Why the super-rich may be indifferent to income growth of their own countries?

Branko Milanovic* and Nassim Nicholas Taleb†
*Luxembourg Income Study Center, City University of New York
†School of Engineering, New York University

Abstract—The superwealthy do not necessarily have their interests aligned with the rest of the economy, as they depend more on inequality than wealth. The further away "in the tail" the more severe the effect. We show how this effect holds across the various families of wealth distributions used by economists. We discuss the mathematical property, that is, the sensitivity of partial tail expectation to dispersion parameters (which includes the Gini coefficient) increases as one goes in the tail, while the effect of the changes in the mean decreases. We show empirical illustrations.

1 Introduction

It is often said that, in very unequal societies, the rich compose a group apart. Not only are their social mores and consumption patterns different, but their fortunes seem dissociated with those of the rest of the population. More recently, there has been an argument that the rich (the top 1%) from different nations form a group apart, a global "superclass" (Freeland, 2012 [1]). The difference is even more acute in the .1%, which represents a large class of close to 700,000 persons. For someone in the business of selling luxury apartments or expensive watches, changes in GDP are far less relevant than changes in Gini or other indices of inequality. This is reflected in the dynamics of real estate prices: many major cities exhibited a rise in the average price of apartments between 2008 and 2015, coupled with a stagnation or even decline in the median value over the same period.

The objective of this note is to consider whether there is theoretical and empirical substance to the claim that in high inequality societies income of the super-rich is in some sense decoupled from the income of the rest of society. We do not mean it in an obvious sense that the rich simply have higher income than the others. What we mean is to look at the income gains that the rich can make from an overall increase in national income (while keeping the distribution unchanged) versus the gains that they can make from a further widening of income distribution (while keeping mean income the same). We shall show that this particular trade-off varies in function of income class, and that especially for the top income classes, the gains from greater inequality tend to be disproportionately high compared to the gains from an increased overall income without a change in the distribution.

1.1 Summary of the Paper

In Section 2, we show a general mathematical derivation of this relation, followed by applications to what distributions are used for wealth and income inequality. In Section 3, we present an empirical analysis, based on household survey data from more than 100 countries. Short conclusions close the paper.

2 Mathematical Properties

2.1 General framework for probability distributions

There is a necessarily mathematical relation, well known by risk takers in mathematical finance (derivatives), namely that remote parts of the distribution –the tails– are less sensitive to changes in the mean, and more sensitive to other parameters controlling the dispersion[1]. The dispersion parameters can be the *scale* (which in the special case of class of finite-variance distributions would be the variance) or the tail exponent –or both. It means that as one's income increases, changes in the mean income (i.e., from GDP increase) would be less relevant than changes in inequality. This mathematical property can be made clear in the following exposition.

We are interested in the expected income (or wealth) for the segment above a certain level K, $\mu_K = \int_K^\infty x \, \mathrm{d}F(x)$, $x \in (l, \infty)$ where $l \geq 0$. The measure under consideration should be as general as possible to cover the various approaches to wealth distributon as well as the following transformation groups.

1. Called in Taleb(2015) [2] the "delta-vega" transfer as remote out of the money options are more reactive to changes in volatility than to those in the underlying security.

Consider a family of continuous one-dimensional unimodal probability distribution with density function $\phi \in C^2 : [0 \wedge l + \delta, \infty) \to [0,1]$, with $s \in \mathbb{R}^+$, where $\lambda(.)$ is a slowly moving function with respect to x defined as $\forall x > 0, \lim_{t\to\infty} \frac{\lambda(tx)}{\lambda(x)} = 1$:

$$\phi(x) := \frac{1}{s}\lambda(x, \alpha+\gamma)\, z\left(\frac{x-\delta}{s}, \alpha+\gamma\right). \tag{1}$$

The general form ϕ allows us to consider most (if not all) forms of one-tailed continuous unimodal distributions used for wealth and income, and location, scale, and shape transformation groups, with two distinct classes.

We define two broad classes of unimodal one-tailed distributions:

1) The function z depends on x not α (Class 1), allowing the location and scale transformation group (with adjustment in the left tail of the support).
2) The function z depends on x and α (Class 2) and $\lambda(.)$ ceases to be a function of x for large values of x , thus allowing both location-scale transformation groups and shape transformation groups.

We can show that there exists a value $x > x_\theta$ defined as the "tail" in Taleb (2015) [2] above which "large values" of x have the following properties.

i) For $K > x_\theta$, the probability density ϕ depends more on the "scale" s (controlling dispersion) than the "shift" δ (controlling the mean).

ii) The effect monotonically increases at higher values of K; in other words for all $\exists x_{\theta'} > x_\theta : \forall K > x'_\theta$, $\left.\frac{\partial^2 \phi(x)}{\partial s^2} \middle/ \frac{\partial^2 \phi(x)}{\partial \delta^2}\right|_{s=1,\delta=0} \geq 0.$

Further, when λ has a nonzero derivative with respect to the tail exponent α,

iii) $\exists x_{\theta''} : \forall K > x_{\theta''}$, $\left.\frac{\partial \phi(x)}{\partial \alpha} \middle/ \frac{\partial \phi(x)}{\partial \delta}\right|_{\delta=0, x=K} \geq 1$, and

iv) $\exists x_{\theta'''} : \forall K > x_{\theta'''}$, $\left.\frac{\partial^2 \phi(x)}{\partial \alpha^2} \middle/ \frac{\partial^2 \phi(x)}{\partial \delta^2}\right|_{\delta=0, x=K} \geq 0.$

This property of densities transfers to any integral transform of $x \in (K, \infty)$ such as partial expectations $\mu_K = \int_K^\infty x \,\mathrm{d}F(x)$ or the integral tranform of any non-decreasing function $f(x)$, namely $\int_K^\infty f(x)\,\mathrm{d}F(x)$, which should also show the same relative dependence on scale vs shift. Under consideration for us is the share of the top $\frac{\int_K^\infty x \,\mathrm{d}F(x)}{\int_l^\infty \mathrm{d}F(x)}$.

