GDP growth rate of China of 10.7% p.a. over the past ten years, 1990-9, is accepted, then there appears to be a marginal improvement in the overall global income distribution during the 90s. On the other hand if, as others [Maddison, Wu, and Keidel] have suggested, the actual GDP growth in China was closer to 8% per annum during that period then global inequality has either remained stable or deteriorated [using the Theil index it increased from 0.908 in 1990 to 0.925 in 1999, whereas the Gini marginally improved from 0.686 to 0.683].

the EKS to remove the substitution bias present in the former.<sup>5</sup> As a result, Chinese and Indian GDP per capita levels were somewhat lower than those found in the World Bank database. The PCE shares in GDP were taken from the World Bank databases.

Additionally, World Bank group aggregation and filling procedures were used to obtain global totals for countries where GNP information was missing. These techniques are detailed in the reference notes to 2001 *World Development Indicators* and are already fairly well known and recognized. The imputation procedures apply, for the most part, to only a few small countries. Overall, the 45 countries in the sample were responsible for about 5/6 of the global PCE and population. The remaining 1/6 was distributed at the regional level according to the regional coverage patterns. The primary objective was to envision how 'global PCE' is broadly distributed across various groups by their respective income levels, and to see whether this changes markedly as a result of applying different assumptions.

Thus, with this approach, it is possible to derive various regional (geographical) sub-aggregates of the global distribution and compare these with [a] other regions and [b] the global position. In principle, this can also be done for specific economic "blocks" such as "low income developing countries," "middle income countries," "industrial countries," trading countries, or any other similar grouping.

No adjustment is made to reported GNP/PCE measures for the under-recording of informal and "shadow economy" transactions, although at the country level these can be significant where the authorities exercise only a weak control over policy management and taxation. In many countries, however, the GNP number will usually incorporate some official estimate to account for such missing values.

REMARK ON REGIONAL COMPOSITION: Africa includes North Africa and Sudan. Some tables show results for Sub-Saharan Africa separately. Other regions are consistent with the UN or World Bank groupings.

### SIMULATIONS

- 1. Distribution-neutral growth, 2000-2015. National distributions kept constant during the period.
- 2. Pro-poor growth, 2000-2015. Incomes of the population below Int. \$700 grow twice the average rate.
- 3. Wealth transfer to the population below Int. \$1,400 from the richest 5% with marginal tax/transfer rates.
- 4. Distribution-neutral growth, 1970-2000. National distributions kept constant during 1970-2000.

<sup>5</sup> The extent of the bias (so called Gerschenkron effect) can attain in some cases 50% or more, and, thus, would seriously distort the overall picture of the income distribution [see Dikhanov (1994)].

### SCENARIO I. DISTRIBUTION-NEUTRAL GROWTH, 2000-2015

The 2015 simulation of global distributions was made using the following assumptions on the per capita income growth, national income distributions and population:

- 1. Per capita PCE were extrapolated from 2000 using the average 1990-2002 growth rates for all the regions except for ECE region where an increase was projected to compensate for the 1990-2002 declines.
- 2. National income distributions were left unchanged from 2000 for all the regions except for ECE region where some improvement was projected.
- 3. Population projections were taken from Human Development Report 2004.

Results are presented in ANNEX III on pp. 44-45.

 $\sim$ 

# SCENARIO II. PRO-POOR GROWTH

# WORLD INCOME DISTRIBUTION, 2015, Pro-poor growth simulation (2x rate for the poor)

Characteristics (estimation)	before	after	0
Income share: Decile 1	0.6%	0.7%	
Decile 2	1.2%	1.3%	
Decile 3	1.8%	1.8%	0
Decile 4	2.4%	2.4%	
Decile 5	3.2%	3.2%	
Decile 6	4.3%	4.3%	0
Decile 7	6.1%	6.1%	0
Decile 8	9.8%	9.8%	
Decile 9	18.8%	18.7%	
Decile 10	51.7%	51.6%	0
Upper boundary, decile 1	700	817	5
Upper boundary, decile 2	1,146	1,143	
Upper boundary, decile 3	1,593	1,590	
Upper boundary, decile 4	2,115	2,111	Ş
Median/Upper			2
boundary, decile 5	2,791	2,785	
Upper boundary, decile 6	3,804	3,796	
Upper boundary, decile 7	5,637	5,624	Ś
Upper boundary, decile 8	9,881	9,859	2
Upper boundary, decile 9	19,919	19,873	
Upper boundary, decile 10	1,000,000	997,700	
Gini-coefficient	0.6542	0.6510	Ś
Theil index	0.8633	0.8376	2
Theil index 2	0.8296	0.8226	
Variance (std.)	1.8340	1.8289	
Mode income	1059	863	9
Mean Income	7557	7557	2
Income less than mean	0.7566	0.7570	
Decile ratio	92.034	70.264	



 $\infty$ 

## Effect of pro-poor growth on poverty, 2015

	WORLD	OECD	LAC	EAP	SAS	AFR	ECE
Pro-poor growth							
poor (mln.)	436	0	25	15	10	386	0
incidence (%)	6.3%	0.0%	4.0%	0.7%	0.6%	37.6%	0.0%
Base scenario							
poor (mln.)	689	(0)	37	95	92	465	0
incidence (%)	10.0%	0.0%	6.0%	4.5%	5.3%	45.2%	0.0%

# WORLD INCOME DISTRIBUTION, 2000, before and after a hypothetical wealth transfer

Characteristics (estimation)	910190 1918nsti	t911£ T912R51	0.350	
Income share: Decile 1	0.5%	0.9%		
Decile 2	1.0%	1.3%		
Decile 3	1.5%	1.7%		
Decile 4	2.0%	2.1%	0.250 -	
Decile 5	2.7%	2.7%		
Decile 6	3.6%	3.6%		
Decile 7	5.3%	5.3%		
Decile 8	9.3%	9.3%	0.150 -	
Decile 9	19.7%	19.7%		
Decile 10	54.4%	53.4%		
Upper boundary, decile 1	450	630	0.050	
Upper boundary, decile 2	701	833	0000	
Upper boundary, decile 3	964	1,046		
Upper boundary, decile 4	1,279	1,299	10	100 10,000 10,000 100,000 100,000
Median/Upper			(0.050) -	
boundary, decile 5	1,700	1,702		
Upper boundary, decile 6	2,364	2,364		\$
Upper boundary, decile 7	3,666	3,666		
Upper boundary, decile 8	7,217	7,217	(0.150) -	
Upper boundary, decile 9	15,569	15,570		
Upper boundary, decile 10	1,000,000	929,634		
Gini-coefficient	0.6842	0.6684		
Theil index	0.9776	0.8834	(0.250) -	
Theil index 2	0.9116	0.8708		
Variance (std.)	1.9209	1.8626		
Mode income	633	783		
Mean Income	5,533	5,533	(0.350) -	
Income less than mean	0.7651	0.7651		
Decile ratio	103.270	59.007		
			(0.450)	

Conditions Marginal transfer rate Marginal tax rate

18.9% to people below \$1,400 threshold7.5% from the richest 5%

## Effect of wealth transfer on poverty, 2000

	WORLD	OECD	LAC	EAP	SAS	AFR	ECE
Wealth transfer							
poor (mln.)	784	0	27	239	168	348	С
incidence (%)	13.4%	0.0%	5.2%	12.9%	12.4%	45.1%	0.7%
Base scenario							
poor (mln.)	1,172	0	43	380	325	416	8
incidence (%)	20.0%	0.0%	8.3%	20.6%	24.0%	53.9%	1.7%

11 SCENARIO IV. DISTRIBUTION-NEUTRAL GROWTH, 1970-2000

### WORLD INCOME DISTRIBUTION, 2000



## WORLD INCOME DISTRIBUTION, 2000, distribution-neutral growth in 1999 Int. Dollars unless otherwise specified

CHARACTER	RISTICS	WORLD	OECD	LAC	EAP	SAS	AFR	E E
Income share:	Decile 1	0.6%	2.9%	0.9%	2.4%	2.4%	0.6%	2.7%
	Decile 2	1.2%	4.4%	1.8%	3.6%	3.9%	1.2%	4.3%
	Decile 3	1.7%	5.5%	2.6%	4.5%	5.0%	1.7%	5.6%
	Decile 4	2.2%	6.5%	3.6%	5.4%	6.1%	2.3%	6.7%
	Decile 5	2.9%	7.6%	4.9%	6.3%	7.3%	3.0%	8.0%
	Decile 6	3.8%	8.9%	6.4%	7.3%	8.7%	4.3%	9.3%
	Decile 7	5.4%	10.4%	8.6%	8.7%	10.4%	6.4%	10.9%
	Decile 8	9.3%	12.5%	11.8%	10.5%	12.6%	10.3%	13.0%
	Decile 9	19.8%	15.7%	17.5%	13.8%	16.3%	16.9%	16.0%
	Decile 10	53.1%	25.8%	41.9%	37.4%	27.5%	53.3%	23.6%
Upper boundary:	Decile 1	491	8,203	818	751	422	161	1,831
	Decile 2	785	10,644	1,336	976	575	244	2,508
	Decile 3	1,076	12,816	1,902	1,175	716	336	3,119
	Decile 4	1,403	15,079	2,570	1,379	864	451	3,717
Media	n/Decile 5	1,826	17,611	3,403	1,610	1,031	615	4,359
	Decile 6	2,471	20,609	4,499	1,890	1,228	885	5,097
	Decile 7	3,704	24,374	6,051	2,260	1,478	1,387	6,008
	Decile 8	7,185	29,601	8,534	2,810	1,832	2,246	7,239
	Decile 9	15,846	39,035	13,672	4,005	2,477	3,838	9,190
	Decile 10	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000
Gini-coefficient		0.6700	0.3394	0.5574	0.4428	0.3726	0.6666	0.3269
Theil index		0.9201	0.1914	0.5937	0.3309	0.2364	0.8945	0.1843
Theil index 2		0.8602	0.1951	0.5843	0.4196	0.2352	0.9607	0.1737
Variance (std.)		1.7861	0.6993	1.4971	1.3288	0.7733	2.6420	0.6226
Mean Income		5,533	21,526	6,104	2,383	1,299	1,731	5,071
Income less than	ı mean	0.7656	0.6268	0.7028	0.7268	0.6313	0.7454	0.5967
Decue ratio		96.022	8.744	48.842	15.291	11.337	84.238	8.811
CONTRIBUTION	OF REGIONS	S TO GLOBAL D	DECILES	0.40/	0.70/	2.20/	E 60/	0.00/
Decile 1		10%	0.0%	0.4%	0.7%	3.3%	5.0% 1.0%	0.0%
Decile 2		10%	0.0%	0.5%	4.9%	4.770	1.9%	0.0%
Decile 4		10%	0.0%	0.5%	4.3%	4.1%	0.7%	0.1%
Decile 5		10%	0.0%	0.3%	5.3%	2.9%	0.7%	0.2%
Decile 6		10%	0.0%	0.9%	5.0%	2.5%	0.8%	0.4%
Decile 7		10%	0.1%	1.3%	4 4%	1.6%	1.0%	1.6%
Decile 8		10%	1.0%	1.9%	2.3%	0.6%	0.9%	3.2%
Decile 9		10%	5.6%	1.5%	1.0%	0.1%	0.3%	1.6%
Decile 10		10%	8.7%	0.7%	0.4%	0.0%	0.1%	0.1%
Total		100%	15.4%	8.8%	31.5%	23.1%	13.1%	8.1%
Note: sums to we	orld deciles by	y row (correspon	ds to vertical slice	es of Figure B)				
REGIONAL POPU	JLATION FAI	LING INTO GLO	OBAL DECILES					
Decile 1		10%	0.0%	4.1%	2.2%	14.3%	42.9%	0.1%
Decile 2		10%	0.0%	5.2%	9.2%	20.5%	14.1%	0.6%
Decile 3		10%	0.0%	5.6%	13.6%	17.7%	7.7%	1.5%
Decile 4		10%	0.0%	6.2%	16.1%	14.8%	5.6%	2.8%
Decile 5		10%	0.0%	7.5%	16.9%	12.6%	5.4%	5.0%
Decile 6		10%	0.1%	9.9%	16.5%	10.1%	6.3%	9.5%
Decile 7		10%	0.4%	14.4%	14.0%	7.1%	7.5%	20.4%
Decile 8		10%	6.2%	22.2%	7.4%	2.8%	7.0%	39.9%
Decile 9		10%	36.5%	17.0%	3.0%	0.2%	2.4%	19.5%
Decile 10		10%	56.8%	7.8%	1.2%	0.0%	1.1%	0.9%
Total		100%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Note: sums to re	gional totals k	oy column						
COMPOSITION C	F WORLD'S	POOR BY REGI	ON	40.0	140 7	201.0	415 0	1.0
	Jow #1 400	998.4	0.0	40.0	149.7	391.9	415.0	1.8
be	700 : D	2,342.7	0.0	109.0	151.9	911.0	541.6	23.2
expenditures in T <u>Reference:</u>	Household Su	urveys	ылрепациres (SN	து terms approxin	uueiy correspond	ລ ເບ 1 <del>7 9</del> 3 ֆ1.08−0	и-шиу т verms of	πουσετιοία

