JEL Classification: C01; J10 https://doi.org/10.35945/gb.2019.07.016

STUDY OF THE DISTRIBUTION OF WEALTH IN THE MIDDLE AND TOP SEGMENTS OF THE POPULATION OF GEORGIA

QETEVAN PIPIA

Phd student
Samtskhe-Javakheti State University, Georgia
qetino.pipia@gmail.com

KEYWORDS: PARETO DISTRIBUTION, LOGNORMAL DISTRIBUTION, MODELS OF ECONOMICAL LAYERS OF SOCIETY.

For citation: Pipia Q. (2019), Study Of The Distribution Of Wealth In The Middle And Top Segments Of The Population Of The Population Of Georgia, *Globalization And Business, №7*, pp. 130-142. https://doi.org/10.35945/gb.2019.07.016

Due to social, economic, political and other differences, the distribution of the population by income is a subject of different laws of distributions, for example, of a normal distribution, a log-normal distribution, Pareto distribution, and so on. In some countries the distribution of elements of society by income per unit of time in the lower and middle groups will be close to normal [10] (later it turned out that the distribution of income in the lower and middle groups better described by a log-normal distribution). The Italian economist Wilfredo Pareto remarked that in the richer segments of society, the income corresponding to the elements of society would be much higher than in the case of a normal distribution, because the large income, in contrast to the small and medium, makes it possible to invest in growth. The study of this phenomenon led to a distribution with two parameters, which is also found in statistics, sociology, etc., and is called the Pareto distribution [1].

The following distribution with two parameters a and ν is called **Pareto distribution:**

$$f = \frac{a}{x^{\nu}}.$$

In [10] the structure of society consisting of three groups is considered: the bottom 10%, with the average distribution of GDP l (low), the average 80% with the average distribution of GDP m (middle), the top 10%, - with the average distribution of GDP t (top). If we denote the share of GDP per capita by d, then we get the balance equation

$$0{,}1l + 0{,}8m + 0{,}1t = d. \qquad (1)$$

Indeed**,**

$$l = \frac{GDP(low)}{N_{low}} = \frac{GDP(low)}{0.1N}, \qquad m = \frac{GDP(middle)}{N_{middle}} = \frac{GDP(middle)}{0.8N}, \qquad t = \frac{GDP(top)}{N_{top}} = \frac{GDP(top)}{0.1N},$$

and it follows the equality (1):

$$0{,}1l + 0{,}8m + 0{,}1t = \frac{GDP(low)}{N} + \frac{GDP(middle)}{N} + \frac{GDP(top)}{N} = \frac{GDP}{N} = d.$$

The ratio of funds is defined as the ratio of incomes of the same number population, but belonging to different income groups. When 10% is selected from the upper and lower levels, this ratio is called **decile ratio** [8].

In the United States, the ratio of the upper 90% -95% layer to the lower 5% layer is often considered [4] since the incomes of the upper 5% layer are subject of other laws of distribution. This ratio is called the **rank coefficient**. We will consider the ratio of the incomes of the upper 10% stratum to the incomes to the lower 10% stratum, i.e. **decimal coefficient** D. Since the number of people in these groups is the same, this ratio is equal to the ratio of the average incomes of these groups:

$$D = \frac{t}{l};$$

Then from (**Error! Reference source not found.**) it follows

$$0,1(D+1)l + 0.8m = d. \quad (1)$$

If N is a size of a population, then the GDP consumed by the lower, middle and upper groups (respectively Q_l, Q_m, Q_t) will be

$$Q_l = 0,1Nl, \qquad Q_m = 0,8Nm, \qquad Q_t = 0,1Nt = 0,1NDl$$

and therefore

$$Q_l + Q_m + Q_t = Q = Nd,$$

there Q is the GDP.

Let us suppose that the number f of members of the top layer who has an income q is expressed by the Pareto distribution

$$f = \frac{a}{q^{\nu}}\ , \quad (2)$$

there $a > 0, \nu > 2$.

Proposition 1. Let q_t be the minimal income of members of the top group, N_t be the number of individuals who has the income more or equal to q_t, $q_n \geq q_t$, N_n is the number of individuals who has the income more or equal to q_n. Then

$$\boldsymbol{N_n = N_t \left(\frac{q_t}{q_n}\right)^{\nu-1}}.$$

Proof. It is clear, that

$$qf = \frac{a}{q^{\nu-1}} \quad (3)$$

is the income of f members of the top group. Then from (2) and (3) we get difeerential equations

$$\frac{dN}{dq} = \frac{a}{q^{\nu}}, \qquad q\frac{dN}{dq} = \frac{dQ}{dq} = \frac{a}{q^{\nu-1}}. \quad (4)$$

Let q_{max} be a maximal income in the top gorup. Then

$$N_n = \int_{q_n}^{q_{max}} \frac{a}{q^{\nu}} dq, \qquad Q_n = \int_{q_n}^{q_{max}} \frac{a}{q^{\nu-1}} dq\,.$$

Since no one has an income more than q_{max}, we can assume that the upper bound in the integral is infinite:

$$N_n = \int_{q_n}^{\infty} \frac{a}{q^{\nu}} dq = \frac{a}{(\nu-1)q_n^{\nu-1}},$$

$$Q_n = \int_{q_n}^{\infty} \frac{a}{q^{\nu-1}} dq = \frac{a}{(\nu-2)q_n^{\nu-2}}.$$

After calculations we get

$$N_n = \int_{q_n}^{\infty} \frac{a}{q^{\nu}} dq = \frac{aq^{1-\nu}}{1-\nu}\Big|_{q_n}^{\infty} = \frac{a}{(\nu-1)q_n^{\nu-1}}, \quad (5)$$

$$Q_n = \int_{q_n}^{\infty} \frac{a}{q^{\nu-1}} dq = \frac{aq^{2-\nu}}{2-\nu}\Big|_{q_n}^{\infty} = \frac{a}{(\nu-2)q_n^{\nu-2}}. \quad (6)$$