2.2 Quantiles for inequality measures

We retain the same symbol K for the threshold in the following exposition. Consider a one-tailed random variable $X \in [x_{\min}, \infty)$. Let us define the *quantile share*

$$\kappa_q := q\frac{\mathbb{E}[X|X > K(q)]}{\mathbb{E}[X]}$$

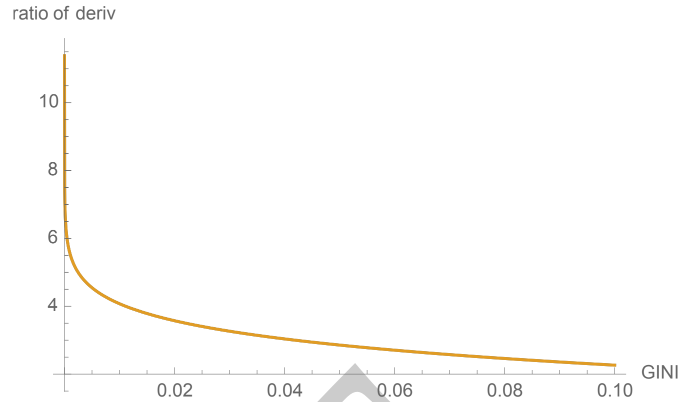

Fig. 1. Lognormal case, ratio of derivatives of mean/inequality in the top 10%, from Eq 6.

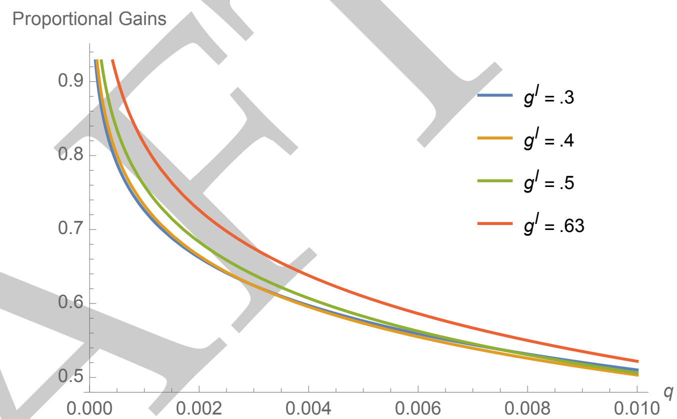

Fig. 2. Lognormal case; sensitivity to inequality in the class ranging from the "very rich" (1%) to the "super-rich" classes (top .1%). We show the relative effect on income from .1 changes in Gini

where $K(q) = \inf\{h \in [x_{min}, +\infty), \mathbb{P}(X > h) \leq q\}$ is the exceedance threshold for the probability q.

For a given sample $(X_k)_{1\leq k\leq n}$, the "empirical" or natural estimator (or observed quantile share) $\widehat{\kappa}_q := \frac{q^{th}\text{percentile}}{total}$, can be expressed, as

$$\widehat{\kappa}_q := \frac{\sum_{i=1}^n \mathbb{1}_{X_i > \hat{K}(q)} X_i}{\sum_{i=1}^n X_i}$$

where $\hat{K}(q)$ is the estimated exceedance threshold for the probability q :

$$\hat{K}(q) := \inf\{K : \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{x>K} \leq q\}$$

2.3 Class 1: Lognormal-Gibrat Case

Class 1 Distributions

The class includes exponential, folded normal, and the more commonly used Lognormal-Gibrat distribution. Further, the Lognormal belongs to the class of subexponential distributions, which is the fattest tail class

TABLE 1
Lognormal Case: The contribution/share in income for class $\frac{1}{m}\mu_K(g^l,\mu)$ scaled by the mean. (Should be interpreted as all the subparts are linear to changes in mean.)

GINI	Top 100%	Bottom 10%	Bottom 50%	Top 50%	10%	1%	.1%	.01%	.001%	.0001%
0.1	1.	0.072	0.429	0.571	0.135	0.016	0.002	0.	0.	0.
0.2	1.	0.051	0.36	0.64	0.178	0.025	0.003	0.	0.	0.
0.3	1.	0.034	0.293	0.707	0.231	0.037	0.005	0.001	0.	0.
0.4	1.	0.022	0.229	0.771	0.295	0.057	0.009	0.001	0.	0.
0.5	1.	0.013	0.17	0.83	0.372	0.085	0.016	0.003	0.	0.
0.6	1.	0.007	0.117	0.883	0.464	0.128	0.029	0.006	0.001	0.
0.7	1.	0.003	0.071	0.929	0.573	0.195	0.052	0.012	0.003	0.001
0.8	1.	0.001	0.035	0.965	0.702	0.304	0.101	0.028	0.007	0.002
0.9	1.	0.	0.01	0.99	0.852	0.5	0.222	0.082	0.026	0.008

TABLE 2
Lognormal Case: "Tradeoff", that is relative change in income from .01 changes in Gini: $\frac{\mu_{K(q}(g^l+0.01)-\mu_{K(q}(g^l)}{\mu_{K(q}(g^l)}$

