Dear Professor Piketty,

When writing an article on the distribution of wealth in the UK last week, I noticed a discrepancy between the contemporary concentration of wealth described in *Capital in the 21st Century* and that reported in the official UK statistics.

Having further investigated the sources and methods you employed to produce the charts 10.1 to 10.6 in chapter 10 of *Capital in the 21st Century*, the Financial Times has found some apparent problems with your methods and calculations. We propose to publish our findings on Friday.

This email is written as a courtesy so that you are not surprised by the articles when they appear. It also gives you an opportunity to reply before publication if you wish to do so.

We will be able to reflect any response you provide regarding our concerns in our reporting so long as we receive it by 15.00 BST tomorrow, Friday 23rd May 2014.

The nature of our proposed reporting is outlined below. It highlights apparent biases in your published results, which have a material effect on the trends in wealth inequalities you outlined, particularly in recent decades.

Yours faithfully

Chris Giles Economics Editor, FT cc. Rebekah White, Harvard University Press Susan Donnelly, Harvard University Press

Problems with the wealth data in Capital in the 21st Century

Fat finger problems

There appear to be "fat finger" problems in your book in which the published results are different from the original source material. An example is the data for the wealth held by the richest 10 per cent and 1 per cent of people in 1920 Sweden. The technical annex to *Capital in the 21st Century* says the source is Waldenstrom (2009), and the relevant extract is copied below.

| | Net worth (net marketable wealth) | | | | | | | | | |
|------|-----------------------------------|-------|-------|--------|-----------------------------|-------|-------|-------|--------|----------|
| | Wealth tax data, market values | | | | Estate tax data, tax values | | | | | |
| v | P90- | P95- | P99- | P99.9- | P99.99- | P90- | P95- | P99- | P99.9- | P99.99- |
| Year | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 25/10/05 |
| 1873 | | | | | | 88.34 | 81.19 | 60.46 | 35.60 | 5.75 |
| 1874 | | | | | | 85.82 | 77.35 | 52.82 | 24.23 | 4.41 |
| 1875 | | | | | | 85.83 | 77.49 | 54.18 | 24.46 | 6.79 |
| 1876 | | | | | | 86.14 | 77.66 | 55.69 | 23.15 | 7.40 |
| 1877 | | | | | | 85.99 | 77.39 | 54.07 | 23.55 | 5.60 |
| 1906 | | | | | | 87.38 | 78.37 | 57.75 | 26.14 | 4.36 |
| 1907 | | | | | | 88.32 | 79.88 | 61.29 | 31.70 | 11.19 |
| 1908 | 86.04 | 76.17 | 53.79 | 28.13 | 13.64 | 88.15 | 79.44 | 61.10 | 27.01 | 3.57 |
| 1920 | 91.69 | 79.25 | 51.51 | 25.37 | 9.60 | | | | | |
| 1930 | 89 49 | 77 35 | 50.02 | 22 35 | 9 23 | | | | | |

It is clear from the above table that the numbers should be 91.69 and 51.51 for (p90-100) and (p99-100) respectively, but as the extract from your spreadsheet below shows, the equivalent numbers used in Capital in the 21^{st} Century are 87.7 and 53.8. Since the number for the top 1 per cent is the same as that for 1908 in Waldenstrom, this appears one of seemingly many fat finger problems in the book.

| (Ohara of the tan | Sweden | | | | | |
|--|---------|--------|-------------|--|--|--|
| (Share of the top x% wealthiest in the total wealth) | Top 10% | Top 1% | Top 0,1% | | | |
| 1810 | 83.9% | 55.9% | | | | |
| 1820 | | | | | | |
| 1830 | | | | | | |
| 1840 | | | | | | |
| 1850 | | | | | | |
| 1860 | | | | | | |
| 1870 | 87.2% | 57.3% | 29.6% | | | |
| 1880 | | | | | | |
| 1890 | | | | | | |
| 1900 | | | | | | |
| 1910 | 88.2% | 61.1% | 29.4% | | | |
| 1920 | 87.7% | 53.8% | | | | |
| 1930 | 83.6% | 42.8% | | | | |
| 1940 | 83.2% | 37.7% | 17.7% | | | |
| 1950 | 77.3% | 32.8% | 9.7% | | | |

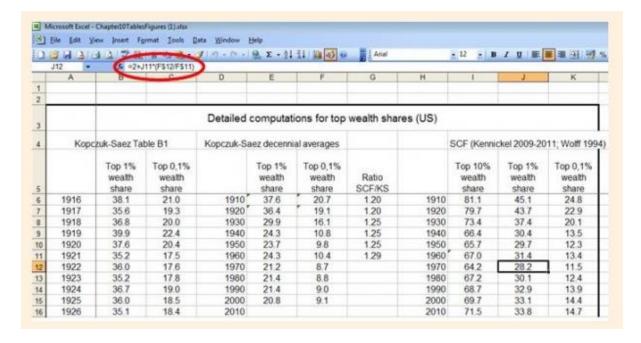
Adjusted data with little explanation

There are numerous examples of data adjusted without explanation in the footnotes or technical annex.