	.005
25.7% 7	4.3%

	WORLD	LAC	EAP	SAS	AFR	ECE
pop. share, %						
700	17.02%	7.78%	8.09%	28.90%	53.79%	0.38%
1,400	39.93%	21.20%	40.95%	67.18%	70.20%	4.90%
inc. share, %						
700	1.31%	0.58%	1.87%	10.65%	10.21%	0.04%
1,400	5.60%	2.89%	16.51%	40.51%	19.59%	1.04%
inc. of group						
700	426	457	550	479	329	578
1,400	777	832	961	783	483	1,077
poverty gap						
700	39.1%	34.7%	21.4%	31.6%	53.1%	17.5%
1,400	44.5%	40.6%	31.4%	44.0%	65.5%	23.1%

AVERAGE INCOMES OF THE POOR AND POVERTY GAP, 2000, distr.-neutral scenario

### HALVING POVERTY

As poverty for OECD and ECE regions [defined as PPP \$700 in PCE terms] is negligible, the main attention was paid to other regions. Both absolute poverty and poverty incidence were studied for two scenarios: (1) distribution-neutral growth, 2000-2015, and (2) pro-poor growth, 2000-2015.

- (1) Distribution Neutral Growth, 2000-2015. For EAP and SAS regions the time needed to halve poverty was found to be 9 and 10 years, respectively, for poverty incidence, and 8 years for both regions for absolute poverty. For Africa and LAC, under the current assumptions, both absolute poverty and poverty incidence cannot be halved earlier than 30 years. At the global level, absolute poverty will be halved in 15 and poverty incidence in 24 years.
- (2) Pro-Poor Growth, 2000-2015. This scenario improves the situation markedly. For EAP and SAS regions the time needed to halve poverty was found to be 4 and 5 years, respectively, for poverty incidence, and 4 years for both regions for absolute poverty. For Africa and LAC, under the current assumptions, poverty incidence will be halved in 30+ and 18 years, respectively, and absolute poverty will be cut in half in 22 and 14 years, respectively. At the global level, absolute poverty will be halved in 9 and poverty incidence in 10 years.

	Poverty	y incidence		Absolu	te poverty
	DNG 2015	PPG 2015		DNG 2015	PPG 2015
WORLD	2	4	10	15	5 9
Latin America	30-	+	18	30+	- 14
East Asia		9	4	8	3 4
South Asia	1	0	5	8	3 4
Africa	30-	+ 3	30+	30+	- 22

### Time to halve poverty (years)

It is quite remarkable that positive growth does not automatically guarantee a decrease in absolute poverty. The Box below explains how this can be possible.

### BOX I. WHEN INCOME GROWTH CANNOT CROWD OUT POVERTY

### **Population vs. Income Growth**

This box illustrates the situation when income growth is not sufficient to decrease numbers of the poor. Two transformations are being applied to the original log-normal distribution (solid line): a horizontal shift equivalent to an income growth, and a scaling-up shift equivalent to a population growth. As elsewhere, the horizontal axis corresponds to income and the vertical axis is the distribution density function in logarithmic terms.

In the case when the poverty cut-off line is above the median, the income growth may not be enough to crowd out poverty. The exact results depend on the distribution characteristics. For example, for the log-normal distribution, if the following is true:

$$d\ln P > d\ln x \frac{F''_x(x_p)x_p}{\int_0^{x_p} F'_x dx}$$
, where  $x_p$  is the poverty line, and  $P$ 

is population, the income growth will not be sufficient to crowd out poverty.



similar way.

### ANNEX I. QUASI-EXACT POLYNOMIAL INTERPOLATION OF DISTRIBUTION FUNCTIONS

Described below are the general principles of the quasi-exact distribution rendering – the technique used for the current paper to convert national group means or interval data into a continuous functional form<sup>6</sup>. The foundation of the technique is based on a polynomial interpolation.

Let's assume that we are given only a set  $\{F(X_i)\}$  of M elements which describes values that the cumulative distribution function takes at  $X_i$ . We need to approximate all other points of the distribution, i.e., to estimate F(x) for  $x \in [0, +\infty]$ . Within each interval  $[X_{i+1}, X_i]$ , we will interpolate the distribution function by a polynomial of the order 3 (*cubic polynomial*) in the form:

$$F_{i,i+1}(x) = \sum_{n=0}^{3} \alpha_i^n \left(\frac{x - X_i}{X_{i+1} - X_i}\right)^n$$

At the boundaries the polynomials are *exact, and are not interpolations*:

i.e., 
$$F_{i,i+1}(X_i) = F_{i-1,i}(X_i) = F(X_i)$$
.

These polynomials are chosen to be *twice continuously differentiable* across the boundaries. This property allows for differential and integral operations with *F* and its derivatives in explicit analytical form. For example, the mean of the distribution would be calculated as follows:  $\mu = \int x dF = \sum_{i=0}^{M} \sum_{n=1}^{3} \alpha_i^n \frac{nX_{i+1} + X_i}{n+1}$ , where M is the number of intervals. Other characteristics of the distribution function can be derived in a

### PRECISION OF THE PROCEDURE

Using logic similar to that behind the remainder term of Taylor formula in Lagrange form, we arrive at the following expression for estimation errors<sup>7</sup>:

$$\|F_{i,i+1}(x) - F(x)\| \le \frac{1}{4!} \left(\frac{X_{i+1} - X_i}{2}\right)^4 \|F^{(4)}(\zeta)\|$$
  
where  $\zeta = \underset{x \in [X_i, X_{i+1}]}{\operatorname{arg\,max}} (\|F^{(4)}(x)\|)$ 

In the case of normal (standard) distribution the above boils down to:

<sup>6</sup> It should be noted that both tails of the distribution, i.e., the last and first group, were forced to be log-normal for the following reasons: (1) both tails are notoriously badly captured in the surveys, if at all; and (2) if the tails are captured, the biases present are different from the rest of the survey. Additionally, attempts were made to utilize household income distributions if personal income distributions were unavailable, using adjustment factor from other years. If distributional characteristics exhibited drastic changes, these data were discarded and a distribution from an adjacent year was used instead. If the restored distribution was found to be drastically non-uniform, the following was attempted: (1) smoothing of the distribution density curve, or (2) merging of the intervals used in the estimate.

<sup>&</sup>lt;sup>7</sup> Dividing interval  $[Y_{i+1}, Y_i]$  in half simply states that, because at the end of the interval the polynomial becomes exact again, maximum errors are attained around the middle of the interval. The coefficient 1/384  $[1/(2^4 \cdot 4)]$  is the absolute theoretical minimum for the errors. The minimum is attained when the polynomial coefficients for the interval are determined (almost) independently of other intervals. In other cases, the inequality is somewhat different, although the order of magnitude for errors remains the same.

$$\left\|F_{i,i+1}(x) - F(x)\right\| \le \frac{1}{384} (X_{i+1} - X_i)^4 \left\|F^{(1)}(\zeta)(3\zeta - \zeta^3)\right\|$$

(For details see Y. Dikhanov, 1996, "Decomposition of Inequality Based on Incomplete Information", World Bank).

In the case when the intervals are separated by  $\sigma/2$ , we obtain that the maximum errors will be in the interval  $[0.5\sigma, \sigma]$  (that can be seen from the first order condition for  $||F^{(1)}(\zeta)(3\zeta - \zeta^3)||$ ), and the errors in this interval are expressed as follows:

$$\left\|F_{i,i+1}(x) - F(x)\right\| \le \frac{1}{384 \cdot 16} \frac{2}{\sqrt{2\pi}} e^{-\frac{3-\sqrt{6}}{2}} \approx 0.01\%$$

Such a precision in interpolation usually exceeds the precision of survey reporting.

It is important to present distributions in the analytical form if one needs to carry out serious analysis of distribution necessary in computations of many functions or in modifying the distributions (for example, through income transfers or subsidies), say, in calculating  $\int \ln(\mu/x) dF(x)$  (for the Theil index), or of any  $\int \Psi(x) dF(x)$  in general (i.e., in integration of the distribution function with other functions). The following example shows importance of presenting distributions in the analytical form:

Let's assume that different income groups within a distribution face different price levels  $P=x/\xi(x)^8$ . Then the Theil index for real incomes (as opposite to nominal incomes) can be recalculated as:

$$\int \ln(\mu/\xi(\mathbf{x})) F'_{\xi} \xi'_{\mathbf{x}} d\mathbf{x}.$$

Additionally, if we can analytically express distribution function F(x), then we can directly calculate all distribution characteristics, such as mean, median, mode, dispersion, various inequality measures, etc.

### "GOOD" CASE FOR INTERPOLATION

As a "good" case, we used ten income intervals for the log-normal distribution LN(5,0.25).

The results are presented in the table below. Graphical results are presented in Figure 1. As can be seen from the graph, the actual distribution cannot be readily distinguished from the simulation. The largest difference is for the mode, which is notoriously difficult to get right.

	Actual	Simulation	Difference
	values		
Mean Income	153.12	153.09	-0.02%
Gini-coefficient	0.14032	0.14023	-0.06%
Median Income	148.41	148.41	0.00%
Mode Income	139.42	139.97	0.39%
Variance	38.887	38.923	0.09%
Income less than mean	0.5497	0.5494	-0.06%
Theil index	0.03125	0.03123	-0.07%
Theil index 2	0.03125	0.03126	0.03%

8 Not an unreasonable assumption, this phenomenon is observed in many household surveys.



Figure 1. Deviation of simulation from actual distribution: a "good" case

### **"BAD" CASE FOR INTERPOLATION**

As a "bad" case, we used only five income intervals for the mixture of two normal distributions N(40,10) and N(60,5). The results are presented in the tables below. Graphical results are presented in Figure 2. As can be seen from the graph, the actual distribution is visually readily distinguishable from the simulation. The largest difference is again for the mode. However, when ten intervals are used, the results of interpolation are much closer to the original distribution.