GINI	Top 100%	Bottom 10%	Bottom 50%	Top 50%	10%	1%	.1%	.01%	.001%	.0001%
0.1	0.	−0.034	−0.016	0.012	0.029	0.046	0.059	0.07	0.08	0.089
0.2	0.	−0.037	−0.019	0.011	0.027	0.044	0.057	0.068	0.079	0.088
0.3	0.	−0.042	−0.022	0.009	0.025	0.042	0.056	0.068	0.079	0.088
0.4	0.	−0.048	−0.027	0.008	0.024	0.042	0.056	0.069	0.08	0.09
0.5	0.	−0.056	−0.033	0.007	0.023	0.042	0.057	0.071	0.083	0.094
0.6	0.	−0.069	−0.042	0.006	0.022	0.042	0.06	0.075	0.089	0.102
0.7	0.	−0.09	−0.057	0.004	0.021	0.044	0.064	0.083	0.099	0.115
0.8	0.	−0.13	−0.087	0.003	0.02	0.048	0.074	0.098	0.12	0.141
0.9	0.	−0.243	−0.175	0.002	0.019	0.057	0.098	0.139	0.178	0.215

below the power laws. Indeed the lognormal has tails thick enough as to be confused for a power law outside extremely large deviations.

Consider a standard lognormal; we parametrize ϕ with $(\log(m) - \frac{1}{2}\sigma^2, \sigma)$ to get a mean of m (independent from σ)[2] and variance of $m\sqrt{(e^{\sigma^2}-1)}$, $z(x) := e^{\frac{\left(-\log(m)+\frac{\sigma^2}{2}+\log(x)\right)^2}{2\sigma^2}}$, $\lambda = \frac{1}{\sqrt{2\pi}\sigma}$. The behavior of the k^{th} moment is examined with $k = m-1$. The first condition for the dispersion to matter more than the mean:

$$K \geq m\, e^{\frac{\sigma^2}{2}}$$

and for second derivatives:

$$K \geq m\, e^{\sigma^2+\frac{1}{2}\sqrt{\sigma^2+4}\sigma}$$

Let g be the scaled (i.e., mean adjusted) standard Gini coefficient, $g^l \in (0,1)$; from the conventional formulations for a scaled Gini. Where F is the cumulative density of the distribution under concern:

$$g = \frac{1}{m}\int \Phi(y)(1-\Phi(y))\mathrm{d}y. \tag{2}$$

which, for a Lognormal distribution can be computed as (the literature, say Theil [3], Cowell [4], etc.) provides results though we could not find derivations, which we redid in Appendix A):

$$g^l = 2\Psi\left(\frac{\sigma}{2}\right) - 1 \tag{3}$$

where $\Psi(.)$ is the cumulative standard Gaussian. Inverting to express σ in terms of Gini:

$$\sigma = 2\ \mathrm{erf}^{-1}(g^l)$$

where $\mathrm{erf}(.)$ is the error function $\mathrm{erf}(z) = \frac{2}{\sqrt{\pi}}\int_0^z e^{-t^2}\mathrm{d}t$. We remark that the Gini coefficient does not directly include the mean.

Recall the partial mean above K, $\mu_K(g^l, m) =$

2. The conventional parametrization of the lognormal (μ, σ) has the problem that its mean depends on the variance, with expectation $\mu - \frac{1}{2}\sigma^2$. We therefore separate the mean.

TABLE 3
Power Law Case: The contribution/share in income for class $\frac{1}{\mu}\mu_K(g^p, \mu)$ scaled by the mean. (Should be interpreted as all the subparts are linear to changes in mean.)

GINI	Top 100%	Bottom 10%	Bottom 50%	Top 50%	10%	1%	.1%	.01%	.001%	.0001%
0.1	1.	0.083	0.433	0.567	0.152	0.023	0.004	0.001	0.	0.
0.2	1.	0.068	0.37	0.63	0.215	0.046	0.01	0.002	0.	0.
0.3	1.	0.055	0.311	0.689	0.289	0.084	0.024	0.007	0.002	0.001
0.4	1.	0.044	0.257	0.743	0.373	0.139	0.052	0.019	0.007	0.003
0.5	1.	0.035	0.206	0.794	0.464	0.215	0.1	0.046	0.022	0.01
0.6	1.	0.026	0.159	0.841	0.562	0.316	0.178	0.1	0.056	0.032
0.7	1.	0.018	0.115	0.885	0.666	0.444	0.296	0.197	0.131	0.087
0.8	1.	0.012	0.074	0.926	0.774	0.599	0.464	0.359	0.278	0.215
0.9	1.	0.006	0.036	0.964	0.886	0.785	0.695	0.616	0.546	0.483

TABLE 4
Power Law Case: Tradeoff for 1 point of Gini.