From these formulas it follows

$$Q_n = \frac{a}{(\nu-2)q_n^{\nu-2}} = \frac{a}{(\nu-1)q_n^{\nu-1}} \frac{(\nu-1)q_n^{\nu-1}}{(\nu-2)q_n^{\nu-2}} = \frac{\nu-1}{\nu-2} N_n q_n. \quad (7)$$

Now let $q_n = q_t$, where q_t is the minimal income of the top layer (equivalently the maximal income of the middle layer). Then

$$N_t = \frac{a}{(\nu-1)q_t^{\nu-1}}, \quad (8)$$

$$Q_t = \frac{a}{(\nu-2)q_t^{\nu-2}}\,. \quad (9)$$

Let us express a by N_t from (9) and substitute in (6). Then we get

$$N_n = \frac{a}{(\nu-1)q_n^{\nu-1}} = \frac{N_t(\nu-1)q_t^{\nu-1}}{(\nu-1)q_n^{\nu-1}} = N_t \left(\frac{q_t}{q_n}\right)^{\nu-1}. \quad (10)$$

We can write (6) and (8) shortly as

$$N_n = N_t \left(\frac{q_t}{q_n}\right)^{\nu-1}, \quad (11)$$

$$Q_n = \frac{\nu-1}{\nu-2} N_n q_n. \quad (12)$$

The proposition is proved.

The formula of the sconomic structure of the upper layer of the society:

Let q_t be a minimal income of the top layer, N_t be the number of the individuals which has the income more or equal to q_t. Then

a) The number of individuals who have the income more or equal to q_n and its whole income are respectively

$$N_n = N_t\left(\frac{q_t}{q_n}\right)^{\nu-1}, \quad Q_n = \frac{\nu-1}{\nu-2}N_n q_n.$$

b) The number of individuals whose income are between q_{n_1} and q_{n_2}, is

$$N_{n_2} - N_{n_1} = N_t\left(\left(\frac{q_t}{q_{n_2}}\right)^{\nu-1} - \left(\frac{q_t}{q_{n_1}}\right)^{\nu-1}\right) = N_t q_t^{\nu-1}\left(\frac{1}{q_{n_2}^{\nu-1}} - \frac{1}{q_{n_1}^{\nu-1}}\right),$$

and its whole income is

$$Q_{n_2} - Q_{n_1} = \frac{\nu-1}{\nu-2}\left(N_{n_2}q_{n_2} - N_{n_1}q_{n_1}\right).$$

The average income of top layer is

$$t = \frac{Q_t}{N_t}.$$

Corollary. Two parameters of Pareto distribution and other parameters are interlinked by formulas:

$$\nu = \frac{2Q_t - N_t q_t}{Q_t - N_t q_t} = \frac{2t - q_t}{t - q_t} = 2 + \frac{q_t}{t - q_t} > 2, \quad \alpha = (\nu - 1)q_t^{\nu-1}N_t. \tag{13}$$

Proof. Let us write (12) for Q_t:

$$Q_t = \frac{\nu-1}{\nu-2}N_t q_t.$$

Now we must take into account that $Q_t = tN_t$. The second formula follows from (8).

Since usually N_t and q_t are known from practice, by these formulas we may calculate the parameters of Pareto distribution ν and α. If we substitute these expressions in (4), we get the following differential equations

$$\frac{dN}{dq} = \frac{(\nu-1)N_t}{q_t}\left(\frac{q_t}{q}\right)^{\nu}, \quad \frac{dQ}{dq} = (\nu-1)N_t\left(\frac{q_t}{q}\right)^{\nu-1}.$$

Let us write (13) and (12) for Q_t and N_t. Then the division of (12) on (11) give us the average income of the top layer t:

$$\frac{Q_t}{N_t} = \frac{\nu-1}{\nu-2}q_t = \bar{q} = t. \tag{14}$$

Let us suppose that the boundary values of low, middle and top layers are interlinked by each other by parameters α, β, γ:

$$\text{Lower boundary of the low group } q_l = \frac{l}{\alpha};$$

$$\text{Top boundary of the low group (since } l \text{ is small enough than } m)\ q_m = \frac{l+m}{\beta} \approx \frac{m}{\beta} = \tag{15}$$

$$= \text{Lower boundary of the middle group};$$

$$\text{Top boundary of the middle group} = \text{Lower boundary of the top group } q_t = \frac{\gamma m}{\beta}.$$

The ratio of the top and lower boundaries of the middle group is

$$\frac{\gamma m}{\beta} : \frac{l+m}{\alpha} = \frac{\alpha\gamma m}{\beta(l+m)} \approx \frac{\alpha\gamma}{\beta}, \tag{16}$$

since l is small enough than m.

Let us substitute $q_t = \frac{\gamma m}{\beta}$ in (14):

$$\frac{\nu-1}{\nu-2}q_t = \frac{\nu-1}{\nu-2}\frac{\gamma m}{\beta} = t, \qquad \frac{\nu-1}{\nu-2} = \frac{\beta t}{\gamma m};$$

If we suppose that $m = kl$ for some k, when we may estimate D:

$$\frac{\nu-1}{\nu-2} = \frac{\beta t}{\gamma m} = \frac{\beta t}{\gamma k l}, \qquad D = \frac{t}{l} = \frac{\gamma k}{\beta}\frac{\nu-1}{\nu-2}.$$

From this expression it is clear that the distribution in the top layer depends on the distribution in the low layer.