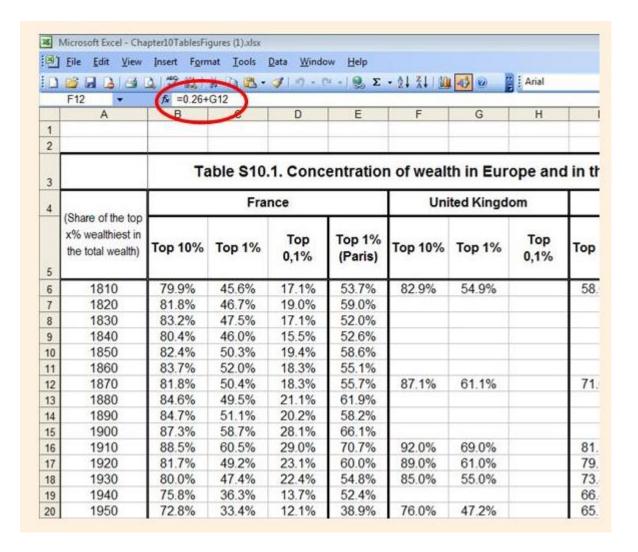
A first example relates to the scale factors used to translate data from estate taxes to the living population wealth distribution in France. These are constant for every year between 1810 and 1960, except for 1910, when the scale factor is increased. In addition, the original source data from the 2006 AER says (footnote 32) that the data is not sufficiently complete to generate a time series for the twentieth century, yet this is done.

| | | | V 15 19 - 0 | 119-0 | - 0, Σ - 21 | 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 | Arial | | - 12 | |
|----|---|---------------------------|-----------------------------|---|----------------------------|---------------------------------------|-----------------------------|--------------------------------------|------|--|
| | A | C =1.0 | 5*C9/100 | E | F | G | н | 1 | | |
| 1 | | | | | | | | | | |
| 2 | | | | | | | | | | |
| 4 | Detailed computations for top wealth shares (Fran | | | | | | | | | |
| |)6 Table 4 (top shidecedents, France) | | | Final estimates used here (top shares among living, France) | | | | | | |
| | (1.5) | Top 1% wealth share | Top 0,1% wealth share | | Top 10% wealth share | Top 1% wealth share | Top 0,1% wealth share | Top 1% wealth share (Paris) | | |
| 6 | 1807 | 43.4 | 16.3 | 1810 | 79.9% | 45.6% | 17.1% | 53.7% | | |
| 7 | 1817 | 44.5 | 18.1 | 1820 | 81.8% | 46.7% | 19.0% | 59.0% | | |
| 8 | 1827 | 45.2 | 16.3 | 1830 | 83.2% | 47.5% | 17.1% | 52.0% | | |
| 9 | 1837 | 43.8 | 14.7 | 1840 | 80.4% | 46.0% | 15.5% | 52.6% | | |
| 10 | 1847 | 47.9 | 18.4 | 1850 | 82.4% | 50.3% | 19.4% | 58.6% | | |
| 11 | 1857 | 49.5 | 17.4 | 1860 | 83.7% | 52.0% | 18.3% | 55.1% | | |
| 12 | 1867 | 48.0 | 17.4 | 1870 | 81.8% | 50.4% | 18.3% | 55.7% | | |
| 13 | 1877 | 47.1 | 20.1 | 1880 | 84.6% | 49.5% | 21.1% | 61.9% | | |
| 14 | 1887 | 48.7 | 19.2 | 1890 | 84.7% | 51.1% | 20.2% | 58.2% | | |
| 15 | 1902 | 51.6 | 23.1 | 1900 | 87.3% | 58.7% | 28.1% | 66.1% | | |
| 16 | 1913 | 54.9 | 26.0 | 1910 | 88.5% | 60.5% | 29.0% | 70.7% | | |
| 17 | 1929 | 50.2 | 24.7 | 1920 | 81.7% | 49.2% | 23.1% | 60.0% | | |
| 18 | 1938 | 42.0 | 19.9 | 1930 | 80.0% | 47.4% | 22.4% | 54.8% | | |
| 19 | 1947 | 29.9 | 11.0 | 1940 | 75.8% | 36.3% | 13.7% | 52.4% | | |
| 20 | 1956 | 30.4 | 11.0 | 1950 | 72.8% | 33.4% | 12.1% | 38.9% | | |
| | 4004 | 04.0 | 00 | 4000 | 00.00/ | 24 000 | 44 FRE | 00.007 | | |