Gini	Top 100%	Bottom 10%	Bottom 50%	50%	10%	1%	.1%	.01%	.001%	.0001%
0.1	0.	−0.019	−0.015	0.011	0.038	0.078	0.12	0.163	0.208	0.254
0.2	0.	−0.02	−0.016	0.01	0.032	0.065	0.1	0.135	0.172	0.21
0.3	0.	−0.021	−0.018	0.008	0.027	0.056	0.085	0.114	0.145	0.176
$\mathbf{0}$.4	0.	−0.023	−0.02	0.007	0.024	0.048	0.072	0.098	0.124	0.15
0.5	0.	−0.026	−0.024	0.006	0.021	0.042	0.063	0.085	0.107	0.13
0.6	0.	−0.031	−0.029	0.005	0.018	0.036	0.055	0.074	0.094	0.113
0.7	0.	−0.039	−0.037	0.005	0.016	0.032	0.049	0.065	0.082	0.1
0.8	0.	−0.055	−0.053	0.004	0.014	0.029	0.043	0.058	0.073	0.089
0.9	0.	−0.104	−0.103	0.004	0.013	0.026	0.039	0.052	0.066	0.079

$\int_K^\infty x f(x, g^l)\, \mathrm{d}x$. Hence

$$\mu_K(g^l, m) = \frac{1}{2} m \left(\operatorname{erf}\left(\frac{2\operatorname{erf}^{-1}(g)^2 - \log(K) + \log(m)}{2\sqrt{2}\operatorname{erf}^{-1}(g)} \right) + 1 \right) \quad (4)$$

We thus are now able to express the total income (or wealth) to those with income (or wealth) in excess K in terms of g a measure of inequality.

Now expressing K in terms of quantiles, $q : K = K(q)$:

$$\mu_{K(q)}(g^l, m) = \frac{1}{2} m \left(\operatorname{erf}\left(\sqrt{2}\operatorname{erf}^{-1}(g^l) - \operatorname{erfc}^{-1}(2q) \right) + 1 \right) \quad (5)$$

We notice in Eq. 5 that expressing K as a function of q removes m from the equation except as a scaling factor. Let us consider ratio of first derivatives, that is, the ratio of sensitivity of the expectation per quantile to the Gini g^l over that of the mean m:

$$\frac{\frac{\partial \mu_K(g^l, m)}{\partial g^l}}{\frac{\partial \mu_K(g^l, m)}{\partial m}} = \frac{\sqrt{2} m \exp\left(\operatorname{erf}^{-1}(g^l)^2 - \left(\operatorname{erfc}^{-1}(2q) - \sqrt{2}\operatorname{erf}^{-1}(g^l) \right)^2 \right)}{\operatorname{erf}\left(\sqrt{2}\operatorname{erf}^{-1}(g^l) - \operatorname{erfc}^{-1}(2q) \right) + 1} \quad (6)$$

Figure 1 shows the effect of the ratio of derivatives from Eq. 6. Figure 2 shows the adjusted discretization of Eq. 6, here the sensitivity for .1 changes in Gini, namely $\frac{\mu_{K(q)}(g^l + 0.1) - \mu_{K(q)}(g^l)}{\mu_{K(q)}(g^l)}$.

2.4 Paretan tailed Class

The results of the next section apply to a broad class, that is a standard Pareto distribution, its Lomax version (shifted Pareto), Singh-Maddala, or Lévy-Stable, as well as distributions of that class we have derived ourselves, such as Folded Student-T or folded Lévy-Stable distributions. Owing to the properties of $\lambda(.)$ in Eq. 1, we can treat the distributions as being statistically "the same" for

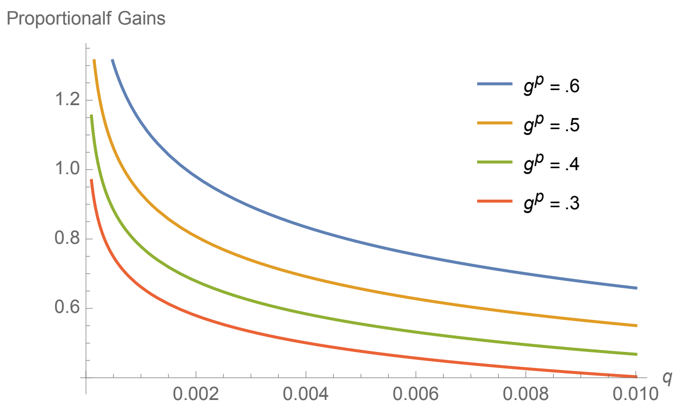

Fig. 3. Power law case; sensitivity to inequality in the class ranging from the "very rich" (1%) to the "super-rich" classes (top .1%). We show the relative effect on income from .1 changes in Gini

large values of K, hence in the lower centiles. Figure 4 illustrates the point.

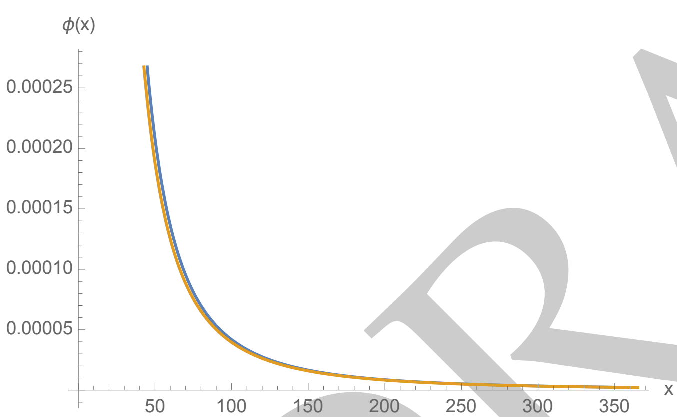

Fig. 4. Comparing the PDF of an Alpha-Stable distribution (α, β, μ, s) with exponent $alpha = 5/4$, symmetry parameter $\beta = 1$, mean $\mu = 4$, and scale $\sigma = 1.652$ to a Pareto distribution with tail $\alpha = 5/4$ and minimum value 1. The distributions are statistically indistinguishable at values above the mean, hence our result does not depend on the specifics of Class 2 distribution.