For example, let be

$$\alpha = \beta = 2, \gamma = 3\,.$$

Then from (16) we get that the ratio of the top and low boundaries of the middle layer is γ. Then

$$\begin{gathered}\text{Lower boundary of the low group } q_l = \frac{l}{2};\\ \text{Top boundary of the low group (since } l \text{ is small enough than } m) \; q_m = \frac{l+m}{2} \approx \frac{m}{2} = \\ = \text{Lower boundary of the middle group;}\\ \text{Top boundary of the middle group} = \text{Lower boundary of the top group } q_t = \frac{3m}{2}.\end{gathered} \quad (17)$$

Let us remark that in some countries, for example for the USA [10] such a choise of γ is justified because a ratio of the top and low boundaries of the middle group is equal to γ=3.

The analysis of the above theoretical conclusions shows that for their application it is necessary to know the boundary between the middle and upper levels q_t and the number of individuals in the upper layer N_t. It is not necessary that the top layer consist of 10 percent of the population.

Let us apply this results to Georgia. According to the World Bank [5], the population of Georgia in 2017 was distributed by Gross Domestic Product (GDP) as follows:

Distribution of the population (%) according to GDP consumption in Georgia

2017 წ.	10%	20%	20%	20%	20%	20%	10%
	2.6	6.7	11.5	16	22.2	43.6	28.1

From this list it follows that decile ratio is

$$D = \frac{28.1}{2.6} = 10.8 \approx 11.$$

The middle group consists with 80% of population and GDP consumed by him is

$$(6.7 - 2.6) + 11.5 + 16 + 22.2 + (43.6 - 28.1) = 69.3\%.$$

Therefore the share of low, middle and top layers in GDP is equal respectively

$$\frac{2.6}{10} = 0.26, \qquad \frac{69.3}{80} = 0.86625, \qquad \frac{28.1}{10} = 2.81.$$

According to Geostat, the population of Georgia, GDP and GDP by purchasing power in 2017 were respectively

$$N = 3700000, \quad GDP = 14,100,000,000\$, \quad GDP_{purch} = 37,600,000,000\$.$$

As a result, purchasing power coefficient in 2017 was

$$\frac{37,600,000,000}{14,100,000,000} = 2.67.$$

GDP per capita (by nominal and purchasing power) will be in Georgia

$$d = \frac{14,100,000,000}{3,700,000} = 3810.811\$, \; d_{purch} = \frac{37,600,000,000}{3,700,000} = 10162.1622\$.$$

Since 10% and 80% of population are

0.1N	0.8N
370000	2960000

,

then for one representative of the population from the lower, middle and upper layers GDP is distributed approximately as follows (in dollars):

$$l = \frac{14100000000 \cdot 0.26}{370000} = 990.81, \qquad m = \frac{14100000000 \cdot 0.86625}{2960000} = 3301.115,$$

$$t = \frac{14100000000 \cdot 2.81}{370000} = 10708.38.$$

The same dates by purchasing power are equal to

$$l_{purch} = 2642, \qquad m_{purch} = 8803, \qquad t_{purch} = 28556.$$

In this example, we limit ourselves to nominal GDP.

By (17) the top bound of the consumption of GDP per capita in middle group, that is equal to the lower bound of the consumption of GDP per capita in top group is

$$q_{t_1} = \frac{3m}{2} = \frac{3 \cdot 3301}{2} = 4952.$$

Similarly, the top bound of the consumption per capita in the middle group i.e. the lower bound of the consumption per capita in the top group (by purchasing power) is

$$q_{t,purch_1} = \frac{3m_{purch}}{2} = \frac{3 \cdot 8803}{2} = 13204.46.$$

Now let's try to evaluate q_t using some well-known economic indicator. According to [12] the subjective margin of well-being of one household in the upper middle class in 2017 was 1978 GEL per month. Since one household in Georgia consists approximately of 3.3 persons [13], then the subjective margin of well-being per capita in the upper middle class is 1978:3.3=599.394 GEL per month, and 599.394x12=7192.727 GEL per year. The dollar rate in 2017 was an average of 2.5 [11]. Therefore the subjective margin of well-being per capita in the upper middle class per year is 7192.727:2.5=2877.091$. If we hypothesize that GDP per person is 1.5 times higher than its direct income, then we will get the maximum income of the upper middle class, that is, the lower boundary of the upper stratum

$$q_{t_2} = 2877.091\text{x}1.5 = 4315.636\$.$$

and because

$$q_{t,purch_2} = 11508.364.$$

If we substitues values of q_{t_1} and q_{t_2} in first formula of (14), we get two values of ν:

$$\nu_1 = 2.860, \qquad \nu_2 = 2.675.$$

The income of top groups will be

$$Q_t = 0.281 \cdot 14{,}100{,}000{,}000 = 3{,}962{,}100{,}000,$$

$$Q_{t,purch} = 0.281 \cdot 37{,}600{,}000{,}000 = 10{,}565{,}600{,}000,$$

and number of members of top group will be

$$N_t = 3{,}700{,}000 \cdot 0.1 = 370{,}000.$$

Now we hypothesize that the rich stratum of the society in Georgia is described by the Pareto distribution - this hypothesis is widespread and is confirmed in the works of many authors, for example in [1], [9], [10]. The fact is that the exponent $e^{-\frac{x^2}{2}}$ decreases faster than the power function $\frac{1}{q^{\nu-1}}$, $\nu > 2$. Therefore, Pareto distribution decreases more slowly than the normal distribution. This adequately reflects the fact that high returns allow you to invest.