Inputs into the procedure

Interval boundaries	Quintiles of population
< 37.4696	Quintile I
37.4696 to 48.10972	Quintile II
48.10972 to 56.60144	Quintile III
56.60144 to 61.47081	Quintile IV
> 61.47081	Quintile V

### Results of the simulation

	Actual values	Simulation	Difference
Mean Income	50.00	49.67	-0.7%
Median Income	53.33	53.23	-0.2%
Mode Income	59.64	58.87	-1.3%
Income less than mean	43.20%	42.82%	-0.9%

(1)

(3)



Figure 2. Deviation of simulation from actual distribution: "bad" case

### COMPARISON TO UNIMODAL FITTING PROCEDURES

An alternative to the quasi-exact interpolation is a unimodal fitting procedure, for example, a quadratic fit for the *Lorenz* curve  $L(F)^9$ . The General Quadratic Lorenz curve (Villasenor, I., and B.C. Arnold, 1989. <u>Elliptical Lorenz Curves</u>, *Journal of Econometrics* 40 (2): 327-338) is used in this analysis. However, the results are similar to using other approaches that fit the Lorenz curves directly. Naturally, only *Lorenz*-type data [means and population shares] can be used in the approximation. The interval data cannot be used directly.

The quality of fit can only be as good as that of the underlying implicit distribution function, especially in modeling multi-modal distribution such as the global distribution. The fit's R<sup>2</sup> could be normally very high: 0.999999 or so. However, the Lorenz curve fits in general suffer of poor resolution [see, e.g., Figures A, B, C and D for quadratic interpolation from below – a simulation of a two-modal distribution – the Lorenz curve

 $F'(x) = f(x) = A((x - x_{min})^2 + A)^{-3/2}$ where x is income level;  $x_{min}$  is minimum income; A is a coefficient.

Correspondingly, after some more simple calculus, the cumulative distribution function F(x) can be written as:

$$F(x) = (1 + A / (x - x_{min})^2)^{-1/2}$$
<sup>(2)</sup>

Or, noticing that  $A=1/3*(x_{median}-x_{min})^2$  [because  $F(x_{median})=1/2$ ], we can simplify expression (2) further to arrive at:

$$F(x) = \frac{1}{\sqrt{1 + \frac{1}{3} \left(\frac{x_{median} - x_{min}}{x - x_{min}}\right)^2}}$$

Then expression (3) is implicitly rescaled to fit a given mean by introducing a third paramether and effectively truncating the distribution at both ends– see plots C and D to this effect. The rescaling makes actual formulas somewhat more complicated. But essentially functions equivalent to expressions (2) and (3) determine the behaviour of such distributions. Additionally, xmin – "minimum" income - can often be negative.

<sup>9</sup> Such a fit can be described as:

can fit very closely, but the predictive power of all distribution moments can be poor]. A small error in the fit of L(F), which is two integrations above the distribution density function f(x), can result in large errors for f(x), especially, in the cases where actual distribution is more complex, or when it's a summary distribution composed of several individual ones. At the same time the polynomial interpolation shows much better precision [see Figures A, B, C and D for polynomial interpolation from below].

### COMPARISON OF GENERAL QUADRATIC [GQ] INTERPOLATION TO ACTUAL DISTRIBUTION



COMPARISON OF "QUASI-EXACT" POLYNOMIAL INTERPOLATION TO ACTUAL DISTRIBUTION

A. Lorenz curve







**B.** Cumulative distribution f-n F(x)



**D.** Distribution density f-n F'(x) [log]



### ANNEX II. DECOMPOSITION OF INEQUALITY MEASURES

### 1. DECOMPOSITION OF GINI - COEFFICIENT

Let's consider a distribution F defined by its cumulative distribution function F(y). The respective distribution density function is F'. The mean of that distribution is defined as  $\mu_i = \int y dF_i(y)$  using Lebesgue-Stiltjes integrals. (Hereinafter a plain integral sign describes integrating from 0 to  $+\infty$ ). Then the essence of the Gini - coefficient can be seen from the graph of the Lorenz curve (see Figure 3).



Gini-coefficient is defined as equal to twice the area between the 45° line and Lorenz curve. Or

$$G = 1 - \frac{2}{\mu} \int (\int_{0}^{r} y d\Phi) dF,$$

or,

$$G = \frac{2}{\mu} \int Fd(\int_{0}^{F} yd\Phi) - 1 = \frac{2}{\mu} \int FydF - 1$$

Let's consider two distributions  $F_i$  and  $F_2$ , where the distributions are defined by their respective cumulative distribution functions  $F_i(y)$ . The respective distribution density functions are  $F_i$ . Means are defined as  $\mu_i = \int y dF_i$ . Thus, we can define Gini - coefficients G for the respective functions as follows:

$$G_{1} = \frac{2}{\mu_{1}} \int y(F_{1} - \frac{1}{2})F_{1}'dy = \frac{2}{\mu_{1}} \int F_{1}y F_{1}'dy - 1$$

$$G_{2} = \frac{2}{\mu_{2}} \int y(F_{2} - \frac{1}{2})F_{2}'dy = \frac{2}{\mu_{2}} \int F_{2}y F_{2}'dy - 1$$
(1)

21

Then, for the combined distribution we can write:

$$G = \frac{2}{\rho_1 \mu_1 + \rho_2 \mu_2} \int y(\rho_1 F_1 + \rho_2 F_2 - \frac{1}{2})(\rho_1 F_1' + \rho_2 F_2')dy$$
(2)

where:

$$\begin{split} \omega_i &= \frac{p_i \mu_i}{p_1 \mu_1 + p_2 \mu_2} \text{ - income share of the } i \text{ distribution} \\ p_i &= \text{ oppulation share} \\ \mu &= \omega_1 \mu_1 + \omega_2 \mu_2 & \text{ - mean income for the combined distribution} \end{split}$$

Or, after some simple operations we will receive:

$$G = \omega_1 G_1 + \omega_2 G_2 - \frac{2p_1 p_2}{\mu} \int y(F_1 - F_2)(F_1' - F_2') dy$$
(3)

Expression (3) is obtained as follows:

$$G = \frac{2}{p_1\mu_1 + p_2\mu_2} \int y(p_1F_1 + p_2F_2)(p_1F_1' + p_2F_2')dy - 1 =$$

$$= \frac{2}{p_1\mu_1 + p_2\mu_2} \int y[p_1^2F_1F_1' + p_2^2F_2F_2' + p_1p_2(F_1F_2' + F_2F_1')]dy - 1 =$$

$$= \frac{2}{\mu} \int y[p_1^2F_1F_1' + p_2^2F_2F_2' + p_1p_2(F_2'(F_2 + F_1 - F_2) + F_1'(F_1 + F_2 - F_1))]dy - 1 =$$

$$:= \frac{2}{\mu} \int y[(p_1^2 + p_1p_2)F_1F_1' + (p_2^2 + p_1p_2)F_2F_2' + p_1p_2(F_2' - F_1')(F_1 - F_2)]dy - 1 =$$

$$= \frac{2}{\mu} \int y[p_1F_1F_1' + p_2F_2F_2' + p_1p_2(F_2' - F_1')(F_1 - F_2)]dy - 1 =$$

and, because  $p_i = \omega_i \frac{\mu}{\mu_i}$  $G = \omega_1 G_1 + \omega_2 G_2 + \frac{2p_1 p_2}{\mu} \int y(F_2' - F_1')(F_1 - F_2) dy$  It is easy to see how the above expression can be expanded for a multi-component case:

$$G = \frac{2}{\sum_{i} p_{i} \mu_{i}} \int y \sum_{i} p_{i} F_{i} \sum_{i} p_{i} F_{i}' dy - 1 =$$

$$= \frac{1}{\sum_{i} p_{i} \mu_{i}} \int y \{2 \sum_{i} [F_{i} F_{i}'(p_{i}^{2} + \sum_{j \neq i} p_{i} p_{j})] - \sum_{i,j} p_{i} p_{j} (F_{i}' - F_{j}')(F_{i} - F_{j})\} dy - 1 =$$

$$= \frac{1}{\sum_{i} p_{i} \mu_{i}} \int y \{2 \sum_{i} F_{i} F_{i}' p_{i} - \sum_{i,j} p_{i} p_{j} (F_{i}' - F_{j}')(F_{i} - F_{j})\} dy - 1$$

The above expression can be rewritten as follows:

$$G = \sum_{i} \omega_{i} G_{i} - \sum_{i,j} \frac{p_{i} p_{j}}{\mu} \int y(F_{i}' - F_{j}')(F_{i} - F_{j}) dy$$
(4)

And, as it is easy to see how the Gini-coefficient can be expressed through the covariance as well:

$$G_i = \frac{2}{Y_i} COV(y, F_i)$$

and the combined Gini-coefficient can be written as:

$$G = \sum_{i} \frac{2\omega_i}{\mu_i} COV(y, F_i) - \sum_{i,j} \frac{p_i p_j}{\mu} COV(y, F_i - F_j)$$
(5)

Or,

$$G = \frac{1}{\mu} \{ 2\sum_{i} p_{i} COV(y, F_{i}) - \sum_{i,j} p_{i} p_{j} COV(y, F_{i} - F_{j}) \}$$

The first component stands for intra-group covariances, whereas the second stands for inter-group covariance.

As we can see from expression (3), the Gini - coefficient for the combined distribution consists of two parts: intra-group and inter-group variances. Similar to the Theil coefficient T1, the individual Gini - coefficients are added up with income weights.

### 2. DECOMPOSITION OF ENTROPY (THEIL) INDEXES

In his book, H. Theil (1967, *Economics and Information Theory*, North-Holland, Amsterdam), introduced, for income inequality measurement, the entropy measure used in thermodynamics and information theory. He suggested using the entropy index in two forms: as income-weighted and population-weighted entropy indexes. In this paper we will call them *T1* and *T2* respectively.

These indexes can be represented as follows:

$$T1 = \sum_{i} \frac{Y_{i}}{Y} \log(\frac{Y_{i}}{Y} / \frac{N_{i}}{N})$$
$$T2 = \sum_{i} \frac{N_{i}}{N} \log(\frac{N_{i}}{N} / \frac{Y_{i}}{Y})$$

where,

 $\begin{array}{ll} Y_i & \text{ is income of group } i; \\ N_i & \text{ is number of people in group } i \end{array}$ 

Or, using Lebesgue-Stiltjes integrals:

$$T1 = \int \frac{y}{\mu} \log(\frac{y}{\mu}) dF(y)$$
$$T2 = \int \log(\frac{\mu}{y}) dF(y)$$

As can be shown, these indexes are easily decomposable in the multi-group case. For the Theil index T1 we have:

$$T1_{i} = \sum_{j} \frac{Y_{i}^{j}}{Y_{i}} \log(\frac{Y_{i}^{j}}{Y_{i}} / \frac{N_{i}^{j}}{N_{i}})$$

where:

 $Y_i$  is income of sub-group *j* of group *i*;

 $N_i$  is number of people in sub-group *j* of group i;

- $Y_i$  is income of group *i*;
- $N_i$  is number of people in group i

Or, using Lebesgue-Stiltjes integrals:

$$T1_i = \int \frac{y}{\mu_i} \log(\frac{y}{\mu_i}) dF_i(y)$$

The Theil index *T1* decomposes into:

$$T1 = \sum_{i} \omega_{i} T1_{i} + \sum_{j} \omega_{i} \log(\frac{\omega_{i}}{p_{i}}) = \sum_{i} \omega_{i} T1_{i} + \sum_{j} \omega_{i} \log(\frac{\mu_{i}}{\mu})$$

T2 decomposes in a similar way with the population weights *p*.