Consider $\phi_\alpha(x)$ the density of a α-Pareto distribution bounded from below by $x_{\min} > 0$, in other words: $\phi_\alpha(x) = \alpha x_{\min}^\alpha x^{-\alpha-1} \mathbb{1}_{x \geq x_{\min}}$, and $\mathbb{P}(X > x) = \left(\frac{x_{\min}}{x}\right)^\alpha$. The mean $\mu = \frac{\alpha x_m}{\alpha - 1}$.

Rewriting Eq. 2 to get g^l, the Gini coefficient for a power law:

$$g^p = \frac{1}{2\alpha - 1} \tag{7}$$

Expressing K in terms of quantiles, $q : K = K(q)$, and getting $K = x_m q^{-1/\alpha}$, we have the partial mean

$$\mu_K(g^p, \mu) = \mu \left(q^{\frac{2g^p}{g^p+1}} \right)^{\frac{1}{2}\left(\frac{1}{g^p} - 1\right)} \tag{8}$$

and finally,

$$\frac{\partial \mu_K(g^p, m)}{\partial g^p} \bigg/ \frac{\partial \mu_K(g^p, \mu)}{\partial \mu} = -\frac{2\mu \log(q)}{(g^p + 1)^2} \tag{9}$$

the discretization of which can seen in Figure 3.

TABLE 5
Share of total income received by each ventile of national income distribution

Ventile	Mean ventile share in total income (in %)	STD of ventile share (in %)	Income gain from 1 STD increased share (in % of own income)
1st	1.06	0.45	42.2
2nd	1.59	0.53	33.3
3rd	1.9	0.56	29.6
4th	2.17	0.56	26.4
5th	2.42	0.58	23.9
6th	2.67	0.58	21.6
7th	2.9	0.57	19.8
8th	3.15	0.56	17.8
9th	3.4	0.55	16.1
10th	3.68	0.53	14.5
11th	3.98	0.5	12.6
12th	4.3	0.47	11
13th	4.67	0.43	9.1
14th	5.09	0.38	7.4
15th	5.6	0.32	5.8
16th	6.22	0.28	4.5
17th	7.04	0.32	4.5
18th	8.22	0.58	7.1
19th	8	0.93	11.6
20th (top)	19.51	5.65	29
Total	100		

3 Empirical Illustration

Income distributions in the nations of the world differ a lot. Table 1 shows, using the data from 116 countries around the year 2008, the average ventile shares and their standard deviations. All ventile shares are calculated from micro data provided by nationally-representative household surveys.[3] Consider the first ventile (the poorest 5 percent of population ranked by income per capita). On average, across countries, the poorest ventile receives just slightly above 1% of total national income. In more equal counties, the share of the bottom is greater (almost 2%), in less equal, it is less (under $\frac{1}{2}$ of 1%). The standard deviation of the