As we know q_{t_1} =4951.672$. Let us compute the number of persons which share in GDP exceeds 10,000$. This means that $q_1 = 10{,}000$. Formula (11) get

$$N_1 = N_t \left(\frac{q_{t_1}}{q_1}\right)^{\nu_1 - 1} = 370000 * \left(\frac{4{,}951.672297}{10{,}000}\right)^{2.86-1} \approx 100090.$$

Also by formula (12) we have

$$Q_1 = \frac{\nu-1}{\nu-2} N_1 q_1 = \frac{2.86-1}{2.86-2} \cdot 100090 \cdot 10{,}000 = 2{,}164{,}532{,}303,$$

which means that the share in GDP of 100090 persons exceeds 10,000$, and they consume from GDP 2,164,532,303$. In the lists below, the distribution of GDP (according to the nominal and the purchasing power) is calculated when the share of one person in GDP is more than 10,000$, more than 100,000$, more than 1,000,000$, more than 10,000,000$ and more than 100,000,000$. From these tables, in particular, we can conclude that in Georgia's economy it is impossible to generate 10,000,000 dollars during a year!

Now let us do the same calculations for $q_{t_2} = 4315.636\$$. As we know in this case $\nu_2 = 2.675$. We have:

$$N_1 = N_t \left(\frac{q_{t_2}}{q_1}\right)^{\nu_2 - 1} = 370000 * \left(\frac{4315.636}{10{,}000}\right)^{2.675-1} \approx 90546;$$

$$Q_1 = \frac{\nu-1}{\nu-2} N_1 q_1 = \frac{2.675-1}{2.675-2} \cdot 90546 \cdot 10{,}000 = 2{,}246{,}726{,}101.$$

t	qt	Nt	Qt	niu	q1	N1	Q1	q2	N2	Q2
10708.38	4951.672	370000	3962100000	2.86015722	10000	100090.4	2164532303	100000	1381.131526	298680487.7
28555.68	13204.46	370000	10565600000	2.86015722	20000	170922	7392643442	100000	8562.564407	1851721480
10708.38	4315.636	370000	3962100000	2.675083767	10000	90546.42	2246726101	100000	1913.319504	474751519.4
28555.68	11508.36	370000	10565600000	2.675083767	20000	146606.8	7275502195	100000	9892.835954	2454707063

List 1.

მილიონი			10 მილ			100 მილ		
q3	N3	Q3	q4	N4	Q4	q5	N5	Q5
1000000	19.05802178	41214461.72	10000000	0.262979	5687120	100000000	0.003629	784757
1000000	118.153511	255516202.8	10000000	1.630382	35258288	100000000	0.022497	4865237
1000000	40.42999906	100318861.8	10000000	0.854319	21198192	100000000	0.018052	4479351
1000000	209.043679	518699590	10000000	4.417263	1.1E+08	100000000	0.09334	23160523

List 2.

If we compute a for example, in the case q_{t_1}, we will get a huge number:

$$a = (\nu - 1)q_{t_1}^{\nu-1}N_t = 5{,}128{,}692{,}385{,}795.29868.$$

A density of a **lognormal distribution** is

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}}e^{-\frac{(lnx-\mu)^2}{2\sigma^2}},$$

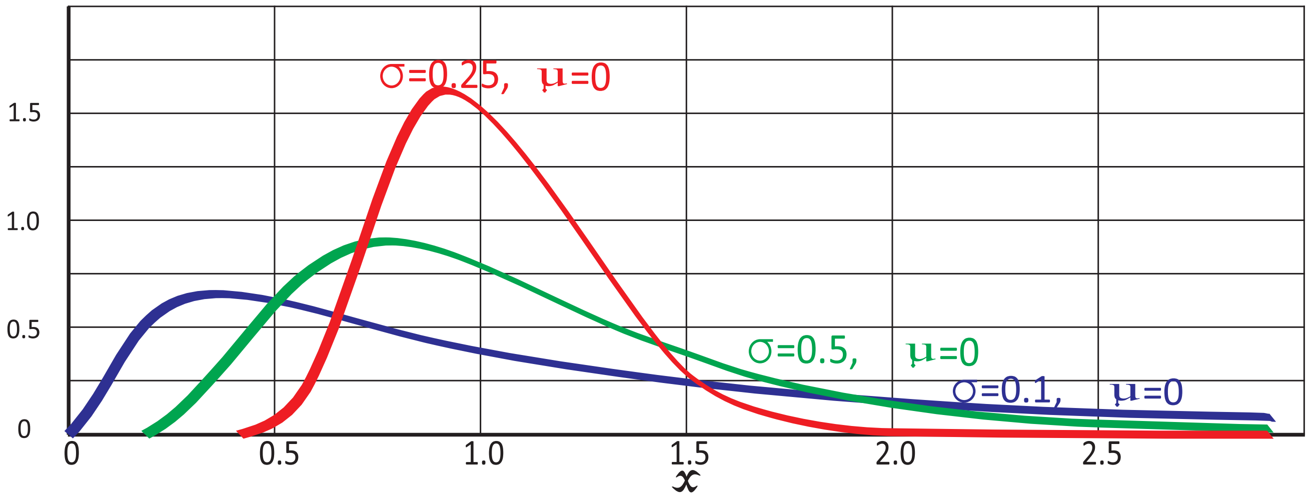
and a cumulative distribution function [9] is

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}}\int_0^x \frac{1}{x}e^{-\frac{(lnx-\mu)^2}{2\sigma^2}}dx.$$

In the lognormal distribution a logarithm of a variables has a normal distribution. Indeed, in this case we must have

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}}\int_0^x e^{-\frac{(lnx-\mu)^2}{2\sigma^2}}d(lnx) = \frac{1}{\sigma\sqrt{2\pi}}\int_0^x \frac{1}{x}e^{-\frac{(lnx-\mu)^2}{2\sigma^2}}dx.$$

The log-normal distribution has two parameters: μ და $\sigma > 0$.



Picture 1.