As has been shown by F. Bourguignon (1979, *Decomposable Income Inequality Measures*, Econometrica, Vol. 47, No. 4.), and A.F.Shorrocks (1980, Inequality Measures, Econometrica, Vol. 48, No 3), the Theil indexes are the only income-weighted and population-weighted indexes respectively that can be decomposed in that way: i. e., weighted sum of individual Theil indexes and the Theil index constructed of individual distributions as if they were elements of the combined distribution. In this sense, the decomposition of the Theil indexes is different from that of the Gini.

### 3. DECOMPOSITION OF NORMALIZED VARIANCE

Normalized variance can be seen as a simple way of describing income inequalities.

$$s^{2}\left(\frac{y}{\mu}\right) = \sum_{i,j} p_{i} p_{j} COV\left(\frac{y_{i}}{\mu}, \frac{y_{j}}{\mu}\right) = \sum_{i,j} \omega_{i} \omega_{j} COV\left(\frac{y_{i}}{\mu_{i}}, \frac{y_{j}}{\mu_{j}}\right)$$
  
Or,  
$$s^{2}\left(\frac{y}{\mu}\right) = \sum_{k} \omega_{k}^{2} s^{2}\left(\frac{y_{k}}{\mu_{k}}\right) + \sum_{i \neq j} \omega_{i} \omega_{j} COV\left(\frac{y_{i}}{\mu_{i}}, \frac{y_{j}}{\mu_{j}}\right)$$

### 4. DECOMPOSITION OF DECILE RATIO

Decile ratio is a simple and transparent inequality measure, however it cannot be meaningfully decomposed.

### 5. LORENZ CURVE

The Lorenz function L is the function of income shares on population shares. The Lorenz curve associated with this function is plotted in Figure 3. The Lorenz curve plays an enormous role in income distribution analysis. Some important relationships between the Lorenz curve and the cumulative distribution function, as well as a graphical representation of the Theil index, are shown below.



Figure 4. First derivative of the Lorenz curve

Figure 4 shows the first derivative of the Lorenz curve, L'(F). It can be easily seen that L'(F) is essentially the normalized income  $y/\mu$ , and, thus, is the inverse (normalized) of the cumulative distribution function. The graph is also related to the Theil (T2) index. The logarithm of this graph is a graphical representation of the index (because the index can be presented as  $T2 = \int \log(\frac{\mu}{v}) dF(y)$ .



Figure 5. Graphical representation of the Theil index (T2)

The second derivative of the Lorenz curve is also an important characteristic of a distribution:  $L''(F) = y'_F/\mu$ . It is essentially the expression for the inverse function of a distribution density function F'(y).



Figure 6. Second derivative of the Lorenz curve

### 6. SOME PROPERTIES OF LOG-NORMAL DISTRIBUTION

Log-normal distribution plays an important role in inequality measurements. It is thought that real distributions of wealth and income at least partially can be approximated by it. An extensive treatment of the log-normal distribution is contained in J. Aitchison and J.Brown (1957, *The Lognormal Distribution*, Cambridge University Press). Here we mention just a few relevant properties.

$$F'(y) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-(\ln y - \theta)^2}{2\sigma^2}} \frac{1}{y}$$
$$\mu = e^{\theta + \sigma^2/2}$$
$$Median = e^{\theta}$$
$$Mode = e^{\theta - \sigma^2}$$
$$S = e^{\theta} \sqrt{e^{\sigma^2} (e^{\sigma^2} - 1)}$$
$$s = \frac{S}{\mu} = \sqrt{(e^{\sigma^2} - 1)}$$
$$E(z^m) = e^{m^2 \frac{\sigma^2}{2}}$$
$$where$$

$$z = \frac{\ln z - \theta}{\sigma}$$

A convenient feature of the log-normal distribution is the simplicity of the Gini and Theil indexes:

$$T1 = \int (\ln y - \ln \mu) dL$$
  

$$T1 = \frac{1}{\mu} \int \frac{\ln y - \ln \mu}{\sqrt{2\pi\sigma}} e^{-\frac{(\ln y - \theta)^2}{2\sigma^2}} dy =$$
  

$$= \frac{1}{\mu} \int y \frac{\ln y}{\sqrt{2\pi\sigma}} e^{-\frac{(\ln y - \theta)^2}{2\sigma^2}} d\ln y - (\theta + \sigma^2 / 2) =$$
  

$$= \frac{1}{\mu} \int \frac{te^t}{\sqrt{2\pi\sigma}} e^{-\frac{(t - \theta)^2}{2\sigma^2}} dt - (\theta + \sigma^2 / 2) =$$
  

$$= \frac{1}{\mu} e^{\theta + \sigma^2 / 2} \int \frac{t}{\sqrt{2\pi\sigma}} e^{-\frac{(t - (\theta + \sigma^2))^2}{2\sigma^2}} dt - (\theta + \sigma^2 / 2) =$$
  

$$= \theta + \sigma^2 - (\theta + \sigma^2 / 2) = \sigma^2 / 2$$

And, in the case of the second Theil index, we can obtain the following expression:

$$T2 = \int (\ln \mu - \ln y) dF$$
  

$$T2 = -\int \frac{\ln y}{\sqrt{2\pi\sigma}} e^{-\frac{(\ln y - \theta)^2}{2\sigma^2}} d\ln y + (\theta + \sigma^2 / 2) =$$
  

$$= -\theta + (\theta + \sigma^2 / 2) = \sigma^2 / 2$$

We can use the test of T1=T2 to examine how close a given distribution approaches a log-normal one.

The relationship of the Theil measures to normalized variance can be expressed as follows:

$$T1 = T2 = \frac{\sigma^2}{2} = \frac{1}{2}\ln(1+s^2)$$

In the log-normal case, we can also think of the Theil indexes as the difference between the mean and median.

$$T1 = T2 = \frac{\sigma^2}{2} = \log(\frac{\mu}{Median})$$

And, finally, as can be easily seen, the Gini coefficient for the log-normal distribution can be written as follows:

$$G = 2\Phi(\sigma/\sqrt{2}) - 1 = 2F(e^{\theta + \sigma^2/\sqrt{2}}) - 1$$

where  $\Phi(.)$  is the standard normal cumulative distribution.

### 7. TWO ALTERNATIVE REPRESENTATIONS OF THE GINI COEFFICIENT

Apart from the traditional visualization of the Gini index using the Lorenz curve, it is possible to represent the Gini using simple graph of the distribution function. Below two such representations are discussed.

1. Let's start from the following expression for the Gini coefficient:

$$G = \frac{2}{\mu} \int y(F - \frac{1}{2}) dF$$
 (6)

Or, as it is easy to see, expression (6) can be written as:

$$G = \frac{2}{\mu} \int (y - \mu) F dF$$
(7)

Expressions (6) and (7) are equivalent to:

$$G = \frac{2}{\mu} Cov(y, F)$$
(8)

We can rewrite Expression (8) using slope coefficient as follows:

$$G = \frac{2}{\mu} \frac{\int (y-\mu)(F-\frac{1}{2})dF}{\int (F-\frac{1}{2})^2 dF} \int (F-\frac{1}{2})^2 dF = \frac{2}{\mu} Slope(F,y)\frac{1}{12}$$
  
because  $\sigma_F^2 = \int (F-\frac{1}{2})^2 dF = \frac{F^3}{3} - \frac{F^2}{2} + \frac{F}{4} \Big|_0^1 = \frac{1}{12}$  (9)  
where *Slope* = slope coefficient<sup>10</sup>

where Slope = slope coefficient

$$G = \frac{1}{6}Slope(F, y_{\mu})$$
(10)
where  $y_{\mu} = \frac{y}{\mu}$ 

Expression (10) can be obtained from Expression (8) in a different way as well. Let's start from rewriting Expression (8) using the correlation coefficient<sup>11</sup>:

$$G = \frac{2}{\mu} Cov(y,F) = \frac{2}{\mu} \rho(y,F) \sigma_y \sigma_F = \frac{1}{\sqrt{3}} \frac{\sigma_y}{\mu} \rho(y,F)$$
(11)

because  $\sigma_F^2 = \frac{1}{12}$ , [see Expression (9)].

<sup>10</sup> 
$$Slope(x,y) = \frac{\sum (x_i - E(x))(y_i - E(y))\omega_i}{\sum (x_i - E(x))^2 \omega_i}$$
, where  $\omega$  *i* are weights, or, in continuous case,  
 $Slope(x,y) = \frac{\int (x - E(x))(y - E(y))dF}{\int (x - E(x))^2 dF}$ 

<sup>11</sup> Discrete case of using correlation coefficients in expressing Gini coefficient [Expression (11)] was shown in Milanovic (1996)

Now, using  $\rho(y,F) = Slope(y,F) \frac{\sigma_F}{\sigma_y}$ , we obtain Expression (10) again:

$$G = \frac{1}{\sqrt{3}} \frac{\sigma_y}{\mu} Slope(F, y) \frac{\sigma_F}{\sigma_y} = \frac{1}{6} Slope(F, y_{\mu})$$



Figure 7. Graphical representation of the Gini coefficient as one sixth of the slope coefficient between income  $y_{\mu}$  and distribution function F.

2. The next representation of the Gini coefficient can be obtained using Expression (7):

$$G = \frac{2}{\mu} \int (y - \mu) F dF = \frac{2}{\mu} \int y F dF - 1 = \int y_{\mu} dF^2 - 1 = \frac{\mu(F^2)}{\mu} - 1$$
(12)

where  $\mu(F^2) = \int y dF^2$ , or the mean for the square of distribution *F*. Or, equivalently:

$$G = \int y_{\mu} dF^{2} - \int y_{\mu} dF = \int F dy_{\mu} - \int F^{2} dy_{\mu}$$
(13)

It is easy to see that distribution  $F^2$  has all the properties of a regular distribution.  $F^2$  is a monotonous transformation of F, and, hence, is itself a monotonously increasing function bounded by [0,1].

Expression (12) essentially says that the Gini coefficient is equal to the difference between regular mean  $\mu$  and the mean for the square of distribution  $\mu$ ( $F^2$ ). The expression is presented in Figure 8 in

30

graphical form. In the case when income normalized by the mean, the Gini coefficient is equal to the area between the distribution function F and the squared distribution function  $F^2$ .



Figure 8. Graphical representation of the Gini coefficient as the area between the distribution function F and the squared distribution function  $F^2$ .

### REFERENCES

Aitchison J. and J.Brown, 1957, *The Lognormal Distribution*, Cambridge University Press, Cambridge.

Bourguignon F., 1979, Decomposable Income Inequality Measures, Econometrica, Vol. 47, No. 4.

Dikhanov Y., 1999, International Income Distribution, World Bank Policy and Research Bulletin, v. 10, n. 3.

Dikhanov Y. and M. Ward, *Evolution of the Global Distribution of Income*, Review of Income and Wealth (forthcoming).