3. The data are available at Harvard database the information of which is available in Milanovic (2015) [5].

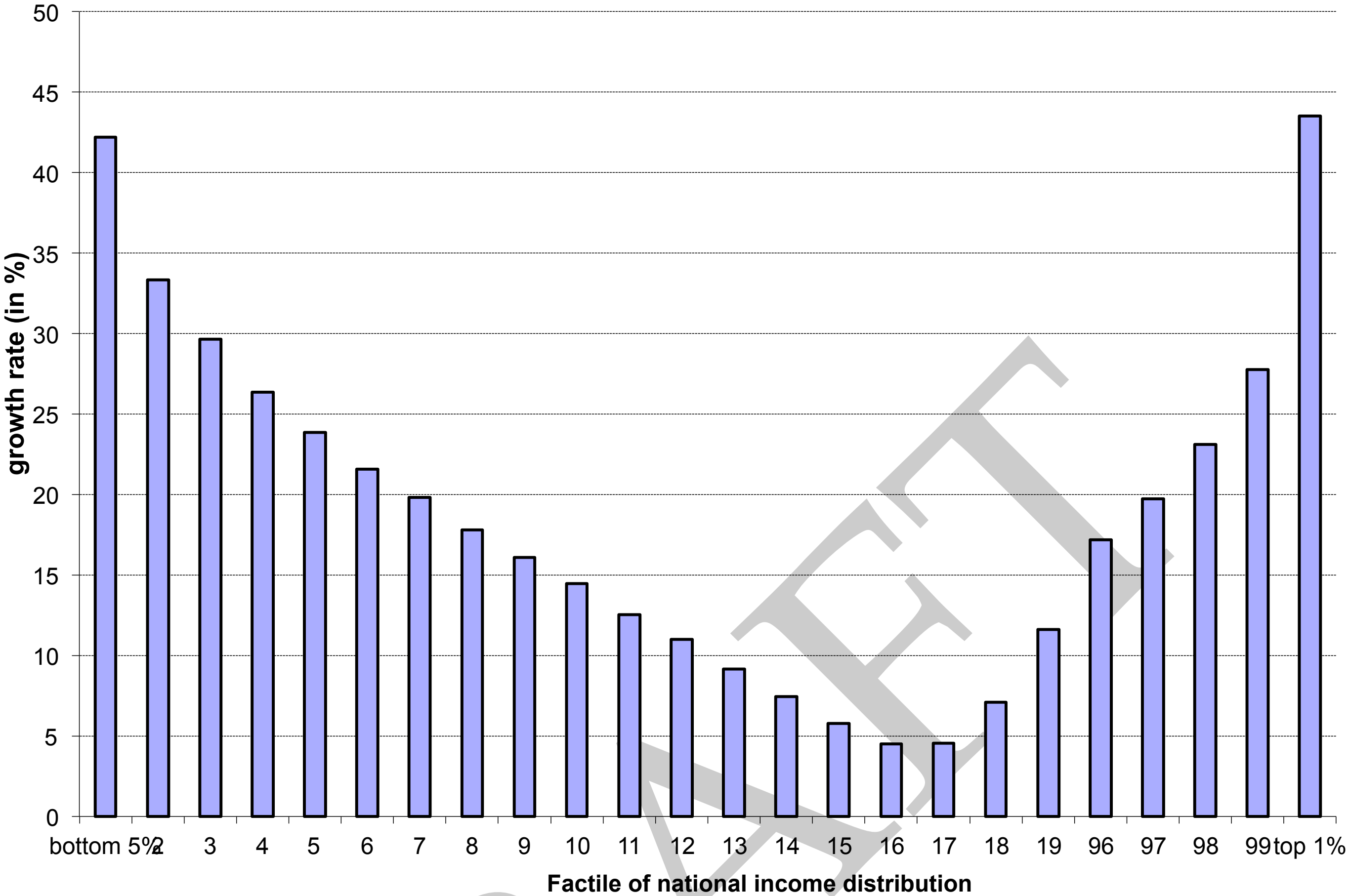

Fig. 5. caption

bottom ventile share is 0.45 percent (column 2). Thus the gain that an average person placed in the bottom ventile would make from moving from a distribution that "allocated" to him the average worldwide share of the bottom ventile to a distribution that would be more favorable to the bottom ventile (by one standard deviation), would be 42 percent (0.45/1.064). We call such a move to "the mean share + 1 standard deviation" a "conceivable" distributional change because the change represents something that is not far-fetched but observable in the empirical reality of national income distributions. The same interpretation of Table 1 applies to all other ventiles.

It can be readily seen that the sensitivity, expressed in terms of own income, is very high for the bottom and top ventiles. For the bottom three ventiles and for the highest ventile it amounts to about 30% of their income. The gain is much more modest for the middle ventiles while for the ventiles 13-18, it stays under 10%. The result is driven by the well-known observation (see Palma 2011 [6], also Milanovic 2008 [7], p. 29)) that middle fractiles tend to get the same share of national income, whether they are in unequal or equal countries. Consequently, if a person belongs to these middle fractiles, his income will not depend on whether his country is equal or unequal, but almost fully on whether the country is rich or poor. In other words, for such a person a way toward higher own income passes through an increase in country's mean income (that is, depends on country's growth rate).

The situation, however, is different for the people placed in the bottom or top of income distribution. The former obviously benefit from more equal and the latter from more unequal distributions. We have seen that for the poorest ventile, moving from a distribution or a country with an "average" distribution (that is, with a share of the poor equal to world average) to a more equal distribution (by 1 standard deviation) results in a substantial real income gain (42 percent). Similarly, for the rich, moving from an "average" distribution to a more unequal distribution produces large income gains: for the top ventile, the gain is 30 percent, but when we disaggregate the top ventile into five top percentiles (labeled 96 to top 1% in Figure 1), we can see that the gains steadily rise: for the 96th percentile, the overall