The lognormal distribution is right asymmetric – it reduces slowly on the right (see Picture 1). In order to obtain the density of a distribution characterized by a left asymmetry, one must consider a distribution whose density is obtained by mirroring the density of the lognormal distribution with respect to a straight line passing through the mode parallel to the OY axis. Since the mode of the lognormal distribution is

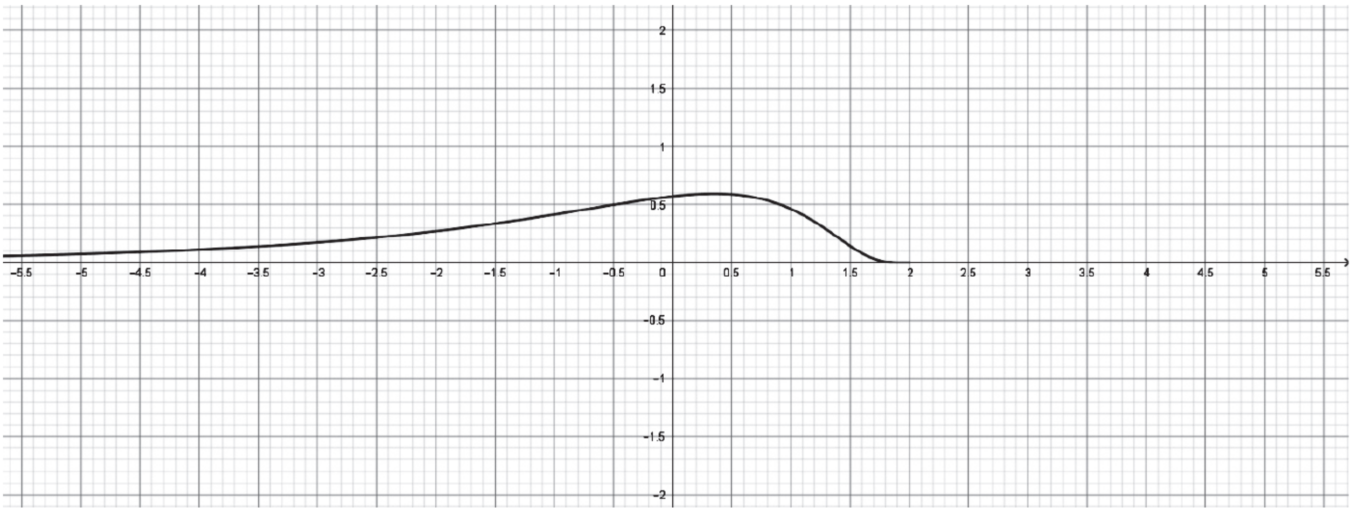
$$X_{mod} = e^{\mu-\sigma^2}, \text{ and } K_x = X_c/X_{mod},$$

then the density and the cumulative distribution function of left-asymmetric distribution will be

$$f(x) = \frac{1}{(2e^{\mu-\sigma^2}-x)\sigma\sqrt{2\pi}}e^{-\frac{(\ln(e^{\mu-\sigma^2}-x)-\mu)^2}{2\sigma^2}},$$

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}}\int_0^x \frac{1}{(2e^{\mu-\sigma^2}-x)}e^{-\frac{(\ln(e^{\mu-\sigma^2}-x)-\mu)^2}{2\sigma^2}}dx.$$

See the plot of the density of the left-asymmetric distribution in Picture 2 below when $a = 1,\ \sigma = 1$.



Picture 2.

Methods for computation of lognormal distribution parameters are described, for example, in [9]. Suppose we know the average monthly salary X_c (which is calculated by counting macroparameters and size of the population) and the value of the logical function when the upper bound of the integral is x_i:

$$\frac{1}{\sigma\sqrt{2\pi}}\int_0^{x_i}\frac{1}{x}e^{-\frac{(\ln(x-\mu)^2}{2\sigma^2}}dx = F(x_i). \quad (18)$$

Let us try to calculate the values of μ and σ. Since a mean of the lognormal distribution is

$$E(X) = e^{\mu+0.5\sigma^2},$$

then [9]

$$X_c = e^{\mu+0.5\sigma^2}. \quad (19)$$

(18) and (19) give the system of two equations with two variables

$$\begin{cases} \dfrac{1}{\sigma\sqrt{2\pi}}\displaystyle\int_0^{x_i}\frac{1}{x}e^{-\frac{(\ln(x-\mu))^2}{2\sigma^2}}dx = F(x_i) \\ e^{\mu+0.5\sigma^2} = X_c \end{cases}. \quad (20)$$

In the first equation of the system (20) let us apply the method of changing the variable

$$z = \frac{\ln(x-\mu)}{\sigma}.$$

Since

$$dz = \frac{1}{\sigma x}dx,$$

then

$$\frac{1}{\sigma\sqrt{2\pi}}\int_0^{x_i}\frac{1}{x}e^{-\frac{(\ln(x-\mu))^2}{2\sigma^2}}dx = \frac{1}{\sqrt{2\pi}}\int_0^{z_i}e^{-\frac{z^2}{2}}dz,$$

where

$$z_i = \frac{\ln(x_i-\mu)}{\sigma}. \quad (21)$$

We got a simpler system:

$$\frac{1}{\sqrt{2\pi}}\int_0^{z_i}e^{-\frac{z^2}{2}}dz = F(x_i),$$

$$e^{\mu-0.5\sigma^2} = X_c.$$

In this system we have two variables μ and σ, because z_i depends on μ.