Dikhanov Y., 1996, *Decomposition of Inequality Based on Incomplete Information*, paper presented at the IARIW 24th General Conference, Lillehammer, Norway, August 18-24, 1996

Shorrocks A.F., 1980, Inequality Measures, Econometrica, Vol. 48, No 3.

Theil H., 1967, Economics and Information Theory, North-Holland, Amsterdam.

Theil H.,1989, *Development of International Inequality*, Journal of Econometrics, Vol. 42, No. 1, North-Holland, Amsterdam.

Villasenor, I., and B.C. Arnold, 1989. *Elliptical Lorenz Curves*, Journal of Econometrics 40 (2): 327-338

### ANNEX III. WORLD INCOME DISTRIBUTION, 1970-2015

33

Tables and Figures

### ABSOLUTE POVERTY, IN MLN.

	1970	1980	1990	2000	2015
below \$700 1999 PPP terms *)					
World	1,408.6	1,543.8	1,355.3	1,172.3	689.2
of which:					
Latin America	43.2	31.6	43.4	42.7	37.4
East Asia	784.2	798.4	529.9	380.4	95.4
South Asia	424.6	516.6	494.2	325.2	91.7
Africa **)	155.8	196.8	287.5	416.1	464.5
East and Central Europe	0.8	0.3	0.2	7.9	0.2
below \$1400 1999 PPP terms					
World	2,024.3	2,377.4	2,548.6	2,537.8	1,778.1
of which:					
Latin America	93.3	82.6	111.1	115.1	109.2
East Asia	1,041.7	1,191.3	1,127.7	911.3	417.4
South Asia	637.9	795.0	895.9	903.4	560.9
Africa	241.9	301.9	407.9	552.8	686.5
East and Central Europe	8.8	6.4	5.9	55.2	4.0

\*) \$700 in 1999 PPP terms approximately corresponds to the \$1-a-day cut-off used by the World bank, when adjusted for 1985-99 inflation and the differences between the Personal Consumption Expenditures (PCE) and expenditures recorded in household surveys.

\*\*) of which Sub-Saharan Africa [w/o Sudan] -- 145, 183, 267, 286 and 430 mln., respectively.

**INCOME GROWTH BY DECILE, 1970-2015** 

(in 1999 PPP dollars)					
			1990-		1970-2000
	1970-80	1980-90	2000	2000-15	cumulative
Decile 1	7.5%	18.4%	11.6%	45.9%	42.1%
Decile 2	9.6%	27.1%	20.7%	60.8%	68.1%
Decile 3	10.5%	29.7%	23.2%	64.7%	76.6%
Decile 4	10.4%	29.4%	23.7%	65.5%	76.8%
Decile 5	8.9%	26.8%	21.7%	64.9%	68.1%
Decile 6	5.3%	16.5%	16.2%	62.6%	42.6%
Decile 7	5.1%	-1.4%	1.6%	57.7%	5.3%
Decile 8	13.7%	-0.1%	-9.6%	45.0%	2.6%
Decile 9	19.4%	8.4%	0.9%	30.1%	30.6%
Decile 10	20.7%	18.6%	11.2%	29.9%	59.2%
Mean Income	17.1%	13.2%	7.1%	36.6%	42.0%
Median Income	7.9%	23.9%	19.9%	64.2%	60.2%

### **INCOME LEVELS BY DECILE, 1970-2015**

(in 1999 PPP dollars)

	1970	1980	1990	2000	2015
Decile 1	205	220	261	291	425
Decile 2	343	376	478	577	928
Decile 3	470	519	673	829	1,365
Decile 4	630	696	901	1,115	1,846
Decile 5	878	957	1,213	1,477	2,436
Decile 6	1,404	1,478	1,723	2,002	3,256
Decile 7	2,778	2,920	2,879	2,926	4,615
Decile 8	4,999	5,682	5,676	5,129	7,439
Decile 9	8,348	9,964	10,800	10,901	14,183
Decile 10	18,895	22,808	27,057	30,081	39,081
Mean Income	3,895	4,562	5,166	5,533	7,557
Median Income	1,061	1,144	1,418	1,700	2,791

### VARIOUS GLOBAL INEQUALITY CHARACTERISTICS

changes			1990-		1970-2000
	1970-80	1980-90	2000	2000-15	cumulative
Gini-coefficient	0.014	0.004	-0.001	-0.030	0.016
Theil index 1	0.057	-0.032	-0.043	-0.114	-0.018
Theil index 2	0.040	0.030	0.021	-0.082	0.090
Income less than mean	0.010	0.023	0.027	-0.009	0.060
Decile ratio	11.30	0.19	-0.41	-11.32	11.08
absolute levels					
	1970	1980	1990	2000	2015
Gini-coefficient	0.668	0.681	0.685	0.684	0.654
Theil index 1	0.996	1.053	1.021	0.978	0.863
Theil index 2	0.822	0.862	0.891	0.912	0.830
Income less than mean	0.705	0.716	0.738	0.765	0.757
Decile ratio	92.28	103.58	103.77	103.36	92.03

Decomposition of inequality using Theil II index								
	1970	1980	1990	2000	2015			
Within-country inequality	0.211	0.220	0.230	0.275	0.276			
	25.7%	25.5%	25.8%	30.2%	33.3%			
Between-country inequality	0.610	0.642	0.661	0.637	0.554			
	74.3%	74.5%	74.2%	69.8%	66.7%			

Absolute Poverty, in mln. (below \$700 1999 PPP terms)











WORLD INCOME DISTRIBUTION, 1970



CHARACTER	NISTICS	WORLI	OECD	LAC	EAP	SAS	AFR	ECE
Income share:	Decile 1	0.5%	2.8%	0.8%	2.4%	2.3%	0.7%	3.3%
	Decile 2	0.9%	4.2%	1.6%	3.6%	3.7%	1.4%	4.9%
	Decile 3	1.2%	5.3%	2.5%	4.4%	4.8%	2.0%	6.0%
	Decile 4	1.6%	6.3%	3.6%	5.3%	6.0%	2.8%	7.1%
	Decile 5	2.3%	7.4%	4.8%	6.2%	7.2%	3.8%	8.1%
	Decile 6	3.6%	8.7%	6.4%	7.3%	8.6%	5.0%	9.3%
	Decile 7	7.1%	10.3%	8.7%	8.7%	10.4%	6.7%	10.7%
	Decile 8	12.8%	12.5%	12.0%	10.7%	12.7%	9.1%	12.5%
	Decile 9	21.4%	15.8%	18.0%	14.5%	16.5%	13.8%	15.4%
	Decile 10	48.5%	26.6%	41.6%	36.9%	27.8%	54.7%	22.7%
Upper boundary:	Decile 1	282	4,372	515	237	236	202	2,202
	Decile 2	403	5,734	869	308	324	317	2,847
	Decile 3	541	6,951	1,264	371	406	451	3,392
	Decile 4	733	8,218	1,737	437	494	614	3,918
Media	n/Decile 5	1,061	9,644	2,331	511	593	818	4,483
	Decile 6	1,899	11,350	3,118	604	710	1,081	5,134
	Decile 7	3,780	13,518	4,235	730	859	1,439	5,950
	Decile 8	6,379	16,562	6,010	926	1,072	2,019	7,092
	Decile 9	10,886	22,113	9,627	1,383	1,457	3,431	8,980
	Decile 10	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000
Gini-coefficient		0.6677	0.3516	0.5609	0.4437	0.3799	0.6487	0.2984
Theil index		0.9961	0.2060	0.6132	0.3293	0.2471	0.8134	0.1488
Theil index 2		0.8216	0.2105	0.5801	0.4026	0.2435	0.9776	0.1466
Variance (std.)		1.6495	0.7321	1.4339	1.2733	0.7836	2.8427	0.5793
Income less than	mean	<b>3,895</b> 0,7052	11,992 0.6326	<b>4,199</b> 0.6974	765 0 7217	<b>753</b> 0.6319	<b>1,877</b> 0 7809	<b>5,170</b>
Decile ratio	mean	92.281	9.463	54.315	15.168	11.978	75.136	6.956
CONTRIBUTION	OF REGIONS	TO GLOBAL D	ECILES					
Decile 1		10%	0.0%	0.3%	5.1%	3.0%	1.6%	0.0%
Decile 2		10%	0.0%	0.2%	6.0%	2.9%	0.9%	0.0%
Decile 3		10%	0.0%	0.3%	5.8%	3.0%	0.9%	0.0%
Decile 4		10%	0.0%	0.4%	5.3%	3.3%	1.0%	0.0%
Decile 5		10%	0.0%	0.7%	4.5%	3.5%	1.3%	0.1%
Decile 6		10%	0.1%	1.4%	3.0%	3.1%	1.8%	0.6%
Decile 7		10%	1.2%	1.8%	1.4%	0.9%	1.2%	3.5%
Decile 8		10%	3.8%	1.2%	0.4%	0.1%	0.4%	4.2%
Decile 9		10%	6.5%	0.8%	0.1%	0.0%	0.2%	2.4%
Decile 10		10%	8.6%	0.6%	0.0%	0.0%	0.2%	0.5%
Total	rld daoilas h	100%	20.2%	7.7%	31.6%	19.6%	9.6%	11.3%
	na aecues by	, row (correspon		з ој гідиге Бј				
Decile 1	LATION FAL	10%	DBAL DECILES	3 7%	16.2%	15.1%	17 1%	0.0%
Decile 2		10%	0.0%	3.1%	18.9%	14.6%	9.6%	0.0%
Decile 3		10%	0.0%	3,9%	18.4%	15.3%	9.1%	0.0%
Decile 4		10%	0.0%	5.5%	16.7%	16.8%	10.3%	0.2%
Decile 5		10%	0.0%	8.8%	14.2%	17.9%	13.2%	0.6%
Decile 6		10%	0.4%	17.9%	9.5%	15.7%	19.1%	5.5%
Decile 7		10%	6.1%	23.4%	4.4%	4.3%	12.8%	31.1%
Decile 8		10%	18.8%	15.1%	1.3%	0.3%	4.2%	36.8%
Decile 9		10%	32.2%	10.4%	0.4%	0.0%	2.2%	20.9%
Decile 10		10%	42.5%	8.1%	0.1%	0.0%	2.4%	4.8%
Total		100%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Note: sums to reg	gional totals b	y column						
COMPOSITION O	F WORLD'S	POOR BY REGI	ON					
b	elow \$700	1,408.6	(0.0)	43.2	784.2	424.6	155.8	0.8
be	low \$1,400	2,024.3	0.7	93.3	1,041.7	637.9	241.9	8.8
Note: 1999 Int.\$7 expenditures in 1 Reference:	700 in Persond Household Su	al Consumption rveys	Expenditures (SN	A) terms approxin	nately correspond	ls to 1993 \$1.08-0	a-day in terms of	household