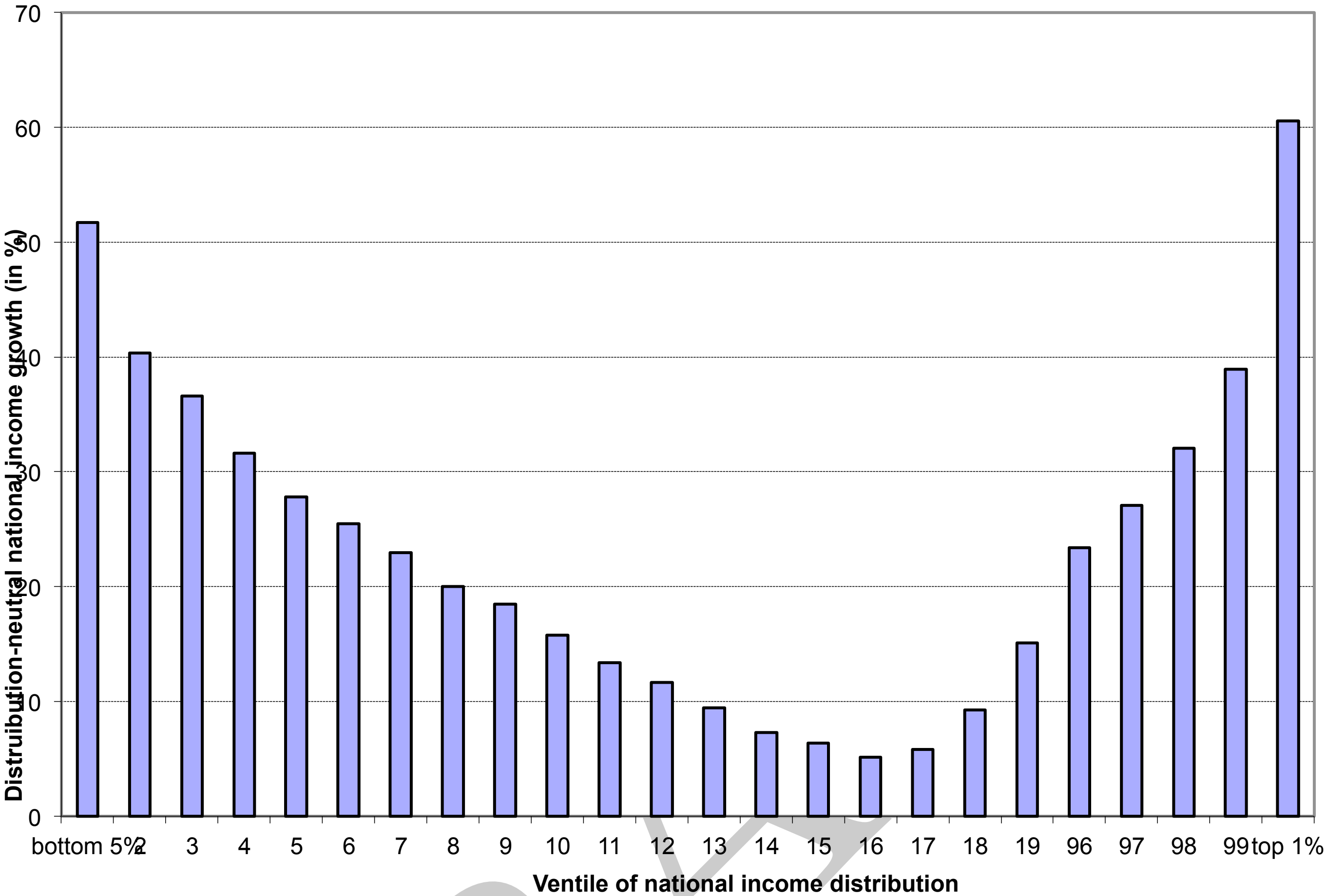

Fig. 6. caption

income gain from a favorable distributional change is 17 percent; for the top 1%, the income gain is almost 45 percent. Thus, for the top 1% to forgo the favorable ("conceivable") distributional change would require a distribution-neutral income growth of almost 45%. The fortunes of both the very poor and the very rich will depend much less on what happens to the mean income and much more on what happens to their country's distribution.

There is a further issue. We have so far defined a "conceivable" distributional change to involve the gain of 1 standard deviation compared to what is the world average. However, the distribution of income shares is not normal. Poor ventiles have income shares that are skewed to the left with a long left tail. This implies that there are quite a few counties with extremely low income shares for the poor ventiles (see Figure 2; left upper panel). Exactly the opposite is the case for the top ventiles. The distribution of income shares of the twentieth ventile is strongly skewed to the right: there are many countries where the richest 5% of population have income shares significantly above the worldwide mean income share of the top ventile. As Figure 2 (right bottom panel) shows, there are countries where the top income share attains 40%, which is twice as much as the worldwide mean share of the top ventile or almost 4 standard deviations above the mean.

Thus, the assumption on which we based our trade-off between income and inequality, namely the advantage of a "conceivable" distributional change of 1 standard deviation, means for the poor ventiles that they move almost to the extreme of what exists in the world, while for the rich ventiles it leaves them with still (empirically) significant possibility to increase own share. Vertical lines in Figure 2 are, for each ventile, drawn at the position equal to the mean (solid line) and "mean + 1 standard deviation" (broken line). For the ventiles 5, 10 and 15, the assumption of "conceivable" distributional change brings them to about the 90th percentile of what is observable in real world. But for the top ventile or top percentile (not shown here), the assumption of "mean + 1 standard deviation" brings it only to about the 75th-80th percentile of what is observable in the world. By pushing for a further increase in inequality, the rich are not pushing against the wall: they can be seen as simply pushing to have a distribution that already exists in other (more unequal) countries.

This then means that if we consider another trade-off such that the favorable distributional change for each ventile means that it reaches an income share equal to the 90th percentile of what is observable, the gains for the rich take an even more extreme form. This is shown in Figure 3. Now the top 1% will require an overall distribution-neutral income growth of 61 percent in order to forgo the possibility of a favorable distributional change. To give an idea of how important is this distributional change: it is equal to the movement from Turkey (where the top 1% receives 7.1 percent of national income) to Mexico (where the top 1% receives 11.4 percent). For other rich ventiles, the distribution-neutral growth needed to forgo the favorable distributional change will also be high, ranging between 23 and 39 percent. But for the middle classes, the equivalent distribution-neutral growth will be around 10 percent.

Finally, it is well known that household surveys tend to underestimate incomes of the top 1%. This is due to rich people's underestimation of capital incomes, non-participation in surveys (see Korinek Mistiaen and Ravallion 2006 [8]), but also to top coding of incomes done by statistical offices. Top coding is a practice introduced by US Census Bureau to set ceilings to various types of incomes, and thus to total income, in order to avoid sudden fluctuations in income shares and measures of inequality as well to preserve confidentiality of information. In Figure 4, we compare micro data from US and Germany; the former applies an aggressive top-coding, the latter does not. It is notable that the very top of US income distribution, around the top 0.1%, drops precipitously, not displaying the long right-end tail that we normally associate with income distributions. Such a long tail however is present in the case of German data. The Pareto line fitted on the top 5% makes us expect to see a much fatter income tail in the US than what we actually see in the data. The contrast is even more striking because US income distribution is significantly more unequal than German: there are many more extremely rich people in the US than in Germany. For example, Forbes 2013 wealth list gives 442 US billionaires and only 58 from Germany; billionaires' wealth/GDP ratios are respectively 12.4 and 8.3 percent. Clearly, something is wrong: either there is a much greater underreporting among the wealthy in the United States, or (more likely) the sharp fall-off is due to the Census Bureau's top-coding.

But whatever is the case, it means that the "true" income shares of the top 1% are higher than recorded, the right-skewness of the top 1% share worldwide probably greater, and thus the equivalent distribution-neutral income growth even higher. It would not be surprising then to posit that empirically the top 1%, interested in maximizing own income, would, in a country where its share is at the world average, have a following choice: (1) use political power to further increase its share to the level of (say) Mexico, or (2) hope or wait until economy's GDP per capita almost doubles in real terms (with no distributional change).