Consequently

$$\frac{1}{\sqrt{2\pi}}\int_0^{z_i} e^{-\frac{z^2}{2}}dz = F(x_i).$$

With the standard method, for example, by Excel, we can find the value of z_i which depends on $F(x_i)$. From the second equation of the system (20) it follows that

$$\mu = lnX_c - 0.5\sigma^2. \quad (22)$$

Let us substitute (22) in (21):

$$z_i = \frac{lnx_i - lnX_c + 0.5\sigma^2}{\sigma}.$$

Then because $\sigma \neq 0$, we get a quadratic equation

$$\sigma^2 - 2z_i\sigma - 2\ln(X_c/x_i) = 0, \quad (23)$$

which solution is

$$\sigma = z_i \pm \sqrt{z_i^2 + 2\ln(X_c/x_i)}.$$

Only the positive root is economically significant:

$$\sigma = z_i + \sqrt{z_i^2 + 2\ln(X_c/x_i)}.$$

From (22) we get μ:

$$\mu = lnX_c - 0.5\sigma^2.$$

Let us now test the hypothesis that the distribution of the declared income in Georgia is lognormal. To do this, we use the data received from Revenue Service of the Ministry of Finance of Georgia (see also [15]) about the declared incomes of the population of Georgia in 2009-2017:

2009	2010	2011	2012	2013	2014	2015	2016	2017
762,855	846,045	987,426	1,118,636	1,104,222	1,181,484	1,116,625	1,185,065	1,189,718

List 3. The number of people with declared income, 2009-2017.

1-100	100-200	200-300	300-400	400-500	500-600	600-800	800-1000	1000-1250	1250-1666
205,435	135,886	93,448	79,784	47,829	34,594	51,159	32,846	25,209	23,315

1666-2083	2083-2500	2500-3333	3333-4166	4166-5000	5000-5833	5833-6666	6666-7500	7500-8333	>8333
11,046	6,518	7,178	3,230	1,766	1,126	703	472	331	1,980

List 4. The distribution of the population by declared incomes, 2009

1-100	100-200	200-300	300-400	400-500	500-600	600-800	800-1000	1000-1250	1250-1666
297469	134559	103185	84494	75525	60955	90192	63427	55643	56682

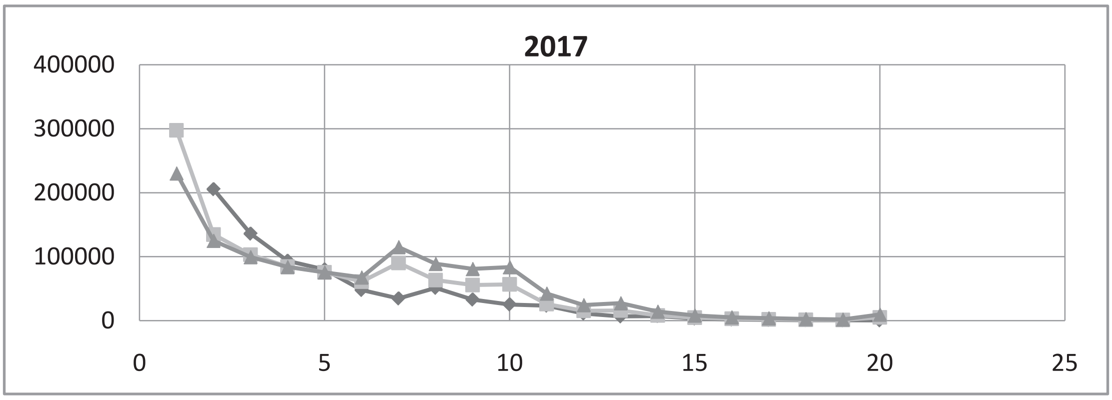
1666-2083	2083-2500	2500-3333	3333-4166	4166-5000	5000-5833	5833-6666	6666-7500	7500-8333	>8333
25878	15266	15887	8093	4646	2929	1937	1339	1097	5128

List 5. The distribution of the population by declared incomes, 2013

1-100	100-200	200-300	300-400	400-500	500-600	600-800	800-1000	1000-1250	1250-1666
230,280	125,013	99,417	83,823	75,525	67,610	115,023	88,848	80,873	83,744

1666-2083	2083-2500	2500-3333	3333-4166	4166-5000	5000-5833	5833-6666	6666-7500	7500-8333	>8333
42,423	24,361	27,309	14,204	8,141	5,233	3,784	2,609	1,884	9,614

List 6. The distribution of the population by declared incomes, 2017



Picture 3. The distribution of the population by declared incomes:
◆ - 2009, ▲ - 2013, ■ - 2017

Based on Table 6 we will create a new Table 7:

2017	x_i	N	$\Delta F(x_i)\%$	$F(x_i)\%$	$F(x_i)$	2017	x_i	N	$\Delta F(x_i)\%$	$F(x_i)\%$	$F(x_i)$
1	0		0	0	0	12	2083	42423	3.565803	91.83512	0.918351
2	100	230280	19.35585	19.35585	0.193558	13	2500	24361	2.047628	93.88275	0.938828
3	200	125013	10.50778	29.86363	0.298636	14	3333	27309	2.295418	96.17817	0.961782
4	300	99417	8.35635	38.21998	0.3822	15	4166	14204	1.193896	97.37207	0.973721
5	400	83823	7.045619	45.2656	0.452656	16	5000	8141	0.68428	98.05635	0.980563
6	500	75525	6.348143	51.61374	0.516137	17	5833	5233	0.439852	98.4962	0.984962
7	600	67610	5.682859	57.2966	0.572966	18	6666	3784	0.318059	98.81426	0.988143
8	800	115023	9.668089	66.96469	0.669647	19	7500	2609	0.219296	99.03355	0.990336
9	1000	88848	7.467988	74.43268	0.744327	20	8333	1884	0.158357	99.19191	0.991919
10	1250	80873	6.797661	81.23034	0.812303	21	16000	6014	0.505498	99.69741	0.996974
11	1666	83744	7.038979	88.26932	0.882693	22	>16000	3600	0.302593	100	1

List 7.