Within-country inequality	0.211	Between-country inequality	0.610
	25.7%		74.3%

WORLD INCOME DISTRIBUTION, 1980



\_

### WORLD INCOME DISTRIBUTION, 1980

CHARACTER	ISTICS	WORLD	OECD	LAC	EAP	SAS	AFR	ECE
Income share:	Decile 1	0.5%	2.9%	0.8%	2.1%	2.3%	0.7%	3.2%
	Decile 2	0.8%	4.4%	1.8%	3.1%	3.7%	1.4%	4.9%
	Decile 3	1.1%	5.5%	2.6%	3.9%	4.8%	2.1%	6.0%
	Decile 4	1.5%	6.5%	3.7%	4.7%	5.9%	2.9%	7.0%
	Decile 5	2.1%	7.6%	4.9%	5.6%	7.1%	4.0%	8.1%
	Decile 6	3.2%	8.9%	6.5%	6.7%	8.6%	5.4%	9.3%
	Decile 7	6.4%	10.4%	8.6%	8.1%	10.3%	7.2%	10.8%
	Decile 8	12.5%	12.4%	11.8%	10.1%	12.7%	9.9%	12.6%
	Decile 9	21.8%	15.7%	17.6%	14.3%	16.5%	14.7%	15.3%
	Decile 10	50.0%	25.7%	41.6%	41.2%	28.1%	51.6%	22.7%
Upper boundary:	Decile 1	308	5,795	761	273	245	208	2,627
	Decile 2	444	7,503	1,251	358	337	327	3,386
	Decile 3	599	9,104	1,786	437	423	462	4,031
	Decile 4	806	10,742	2,417	522	515	635	4,687
Media	n/Decile 5	1,144	12,519	3,203	620	618	863	5,404
	Decile 6	1,954	14,584	4,236	741	742	1,163	6,230
	Decile 7	4,120	17,188	5,695	903	899	1,573	7,234
	Decile 8	7,451	20,958	8,025	1,168	1,126	2,209	8,562
	Decile 9	13,165	27,717	12,816	1,863	1,534	3,553	10,732
	Decile 10	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000
Gini-coefficient		0.6814	0.3387	0.5556	0.4885	0.3837	0.6305	0.3009
Theil index		1.0527	0.1908	0.5913	0.4038	0.2520	0.7691	0.1524
Theil index 2		0.8616	0.1942	0.5776	0.4963	0.2492	0.8982	0.1487
Variance (std.)		1.6965	0.6968	1.4753	1.4494	0.7970	2.6386	0.5815
Mean Income		4,562	15,231	5,707	1,009	791	1,877	6,215
Decile ratio	mean	103 582	0.6275	0.7007 49.089	0.7479	0.6345 12.249	0.7543 69.491	0.5983
		100.002	0.700	19.005	15.777	12.219	05.151	7.100
CONTRIBUTION	OF REGIONS	S TO GLOBAL D	eciles	0.2%	4 5%	3.5%	1.9%	0.0%
Decile 2		10%	0.0%	0.2%	5.5%	3.3%	1.5%	0.0%
Decile 3		10%	0.0%	0.2%	5.5%	3.3%	1.0%	0.0%
Decile 4		10%	0.0%	0.3%	5.3%	3.4%	1.0%	0.0%
Decile 5		10%	0.0%	0.6%	4.8%	3.4%	1.2%	0.1%
Decile 6		10%	0.0%	1.2%	3.6%	3.0%	1.8%	0.4%
Decile 7		10%	0.6%	2.2%	1.9%	1.0%	1.6%	2.8%
Decile 8		10%	3.0%	1.6%	0.8%	0.1%	0.5%	4.1%
Decile 9		10%	6.1%	1.0%	0.3%	0.0%	0.2%	2.4%
Decile 10		10%	8.5%	0.8%	0.1%	0.0%	0.2%	0.5%
Total		100%	18.2%	8.2%	32.1%	20.8%	10.4%	10.2%
Note: sums to we	orld deciles by	y row (correspon	ds to vertical slice	s of Figure B)				
REGIONAL POPU	LATION FAI	LING INTO GLO	DBAL DECILES					
Decile 1		10%	0.0%	1.9%	13.9%	16.7%	18.4%	0.0%
Decile 2		10%	0.0%	2.1%	17.1%	15.7%	10.3%	0.0%
Decile 3		10%	0.0%	2.8%	17.1%	15.8%	9.4%	0.0%
Decile 4		10%	0.0%	4.1%	16.4%	16.2%	9.6%	0.1%
Decile 5		10%	0.0%	6.9%	14.9%	16.2%	11.7%	0.5%
Decile 6		10%	0.1%	15.0%	11.3%	14.3%	17.1%	3.5%
Decile 7		10%	3.0%	26.2%	6.0%	4.8%	15.4%	27.2%
Decile 8		10%	16.5%	19.0%	2.4%	0.3%	4.5%	40.5%
Decile 9		10%	33.7%	12.4%	0.8%	0.0%	1.9%	23.4%
Decile 10		10%	46.7%	9.6%	0.2%	0.0%	1.7%	4.7%
Total		100%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Note: sums to reg	gional totals b	oy column						
COMPOSITION O	F WORLD'S	POOR BY REGI	ON					
b	elow \$700	1,543.8	0.0	31.6	798.4	516.6	196.8	0.3
be	low \$1,400	2,377.4	0.2	82.6	1,191.3	795.0	301.9	6.4
Note: 1999 Int.\$7 expenditures in 1 <u>Referenc</u> e:	(00 in Person Household Su	al Consumption . urveys	Expenditures (SN	A) terms approxin	nately correspond	ls to 1993 \$1.08-c	a-day in terms of	household

Within-country inequality	0.220	Between-country inequality	0.642
	25.5%		74.5%





CHARACTER	RISTICS	WORLD	OECD	LAC	EAP	SAS	AFR	ECE
Income share:	Decile 1	0.5%	2.8%	0.9%	1.9%	2.3%	0.6%	3.0%
	Decile 2	0.9%	4.2%	1.8%	3.0%	3.7%	1.2%	4.7%
	Decile 3	1.3%	5.3%	2.7%	3.9%	4.8%	1.8%	5.9%
	Decile 4	1.7%	6.3%	3.7%	4.8%	6.0%	2.5%	7.0%
	Decile 5	2.3%	7.4%	5.0%	5.7%	7.2%	3.5%	8.1%
	Decile 6	3.3%	8.7%	6.5%	6.9%	8.6%	4.9%	9.4%
	Decile 7	5.6%	10.3%	8.6%	8.4%	10.3%	7.0%	10.8%
	Decile 8	11.0%	12.4%	11.8%	10.6%	12.7%	10.1%	12.6%
	Decile 9	20.9%	15.7%	17.5%	14.7%	16.5%	15.5%	15.4%
	Decile 10	52.4%	26.9%	41.5%	40.0%	27.9%	52.7%	23.0%
Upper boundary:	Decile 1	382	6,969	706	398	313	179	2,878
	Decile 2	573	9,178	1,154	539	428	282	3,745
	Decile 3	778	11,146	1,638	670	538	399	4,513
	Decile 4	1,037	13,141	2,205	809	654	548	5,293
Media	n/Decile 5	1,418	15,359	2,905	971	783	759	6,123
	Decile 6	2,119	18,005	3,820	1,172	938	1,070	7,054
	Decile 7	3,935	21,427	5,107	1,442	1,138	1,536	8,171
	Decile 8	7,743	26,279	7,158	1,866	1,424	2,253	9,664
	Decile 10	14,734	34,932	11,386	2,878	1,929	3,683	12,249
	Decile 10	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000
Gini-coefficient		0.6855	0.3532	0.5521	0.4854	0.3813	0.6506	0.3065
Theil index		1.0205	0.2084	0.5805	0.4016	0.2486	0.8396	0.1595
Theil index 2		0.8911	0.2156	0.5735	0.4792	0.2459	0.9424	0.1544
Variance (std.)		1.8065	0.7542	1.4834	1.3904	0.7902	2.6991	0.5921
Income less than	mean	<b>5,100</b> 0,7383	0.6368	<b>5,127</b> 0,7013	<b>1,546</b> 0,7300	0.6342	<b>1,844</b> 0 7487	<b>7,022</b> 0,5968
Decile ratio	mean	103.773	9.623	47.196	20.559	12.029	81.935	7.563
CONTRIBUTION	OF REGIONS	S TO GLOBAL D	ECILES					
Decile 1		10%	0.0%	0.3%	2.9%	3.5%	3.4%	0.0%
Decile 2		10%	0.0%	0.3%	4.4%	3.8%	1.5%	0.0%
Decile 3		10%	0.0%	0.4%	4.9%	3.6%	1.1%	0.0%
Decile 4		10%	0.0%	0.5%	5.0%	3.5%	1.0%	0.0%
Decile 5		10%	0.0%	0.7%	5.0%	3.2%	1.0%	0.1%
Decile 6		10%	0.0%	1.1%	4.6%	2.7%	1.3%	0.3%
Decile 7		10%	0.2%	1.9%	3.2%	1.5%	1.5%	1.7%
Decile 8		10%	2.0%	1.8%	1.3%	0.2%	0.7%	4.1%
Decile 9		10%	5.6%	1.0%	0.6%	0.0%	0.2%	2.6%
Decile 10		10%	8.7%	0.5%	0.2%	0.0%	0.2%	0.5%
Total	orld deciles hi	100%	16.4%	8.6%	32.0%	22.0%	11.8%	9.3%
				,e oj 1 (gul e D)				
Decile 1	LATION FAL	10%	0.0%	3.4%	8.9%	15.8%	28.6%	0.0%
Decile 2		10%	0.0%	3.7%	13.7%	17.3%	12.8%	0.0%
Decile 3		10%	0.0%	4.5%	15.2%	16.6%	9.3%	0.1%
Decile 4		10%	0.0%	5.8%	15.7%	15.7%	8.4%	0.3%
Decile 5		10%	0.0%	8.2%	15.6%	14.5%	8.7%	0.9%
Decile 6		10%	0.1%	13.0%	14.5%	12.4%	10.6%	2.9%
Decile 7		10%	1.3%	22.5%	10.0%	6.8%	12.5%	18.2%
Decile 8		10%	11.9%	20.9%	4.0%	0.9%	5.7%	44.0%
Decile 9		10%	34.0%	11.6%	1.7%	0.0%	1.9%	28.3%
Decile 10		10%	52.7%	6.4%	0.5%	0.0%	1.4%	5.3%
Total		100%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Note: sums to re	gional totals b	y column						
COMPOSITION C	F WORLD'S	POOR BY REGI	ON					
1	elow \$700	1,355.3	0.0	43.4	529.9	494.2	287.5	0.2
be	low \$1,400	2,548.6	0.1	111.1	1,127.7	895.9	407.9	5.9
Note: 1999 Int.\$ expenditures in T Reference:	700 in Persond Household Su	al Consumption rveys	Expenditures (SN	A) terms approxin	nately correspond	ls to 1993 \$1.08-c	a-day in terms of	household