REFERENCES

[1] C. Freeland, *Plutocrats: The New Golden Age*. Penguin, 2012.

[2] N. N. Taleb, "Silent risk: Lectures on probability, vol 1," *Available at SSRN 2392310*, 2015.

[3] H. Theil, "World income inequality and its components," *Economics Letters*, vol. 2, no. 1, pp. 99–102, 1979.

[4] F. A. Cowell, "Measurement of inequality," *Handbook of income distribution*, vol. 1, pp. 87–166, 2000.

[5] B. Milanovic, "Global inequality of opportunity: how much of our income is determined by where we live," *Review of Economics and Statistics*, 2015.

[6] J. G. Palma, "Homogeneous middles vs. heterogeneous tails, and the end of the 'inverted-u': It's all about the share of the rich," *Development and Change*, vol. 42, no. 1, pp. 87–153, 2011.

[7] "Where in the world are you? assessing the importance of circumstance and effort in a world of different mean country incomes and (almost) no migration," *World Bank Working Paper No. 4493*, 2008.

[8] A. Korinek, J. A. Mistiaen, and M. Ravallion, "Survey nonresponse and the distribution of income," *The Journal of Economic Inequality*, vol. 4, no. 1, pp. 33–55, 2006.

APPENDIX A: DERIVING THE GINI FOR A LOGNORMAL DISTRIBUTION

As we mentioned earlier we could find no derivation of the Gini for a lognormal, only the result, with some amount of circularity in the citations. So we derived the result in Equation 3, namely that

$$g^l = 2\Psi\left(\frac{\sigma}{2}\right) - 1$$

for safety.

Proof. Where g^l is the Gini coefficient and X and X' are independent (etc., etc.) with mean μ:

$$g^l = \frac{1}{2}\frac{\mathbb{E}\left(|X - X'|\right)}{\mu}. \qquad (10)$$

In other words the Gini is the mean expected deviation between any two random variables scaled by the mean. If we know the distribution, then Equation 10 is rather simple. In the event of known cumulative distribution function Φ, consider that $|X-X'|= X+X'-2\min(X, X')$. Hence the expectation becomes:

$$\mathbb{E}\left(|X - X'|\right) = 2\left(\mu - \mathbb{E}(X, X')^-\right)$$

We have the joint cumulative

$$F\left((x,x')^-\right) = 1 - \mathbb{P}(X > x)\mathbb{P}(X' > x)$$

hence:

$$g^l = 1 - \frac{1}{\mu}\int_0^\infty \left(1 - \Phi(x)\right)^2 \, \mathrm{d}x \tag{11}$$

Let X follow a lognormal distribution $Ln\left(\log(\mu) - \frac{\sigma^2}{2}, \sigma\right)$ with PDF

$$\phi(x) = \frac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{\left(-\log(\mu)+\frac{\sigma^2}{2}+\log(x)\right)^2}{2\sigma^2}},$$

CDF

$$\Phi(x) = \frac{1}{2}\text{erfc}\left(\frac{\log(\mu) - \frac{\sigma^2}{2} - \log(x)}{\sqrt{2}\sigma}\right),$$

where erfc is the complementary error function, and mean $\mathbb{E}(X) = \mu$.

Equation 11 can be derived by parts:

$$g^l = \Big[x(\Phi(x)-1)\Phi(x)\Big]_0^\infty - \frac{1}{\mu}\int_0^\infty x(1-2\Phi(x))\Phi'(x)\, \mathrm{d}x \tag{12}$$

hence:

$$G = \frac{1}{\sqrt{2\pi}\mu\sigma} \int_0^\infty e^{-\frac{\left(-\log(\mu)+\frac{\sigma^2}{2}+\log(x)\right)^2}{2\sigma^2}} \text{erf}\left(\frac{\log(\mu) - \frac{\sigma^2}{2} - \log(x)}{\sqrt{2}\sigma}\right) \mathrm{d}x. \tag{13}$$

Substituting $u = \frac{\log(\mu) - \frac{\sigma^2}{2} - \log(x)}{\sqrt{2}\sigma}$ and changing the band of integration:

$$g^l = \int_{-\infty}^\infty \frac{\text{erf}(u) e^{-\frac{\sigma^2}{2} - u^2 - \sqrt{2}\sigma u}}{\sqrt{\pi}} \, \mathrm{d}u$$

Integrating by parts:

$$g^l = \frac{1}{\sqrt{\pi}}\int_{-\infty}^\infty e^{-u^2}\text{erf}\left(\frac{\sigma}{\sqrt{2}} + u\right) \mathrm{d}u.$$

Since

$$\frac{\partial g^l}{\partial \sigma} = \frac{\partial}{\partial \sigma}\left(\int_{-\infty}^\infty \frac{e^{-u^2}\text{erf}\left(\frac{\sigma}{\sqrt{2}} + z\right)}{\sqrt{\pi}} dz\right) = \frac{e^{-\frac{\sigma^2}{4}}}{\sqrt{\pi}},$$

we finally get

$$g^l = \int_0^\sigma \frac{e^{-\frac{t^2}{4}}}{\sqrt{\pi}} \mathrm{d}t = \text{erf}\left(\frac{\sigma}{2}\right). \tag{14}$$

We note that $\text{erf}\left(\frac{\sigma}{2}\right) = 2\Psi\left(\frac{\sigma}{2}\right) - 1$. □

———