A second example comes from the US data. The wealth share of the top 1 per cent is increased by 2 percentage points for 1970 as is clear from the screen grab below for 1970. The 1970 datapoint is also interesting for the top 1 per cent wealth share because it is based on movements in the KS series for the top 0.1 per cent share. This assumption is not explained and may also be problematic.

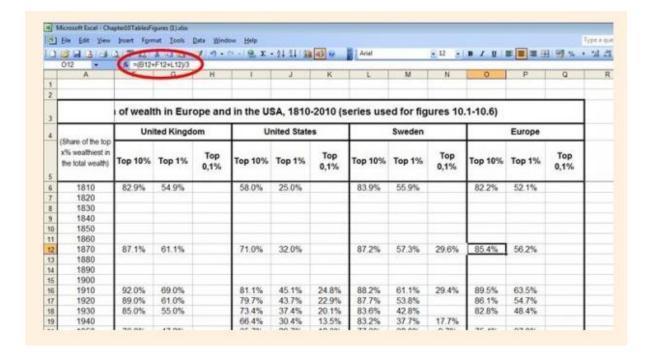


Another example comes from the British data. In cell F12, you say the wealth share of the top 10 per cent in 1870 is equivalent to the top 1 per cent plus 26 percentage points. The Lindert 2006 JPE paper gives estimates for this figure so there does not appear to be a need to add a constant, nor for using a different scale factor in 1810 of 28 percentage points. There is no discussion of the need to use these scale factors nor to change them for different years. Lindert estimates exist, which are accepted in the book for the top 1 per cent share, but not the top 10 per cent share.



Averaging problems

When averaging the data between Britain, France and Sweden to arrive at combined European estimates, the book uses a simple average, giving roughly seven times more weight to a person in Sweden than Britain or France. This is shown in the screen grab below.



Data generation issues

Some of the data in the spreadsheets is difficult to reconcile with the sourcing given.

One example is that there is no historical source data for the top 10 per cent wealth share in the US between 1910 and 1950. You assume the top 10 per cent wealth share is your estimate for the top 1 per cent share plus 36 percentage points.

The data for the top 1 per cent share of the UK data in 2010 similarly has no source and no possible source given the citations in the technical annex.

Problems with the choice of points of comparison

There is no doubt that the source data is sketchy. It is difficult to find data that relates to the start of each decade that the figures 10.1 to 10.6 demand, so it is natural, for example, to choose 1908 as a reasonable data point for 1910 on the graph.

It is harder to explain why the chart uses data from 1935 Sweden for the 1930 datapoint, especially when 1930 data exists in the source material, or why the UK source data for 1938 should be used for 1930 rather than 1940.

Problems with definitions

In the source notes to the spreadsheets, you state that the wealth data for the countries you compare is consistent data from the same methods applied to different countries.

"Note: as explained in the text, these are for all countries estimates of inequality of net worth between living adults (using mortality multiplier methods)."

This does not appear to be the case. The US data does not use estate tax records for 1960 and after 1970, using estimates coming from a completely different type of source, the cross section Survey of Consumer Finances.

This might be well justified, but for the UK, there has not been a similar decision to switch this century to the new Office for National Statistics wealth and assets survey, which could have provided more robust estimates for 2000 and 2010. Instead, the book bases the latest UK numbers on the UK's HMRC distribution of personal wealth statistics. HMRC is clear in its footnotes that this data cannot be used for estimating total distribution of wealth. The data notes state clearly that coverage problems in the HMRC data "limits the use of the data for assessing the wealth of the whole population, particularly for the smaller estates which are likely to be under-represented in the data. It means it is not a suitable data source for estimating total wealth in the UK, or wealth inequality across the whole of the wealth population; the Wealth and Asset survey is more suitable for those purposes".

Cherry-picking data sources

There is little consistency in the way that Capital in the 21st Century combines different data sources, which are often measured in very different ways and have different definitions of wealth.

Sometimes, as in the US, your numbers tend to favour cross-sectional surveys of living households rather than estate tax records. For the UK, the book appears to favour data that come from a cross sectional survey of living households.

In both cases, your choice of which type of data to use has had the effect of showing wealth inequality rising, rather than staying constant (US) or falling (UK).