Within-country inequality	0.230	Between-country inequality	0.661
	25.8%		74.2%





CHARACTER	RISTICS	WORLD	OECD	LAC	EAP	SAS	AFR	ECE
Income share:	Decile 1	0.5%	2.7%	0.8%	1.4%	3.0%	0.7%	1.9%
	Decile 2	1.0%	4.1%	1.7%	2.4%	4.4%	1.2%	3.1%
	Decile 3	1.5%	5.2%	2.5%	3.3%	5.5%	1.8%	4.1%
	Decile 4	2.0%	6.2%	3.5%	4.3%	6.5%	2.4%	5.3%
	Decile 5	2.7%	7.2%	4.7%	5.4%	7.7%	3.1%	6.5%
	Decile 6	3.6%	8.5%	6.2%	6.7%	9.0%	4.2%	8.1%
	Decile 7	5.3%	10.1%	8.3%	8.4%	10.5%	6.1%	10.1%
	Decile 8	9.3%	12.2%	11.5%	11.0%	12.5%	9.7%	12.8%
	Decile 9	19.7%	15.6%	17.4%	15.8%	15.7%	16.5%	17.2%
	Decile 10	54.4%	28.3%	43.6%	41.3%	25.2%	54.4%	30.9%
Upper boundary:	Decile 1	450	7,566	785	471	499	170	1,307
	Decile 2	700	10,018	1,277	688	646	259	1,840
	Decile 3	963	12,160	1,817	905	781	354	2,373
	Decile 4	1,278	14,373	2,456	1,144	922	467	2,971
Media	n/Decile 5	1,700	16,860	3,259	1,423	1,078	620	3,682
	Decile 6	2,363	19,849	4,327	1,777	1,259	859	4,568
	Decile 7	3,666	23,683	5,860	2,267	1,482	1,295	5,732
	Decile 8	7,216	29,141	8,365	3,045	1,793	2,132	7,405
	Decile 9	15,569	39,502	13,701	4,814	2,352	3,843	10,354
	Decile 10	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000
Gini-coefficient		0.6842	0.3684	0.5712	0.5204	0.3338	0.6676	0.4280
Theil index		0.9777	0.2276	0.6245	0.4804	0.1856	0.8840	0.3198
Theil index 2		0.9117	0.2402	0.6223	0.5284	0.1872	0.9699	0.3153
Variance (std.)		1.9210	0.8198	1.5774	1.4276	0.6788	2.6577	0.9311
Mean Income		5,533	21,526	6,104	2,383	1,299	1,731	5,071
Decile ratio	mean	103.359	0.6473 10.680	0.7125 52.602	29.768	0.6200 8.486	0.7593 82.077	0.6470 16.356
CONTRIBUTION	OF REGIONS	5 TO GLOBAL D	DECILES	0.3%	2.9%	1 7%	5.1%	0.0%
Decile 2		10%	0.0%	0.4%	3.6%	3.9%	2.0%	0.1%
Decile 3		10%	0.0%	0.5%	3.8%	4.3%	1.2%	0.2%
Decile 4		10%	0.0%	0.6%	3.9%	4.2%	0.9%	0.4%
Decile 5		10%	0.0%	0.7%	4.1%	3.8%	0.8%	0.6%
Decile 6		10%	0.0%	0.9%	4.3%	2.9%	0.8%	1.0%
Decile 7		10%	0.1%	1.4%	4.2%	1.8%	1.0%	1.6%
Decile 8		10%	1.2%	1.9%	3.1%	0.5%	0.9%	2.4%
Decile 9		10%	5.6%	1.4%	1.3%	0.0%	0.3%	1.4%
Decile 10		10%	8.5%	0.7%	0.4%	0.0%	0.1%	0.3%
Total		100%	15.4%	8.8%	31.5%	23.1%	13.1%	8.1%
Note: sums to wo	orld deciles by	y row (correspon	ds to vertical slice	es of Figure B)				
REGIONAL POPU	LATION FAL	LING INTO GLO	OBAL DECILES					
Decile 1		10%	0.0%	3.7%	9.1%	7.3%	38.7%	0.4%
Decile 2		10%	0.0%	4.6%	11.5%	16.7%	15.3%	1.3%
Decile 3		10%	0.0%	5.3%	12.0%	18.7%	9.1%	2.8%
Decile 4		10%	0.0%	6.4%	12.5%	18.2%	6.7%	5.0%
Decile 5		10%	0.0%	7.9%	13.0%	16.5%	5.8%	7.8%
Decile 6		10%	0.1%	10.7%	13.5%	12.7%	6.4%	12.5%
Decile 7		10%	0.7%	15.5%	13.3%	7.6%	7.4%	20.0%
Decile 8		10%	7.9%	21.9%	9.9%	2.1%	6.9%	29.3%
Decile 9		10%	36.2%	15.7%	4.0%	0.1%	2.7%	17.6%
Decile 10		10%	55.0%	8.1%	1.3%	0.0%	1.1%	3.4%
Total		100%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Note: sums to reg	gional totals b	y column						
COMPOSITION O	F WORLD'S	POOR BY REGI	ON					
b	elow \$700	1,172.3	0.0	42.7	380.4	325.2	416.1	7.9
be	low \$1,400	2,537.8	0.1	115.1	911.3	903.4	552.8	55.2
Note: 1999 Int.\$7 expenditures in I Reference:	700 in Person Household Su	al Consumption rveys	Expenditures (SN	A) terms approxin	nately correspond	ls to 1993 \$1.08-0	a-day in terms of	household

Within-country inequality	0.275	Between-country inequality	0.637
	30.2%		69.8%





CHARACTER	ISTICS	WORLI	OECD	LAC	EAP	SAS	AFR	ECE
Income share:	Decile 1	0.6%	2.6%	0.8%	1.5%	3.0%	0.7%	2.3%
	Decile 2	1.2%	4.0%	1.6%	2.7%	4.5%	1.3%	3.7%
	Decile 3	1.8%	5.0%	2.5%	3.6%	5.5%	1.9%	4.8%
	Decile 4	2.4%	6.0%	3.4%	4.5%	6.6%	2.6%	5.9%
	Decile 5	3.2%	7.1%	4.6%	5.6%	7.7%	3.4%	7.2%
	Decile 6	4.3%	8.3%	6.1%	6.9%	9.0%	4.5%	8.6%
	Decile 7	6.1%	9.9%	8.2%	8.7%	10.5%	6.3%	10.3%
	Decile 8	9.8%	12.0%	11.4%	11.3%	12.5%	9.9%	12.7%
	Decile 9	18.8%	15.8%	17.4%	16.1%	15.7%	17.2%	16.4%
	Decile 10	51.7%	29.3%	44.0%	39.0%	25.0%	52.2%	28.1%
Upper boundary:	Decile 1	700	9,846	949	982	855	221	2,924
	Decile 2	1,146	12,850	1,551	1,405	1,108	339	3,984
	Decile 3	1,593	15,682	2,211	1,825	1,337	462	5,004
	Decile 4	2,115	18,648	2,997	2,279	1,576	608	6,081
Mediar	n/Decile 5	2,791	21,930	3,986	2,814	1,841	798	7,282
	Decile 6	3,804	25,820	5,304	3,492	2,149	1,083	8,713
	Decile 7	5,637	30,864	7,210	4,427	2,530	1,603	10,559
	Decile 8	9,881	38,556	10,325	5,908	3,054	2,638	13,221
	Decile 9	19,919	53,657	17,010	9,209	3,988	4,798	17,913
	Decile 10	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000
Gini-coefficient		0.6542	0.3792	0.5760	0.4981	0.3312	0.6499	0.3823
Theil index		0.8633	0.2401	0.6371	0.4347	0.1826	0.8257	0.2493
Theil index 2		0.8296	0.2555	0.6350	0.4713	0.1838	0.8981	0.2489
Variance (std.)		1.8340	0.8522	1.6041	1.2934	0.6699	2.4754	0.8032
Mean Income	magn	<b>7,557</b>	<b>28,553</b>	<b>7,567</b>	<b>4,514</b>	<b>2,211</b>	<b>2,079</b>	<b>9,318</b>
Decile ratio	mean	92.034	11.135	54.692	25.254	8.354	72.674	11.991
CONTRIBUTION	OF REGIONS	TO GLOBAL D	ECILES					
Decile 1		10%	0.0%	0.5%	1.4%	1.3%	6.7%	0.0%
Decile 2		10%	0.0%	0.7%	2.8%	4.1%	2.4%	0.0%
Decile 3		10%	0.0%	0.7%	3.3%	4.7%	1.2%	0.1%
Decile 4		10%	0.0%	0.7%	3.7%	4.6%	0.9%	0.2%
Decile 5		10%	0.0%	0.8%	4.0%	4.1%	0.8%	0.3%
Decile 6		10%	0.0%	1.0%	4.3%	3.3%	0.8%	0.6%
Decile 7		10%	0.2%	1.2%	4.5%	2.0%	0.9%	1.2%
Decile 8		10%	1.2%	1.5%	3.8%	0.7%	0.7%	2.0%
Decile 9		10%	4.8%	1.2%	1.9%	0.1%	0.3%	1.7%
Total		100%	1/ 1%	0.776	30.5%	24.0%	14 0%	6.6%
Note: sums to wo	orld deciles by	y row (correspon	ds to vertical slice	s of Figure B)	30.370	21.970	14.970	0.070
REGIONAL POPU	LATION FAL	LING INTO GL	OBAL DECILES					
Decile 1		10%	0.0%	6.0%	4.5%	5.3%	45.2%	0.0%
Decile 2		10%	0.0%	7.3%	9.2%	16.3%	16.4%	0.3%
Decile 3		10%	0.0%	7.4%	10.8%	19.0%	8.2%	1.1%
Decile 4		10%	0.0%	8.0%	12.0%	18.3%	5.9%	2.5%
Decile 5		10%	0.1%	8.9%	13.1%	16.5%	5.2%	4.8%
Decile 6		10%	0.3%	10.8%	14.1%	13.1%	5.4%	9.3%
Decile 7		10%	1.3%	13.7%	14.8%	8.2%	5.8%	17.7%
Decile 8		10%	8.5%	16.8%	12.5%	2.9%	4.9%	30.7%
Decile 9		10%	33.9%	13.3%	6.3%	0.3%	2.1%	25.8%
Decile 10		10%	56.0%	7.8%	2.6%	0.0%	0.9%	7.0%
Total	tional totals h	100%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	<b>_ </b>							
COMPOSITION O	elow \$700	689.2	(0.0)	37.4	95.4	91.7	464.5	0.2
be	low \$1,400	1,778.1	0.0	109.2	417.4	560.9	686.5	4.0
Note: 1999 Int.\$7 expenditures in F	700 in Person Household Su	al Consumption rveys	Expenditures (SN.	A) terms approxin	nately correspond	's to 1993 \$1.08-a	a-day in terms of	household

Within-country inequality	0.276	Between-country inequality 0.55	4
	33.3%	66.7	%
			_

			- /			
	WORLD	LAC	EAP	SAS	AFR	ECE
pop. share, %						
700	38.56%	15.31%	67.89%	59.21%	44.51%	0.19%
1,400	55.41%	33.10%	90.19%	88.94%	69.08%	2.14%
inc. share, %						
700	3.97%	1.53%	35.88%	31.94%	8.53%	0.02%
1,400	8.17%	5.92%	63.40%	70.23%	21.74%	0.45%
inc. of group						
700	401	421	404	406	360	600
1,400	574	751	538	595	591	1,090
poverty gap						
700	42.8%	39.9%	42.3%	41.9%	48.6%	14.3%
1,400	59.0%	46.4%	61.6%	57.5%	57.8%	22.2%

### AVERAGE INCOMES OF THE POOR AND POVERTY GAP, 1970

### AVERAGE INCOMES OF THE POOR AND POVERTY GAP, 1980

	WORLD	LAC	EAP	SAS	AFR	ECE
pop. share, %						
700	35.33%	8.78%	56.89%	56.85%	43.16%	0.08%
1,400	54.40%	22.93%	84.87%	87.48%	66.22%	1.43%
inc. share, %						
700	3.20%	0.69%	24.05%	29.54%	8.28%	0.01%
1,400	7.29%	3.29%	50.66%	67.27%	20.71%	0.26%
inc. of group						
700	413	450	426	411	360	593
1,400	611	819	602	608	587	1,109
poverty gap						
700	41.0%	35.7%	39.1%	41.3%	48.6%	15.3%
1,400	56.4%	41.5%	57.0%	56.6%	58.1%	20.8%