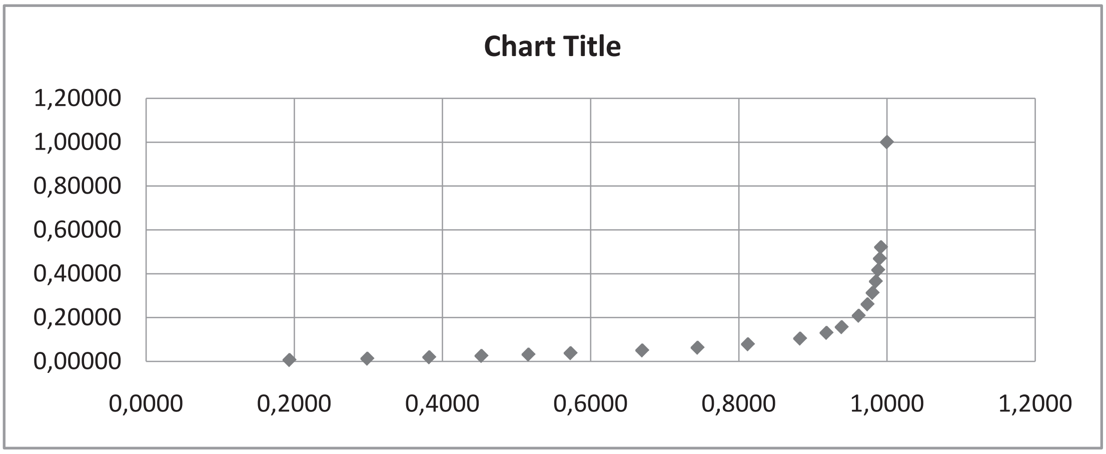
The first column shows the numbers of the intervals, in the second - the intervals of declared income, in the third - the number of persons corresponding to the intervals. For example, this table shows that the monthly income from 0 to 100 GEL have 230280 people, from 100 to 200 GEL - 125013 people, etc. The fourth column shows how many percent are the third column data from the number of persons with declared income (1189718 person), in the fifth column - the accumulated value of these percent. In the sixth column, these percentages are converted to fractions of a unit.

At first, let us calculate Gini index of the declared monthly income for 2017. Let's agree that when calculating the Gini index we will not take into account incomes exceeding 16,000 GEL per month, since in this case the Gini index will be inadequate. Let us construct the Lorenc curve. In List 8 the data are translated into a convenient form for constructing Lorenz curve:

x_i	$F(x_i)$		x_i	$x_i/16000$		$F(x_i)$	$x_i/16000$
0	0		0				
100	0.193558		100	0.00625		0.193558	0.00625
200	0.298636		200	0.0125		0.298636	0.0125
300	0.3822		300	0.01875		0.3822	0.01875
400	0.452656		400	0.025		0.452656	0.025
500	0.516137		500	0.03125		0.516137	0.03125
600	0.572966		600	0.0375		0.572966	0.0375
800	0.669647		800	0.05		0.669647	0.05
1000	0.744327		1000	0.0625		0.744327	0.0625
1250	0.812303		1250	0.078125		0.812303	0.078125
1666	0.882693		1666	0.104125		0.882693	0.104125
2083	0.918351		2083	0.130188		0.918351	0.130188
2500	0.938828		2500	0.15625		0.938828	0.15625
3333	0.961782		3333	0.208313		0.961782	0.208313
4166	0.973721		4166	0.260375		0.973721	0.260375
5000	0.980563		5000	0.3125		0.980563	0.3125
5833	0.984962		5833	0.364563		0.984962	0.364563
6666	0.988143		6666	0.416625		0.988143	0.416625
7500	0.990336		7500	0.46875		0.990336	0.46875
8333	0.991919		8333	0.520813		0.991919	0.520813
16000	0.996974		16000	1		**1**	1
>16000	1		>16000				

List 8.

Here we divide x_i to 16000, then rearrange $x_i/16000$ and $F(x_i)$. The last two columns will give us a Lorenz curve:



Picture 4

If p_i part of population, $i = 1,2, \dots, n$, $p_1 + p_2 + \cdots + p_n = 1$, owns q_i part of a wealth, where $i = 1,2, \dots, n$, $q_1 + q_2 + \cdots + q_n = 1$, with a cumulative form $s_i = q_1 + q_2 + \cdots + q_i$, then Gini index can be calculated by formula [2]

$$G = 1 - \sum_{i=1}^{n} p_i(s_i + s_{i-1}).$$

The plot in Picture 4 is almost rectangular, therefore the distribution of declared income is extremely uneven. It is easy to calculate that the Gini index corresponding to the plot on Picture 4 is equal to $G = 0.89$.

According to Geostat [14] the average annual salary in the public and private sectors of Georgia in 2017 was X_c =989 GEL. For example, let us choose i = 1; According to table 3,

$$F(x_1) = 0.193558,$$

i.e.

$$\frac{1}{\sqrt{2\pi}}\int_0^{z_1} e^{-\frac{z^2}{2}}dz = 0.193558.$$

By Excel we can calculate that

$$z_1 = -0.86486.$$

Since

$$\ln(X_c/x_i) = 2.291524,$$

the quadratic equation (23) takes the form

$$\sigma^2 + 1.72972\sigma - 4.583048291 = 0.$$

It has the positive solution

$$\sigma = 1.444044073.$$

Then from (22) we have

$$\mu = ln989 - 0.51.444044073^2 = 5.85406269.$$

The cases $i = 2,3, \dots ,20$ are considered in List 9:

i	z_i	X_c/x_i	$\ln\left(\frac{X_c}{x_i}\right)$	σ	μ	X_{mod}	K_x
1	-0.86486	9.89	2.291524	1.444044	5.854063	43.32801	22.82587876
2	-0.52833	4.945	1.598377	1.336045	6.004186	67.97694	14.54905095
3	-0.29971	3.296667	1.192912	1.273711	6.085524	86.76335	11.39882192
4	-0.11895	2.4725	0.90523	1.231827	6.137995	101.5546	9.738605958
5	0.040462	1.978	0.682086	1.20914	6.165684	110.3508	8.962327139
6	0.183931	1.648333	0.499765	1.200474	6.176126	113.8622	8.68593753
7	0.438938	1.23625	0.212083	1.224325	6.147208	104.4007	9.473114755
8	0.656743	0.989	-0.01106	1.296422	6.05634	79.48999	12.4418183
9	0.886416	0.7912	-0.2342	1.449733	5.845832	42.27126	23.39651375
10	1.188558	0.593637	-0.52149	1.796586	5.282834	7.807727	126.6693799
11	1.394067	0.474796	-0.74487	2.067625	4.759157	1.622688	609.4826824
12	1.545005	0.3956	-0.92735	2.274621	4.309745	0.421408	2346.894138
13	1.771747	0.29673	-1.21493	2.613898	3.480462	0.035014	28245.6943
14	1.938529	0.237398	-1.43802	2.877604	2.756392	0.003989	247931.6508
15	2.065528	0.1978	-1.6205	3.078152	2.159184	0.000665	1487386.178
16	2.169088	0.169553	-1.77459	3.244149	1.634444	0.000138	7179570.988
17	2.261717	0.148365	-1.90808	3.401544	1.111444	2.87E-05	34475203.52
18	2.339125	0.131867	-2.02596	3.530586	0.664175	7.5E-06	131900542
19	2.405241	0.118685	-2.13128	3.639183	0.274867	2.33E-06	424101439.1
20	2.744959	0.061813	-2.78365	4.147634	-1.70474	6.15E-09	1.6094E+11

List 9.

For example let us write the lognormal distribution for data from the row 8:

$$F(x_i) = \frac{1}{1.2964\sqrt{2\pi}}\int_0^{x_i}\frac{1}{x}e^{-\frac{(\ln(x-6.0563))^2}{2*1.2964^2}}dx$$

From List 9 it is clear that with a certain accuracy it can be argued that the intervals of incomes 300-800 GEL and 4166-7500 GEL is described by the log-normal distribution. Such a conclusion can be drawn because the variance and the expectation change at the same time in an insignificant way.

REFERENCES:

1. Arnold, B. C., (2015). Pareto Distributions. London: CRC Press.
2. Bellu, L. & Liberati, P., (2006). Inequality and Axioms for its Measurement. N.-Y.: FAO.
3. Cowell, F. A., (2009). Measuring Inequality. London: Oxford University Press.
4. University of Texas., 2018. Measuring Inequality Project. https://utip.lbj.utexas.edu/tutorials.html.
5. World Bank, (2017). Percentage share of income or consumption. http://wdi.worldbank.org/table/1.3.
6. Goskomstat Rossii, (1996). Methodological provisions on statistics. Moscow (in Russian)
7. Kolmakov I. B., (2008). Methods of forecasting poverty indicators, taking into account disadvantaged groups. Problemy Prognozirovania, Issue 5, pp. 95-109 (in Russian).
8. Kolmakov I. B., (2015). Methodology of calculations and analyzes of integral estimates of the polarization of money incomes of the population. Voprosy Statistiki, pp. 23-36 (in Russian).
9. Kolmakov I. B., (2016). Conjugation of the logarithmically normal distribution of the population in terms of household income with the Pareto distribution. Audit i Finansovyj Analiz, pp. 124-131 (in Russian).
10. Rozanov V. B., (2007). Economic structure of the Russian society. In: Econophysics. Moscow, pp. 560-600 (in Russian).
11. 2018. https://m2b.ge. [Online] Available at: https://m2b.ge/post/177009-laris-kursis-dinamika-2017-wels-mizezebi-da-molodinebi [Accessed 7 July 2019].
12. Kakulia M., (2018). Middle class in Georgia: quantitative assessment, dynamics And profile. Tbilisi in Georgian).
13. GeoStat, (2016). http://census.ge/. [Online] Available at: http://census.ge/files/results/Census%20Release_GEO.pdf [Accessed 7 July 2019].
14. GeoStat, (2018). [Online] Available at: http://www.geostat.ge/?action=page&p_id=148&lang=geo [Accessed 27 09 2018].
15. Tabula, (2018). Only 0.6% of Georgian population has more than 5,000 GEL per month [Online] Available at: http://tbl.ge/2t1g [Accessed 29. 09. 2018] (in Georgian).

JEL Classification: C01; J10 https://doi.org/10.35945/gb.2019.07.016

STUDY OF THE DISTRIBUTION OF WEALTH IN THE MIDDLE AND TOP SEGMENTS OF THE POPULATION OF GEORGIA

QETEVAN PIPIA

Phd student
Samtskhe-Javakheti State University, Georgia
qetino.pipia@gmail.com

KEYWORDS: PARETO DISTRIBUTION, LOGNORMAL DISTRIBUTION, MODELS OF ECONOMICAL LAYERS OF SOCIETY.

For citation: Pipia Q. (2019), Study Of The Distribution Of Wealth In The Middle And Top Segments Of The Population Of The Population Of Georgia, *Globalization And Business, №7*, pp. 130-142. https://doi.org/10.35945/gb.2019.07.016

SUMMARY

Due to economic, social, political and other differences, different sectors of society are subject to different laws of distribution. Among these laws are Pareto distribution, the normal distribution, the lognormal distribution, and so on. It is noteworthy that the higher, richer stratum of a society more often depends on the Pareto distribution. As for the poor and middle class, there was an attempt to build their model using a normal distribution. But later it turned out that more accurate results are provided by a lognormal distribution. The article attempts to build a model of the distribution of the upper layers of the population of Georgia in terms of per capita GDP consumption (according to the World Bank) using Pareto distribution. As for the other layers, due to the lack of data in GeoStat, when trying to build a model using a lognormal distribution, data on the population's declared income are used, obtained from the Revenue Service of the Ministry of Finance of Georgia, hoping that this data correlates with the population distribution by GDP consumption.