### AVERAGE INCOMES OF THE POOR AND POVERTY GAP, 1990

	WORLD	LAC	EAP	SAS	AFR	ECE
pop. share, %						
700	26.39%	9.86%	32.26%	43.75%	47.56%	0.04%
1,400	49.62%	25.24%	68.66%	79.31%	67.48%	1.24%
inc. share, %						
700	2.22%	0.86%	9.90%	19.37%	8.82%	0.00%
1,400	6.72%	4.00%	33.48%	54.65%	19.67%	0.20%
inc. of group						
700	434	448	474	442	342	574
1,400	700	812	754	689	537	1,120
poverty gap						
700	38.0%	36.0%	32.3%	36.8%	51.2%	18.0%
1,400	50.0%	42.0%	46.1%	50.8%	61.6%	20.0%

### AVERAGE INCOMES OF THE POOR AND POVERTY GAP, 2000

	WORLD	LAC	EAP	SAS	AFR	ECE
pop. share, %						
700	19.98%	8.31%	20.55%	23.98%	53.93%	1.68%
1,400	43.25%	22.39%	49.24%	66.61%	71.65%	11.67%
inc. share, %						
700	1.57%	0.62%	3.99%	9.46%	10.63%	0.18%
1,400	5.87%	3.04%	16.37%	42.86%	20.72%	2.33%
inc. of group						
700	434	457	462	512	341	538
1,400	751	829	792	836	501	1,014
poverty gap						
700	38.0%	34.7%	33.9%	26.8%	51.2%	23.1%
1,400	46.4%	40.8%	43.4%	40.3%	64.2%	27.6%

In Eleion Hier	owned of the rook.		015			
	WORLD	LAC	EAP	SAS	AFR	ECE
pop. share, %						
700	10.01%	6.01%	4.54%	5.34%	45.23%	0.04%
1,400	25.82%	17.54%	19.88%	32.70%	66.85%	0.89%
inc. share, %						
700	0.56%	0.37%	0.52%	1.35%	8.19%	0.00%
1,400	2.77%	1.98%	4.16%	14.65%	18.42%	0.11%
inc. of group						
700	425	465	517	558	376	589
1,400	810	852	944	991	573	1,137
poverty gap						
700	39.3%	33.5%	26.2%	20.3%	46.2%	15.8%
1,400	42.1%	39.1%	32.6%	29.2%	59.1%	18.8%

AVERAGE INCOMES OF THE POOR AND POVERTY GAP, 2015

Poverty Gap (below \$700 income group)

	1970	1980	1990	2000	2015
Latin America	39.9%	35.7%	36.0%	34.7%	33.5%
East Asia	42.3%	39.1%	32.3%	33.9%	26.2%
South Asia	41.9%	41.3%	36.8%	26.8%	20.3%
Africa	48.6%	48.6%	51.2%	51.2%	46.2%
East and Central Europe	14.3%	15.3%	18.0%	23.1%	15.8%

### ANNEX IV. HOW TO READ THE DISTRIBUTION DIAGRAM

The diagram on the right describes the global income distribution [for actual charts see ANNEX III]. The diagram consists of three parts:

(1) The top part shows the population distribution density function (in logarithmic form)  $F'_{\ln x} = \frac{\partial F}{\partial \ln x}$ , decomposed by

region. The overall area under the curve is equal to one:  $\int F'_{\ln x} d \ln x = 1$ , and the regional areas are equal to the respective regional population shares:  $\int F'_{\ln x}(\text{region}) d\ln x = s_{\text{region}}$ ,  $\sum_{\forall \text{in regions}} s_{\text{regioni}} = 1$ . The density functions are plotted on the logarithmic scale to reflect the fact that actual income distributions are approximately log-normal. This presentation allows comparing shapes of distributions from different years or countries and provides a common WORLD INCOME DISTRIBUTION, 2000 basis for aggregation. It shows clearly the twin-peak nature of the global distribution, as well as relative 0.400 importance of various regions in population terms. A region whose income changes would shift in the 0.300 horizontal direction, if that region becomes more POI 0 250 populous then its density function would scale proportionately in the vertical direction. 0.200 0.150

(2) The middle part of the diagram shows the income distribution density function (in logarithmic form):

$$F'_{\ln x} \frac{x}{x_{mean}} = \frac{\partial \left( \int_0^x F'_x \tilde{x} \, d\tilde{x} \right)}{\partial x} \frac{x}{x_{mean}}, \text{ again}$$

decomposed by region. The overall area under the

curve is also equal to one:  $\int F'_{\ln x} \frac{x}{x_{mean}} d\ln x = 1$ ,

and the regional areas are equal to the respective regional income shares. In contrasting the middle part to the top part, one can see immediately how unequal the income distribution by region is. The share of OECD in income, for example, is much larger to that in population.

(3) The bottom part of the diagram shows the regional composition of global income percentiles. One can immediately see who is populating one or another global decile: for example, the lowest global decile is dominated by Africa, whereas the richest one is

0.10 0.05 OFCD Den eitz (0.050 OECD (0.100 (0.150) (0.200) (0.250 (0.300) (0.350) (0.400) Correspondence between entiles and ir (0.450) 'ECE 111111 omposition of globa OFICE

originating mostly from OECD. All three parts of the diagram are interconnected by the global decile boundary lines (dotted vertical lines). These lines originate from the bottom part of the diagram and cut through the corresponding income levels of the upper parts. For example, the 50% line [global median] happens to cross the income line of the upper parts of the diagram at \$1,418.

Next let us look at the set of tables that accompany the diagram. The set consists of four parts:

- The top part describes the following characteristics: income shares by decile, both global and regional, interval boundaries for these deciles, as well as main characteristics – the Gini, two Theils indices, variance, mean income, the share of population with incomes less than the mean, decile ratio [defined as the ratio of incomes of the top and the bottom deciles].
- (2) The second part shows the contribution of regions to global deciles, or where the populations of global deciles come from. The regional sums equal to the regional population shares. For example, one can see that 79% of the top global decile comes from the OECD. At the same time the OECD's share in global population is 15.4%. This table corresponds to the Part 3 [bottom] of the diagram.
- (3) The third part of the set shows shares of regional populations falling into global deciles. For example, 55% of the OECD population falls into the top global decile.
- (4) The bottom table displays composition of the world's poor by region. As a reference, below is shown a decomposition of the global inequality into the within-country and between country parts using the Theil 2 (income weighted) index.

March 4, 2005

	WORLI	INCOM	E DIST	RIBUTIC	<b>DN, 2000</b>	)	
	RLD	e.				~	
CHARACTERISTICS	ow	OBO	IVC	EME	SVS	AFF	BCI
Income share: Decile 1	0.5%	2.7%	0.8%	1.4%	3.0%	0.7%	1.9%
Decile 2	1.0%	4.1%	1.7%	2.4%	4.4%	1.2%	3.19
Decile 3	1.5%	5.2%	2.5%	3.3%	5.5%	1.8%	4.19
Decile 4	2.0%	6.2%	3.5%	4.3%	6.5%	2.4%	5.3%
Decile 5	2.7%	7.2%	4.7%	5.4%	7.7%	3.1%	6.5%
Decile 6	3.6%	8.5%	6.2%	6.7%	9.0%	4.2%	8.19
Decile 7	0.3%	10.1%	8.3%	8.4%	10.5%	0.1%	10.17
Decile 9	19.7%	15.6%	17.4%	15.8%	15.7%	16.5%	12.87
Decile 10	54.4%	28.3%	43.6%	41.3%	25.2%	54.4%	30.9%
	450	7.566	705	171	100	170	1.005
Upper boundary: Decile 1	450	7,566	785	471	499	170	1,307
Decile 2	963	12,160	1,277	905	781	354	2 373
Decile 4	1 278	14 373	2 456	1 144	922	467	2,070
Median/Decile 5	1.700	16.860	3.259	1.423	1.078	620	3.682
Decile 6	2,363	19,849	4,327	1,777	1,259	859	4,568
Decile 7	3,666	23,683	5,860	2,267	1,482	1,295	5,732
Decile 8	7,216	29,141	8,365	3,045	1,793	2,132	7,405
Decile 9	15,569	39,502	13,701	4,814	2,352	3,843	10,354
Decile 10	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000
Gini-confficient	0.6842	0 3684	0 5712	0 5204	0 3338	0.6676	0 4280
Theil index	0.9777	0.2276	0.6245	0.4804	0.1856	0.8840	0.3198
Theil index 2	0.9117	0.2402	0.6223	0.5284	0.1872	0.9699	0.3153
Variance (std.)	1.9210	0.8198	1.5774	1.4276	0.6788	2.6577	0.9311
Mean Income	5,533 0.7651	21,526	6,104 0.7125	2,383	1,299	1,731	5,071 0.647/
Decile ratio	103.359	10.680	52.602	29.768	8.486	82.077	16.350
Decile 1	10%	L DECILES	0.3%	2.9%	1.7%	5.1%	0.03
Decile 2	10%	0.0%	0.4%	3.6%	3.9%	2.0%	0.19
Decile 3	10%	0.0%	0.5%	3.8%	4.3%	1.2%	0.29
Decile 4	10%	0.0%	0.6%	3.9%	4.2%	0.9%	0.4%
Decile 5	10%	0.0%	0.7%	4.1%	3.8%	0.8%	0.6%
Decile 6	10%	0.0%	0.9%	4.3%	2.9%	0.8%	1.0%
Decile 7	10%	0.1%	1.4%	4.2%	1.8%	1.0%	1.6%
Decile 9	10%	5.6%	1.9%	1 3%	0.0%	0.3%	1.49
Decile 10	10%	8.5%	0.7%	0.4%	0.0%	0.1%	0.3%
Total	100%	15.4%	8 8%	31 5%	23.1%	13.1%	8 19
Note: sums to world deciles	by row (corres)	oonds to vertical	slices of Figure .	B)			
Decile 1	10%	0.0%	3.7%	9.1%	7.3%	38.7%	0.49
Decile 2	10%	0.0%	4.6%	11.5%	16.7%	15.3%	1.39
Decile 3	10%	0.0%	5.3%	12.0%	18.7%	9.1%	2.8%
Decile 4	10%	0.0%	6.4%	12.5%	18.2%	6.7%	5.0%
Decile 5	10%	0.0%	7.9%	13.0%	16.5%	5.8%	7.8%
Decile 6	10%	0.1%	10.7%	13.5%	12.7%	6.4%	12.5%
Decile 7	10%	0.7%	15.5%	13.3%	7.6%	7.4%	20.0%
Decile 8	10%	7.9%	21.9%	9.9%	2.1%	0.9%	29.39
Decile 10	10%	55.0%	8.1%	1.3%	0.1%	2.7%	3.49
T-tal	100%	100.0%	100.0%	100.0%	100.0%	100.0%	100.00
Note: sums to regional total	is bu column	100.0%	100.0%	100.0%	100.0%	100.0%	100.09
note: sums to regional total	a by countr						
COMPOSITION OF WORLI	1 172.3	EGION	42.7	380.4	325.2	416.1	7 0
below \$1,400	2,537.8	0.1	115.1	911.3	903.4	552.8	55.2
Note: 1999 Int.\$700 in Per-	sonal Consumpt	ion Expenditure:	s (SNA) terms ap	proximately corre	esponds to 1993	\$1.08-a-day in	terms of
household expenditures in	Household Suri	æys					
Reference:							
Within-country	y inequality	0.275		1	Between-countr	ry inequality	0.637
1		30.2%					69.8%