Triangulation in Economic and Social Measurement

The 2018 Nancy and Richard Ruggles Memorial Lecture

August 21, 2018

Tim Smeeding
Lee Rainwater Distinguished Professor of Public Affairs and Economics
Hello —and thanks

• To the organization where I have had the privilege of being involved since 1981: Council in 1996-2001; Chair of Council 2002-2004; Board of Editors, RoIW 1987-1992 and 1997-2018 (26 years!)

• First “income distribution” guy amongst the national accountants but learned from Atkinson and others (Anne Harrison, Alice Nakamura) about the importance of macro-data with micro-data

• Many to thank: peers, older RIW leaders, and wise men (one exemplar Peter Hill and OECD story):


--- Current co-authors whose work has influenced this presentation --David Johnson, Janet Gornick, Jeffery Thompson, Jonathan Fisher,

Themes

• Better evidence (essential for better policy) comes from both data and conceptual ‘triangulations’

• Three sources together better than any one alone (lots of exemplars, choose 3 here)

• Focus on distributional outcomes (‘heterogeneity’) but broad applications to productivity, consumption, macros measurement topics (Torben A., this am!!)

• Sub title—”the growing importance of capital income and wealth”

• Precursor? Richard and Nancy Ruggles “THE INTEGRATION OF MACRO AND MICRO DATA FOR THE HOUSEHOLD SECTOR”, RIW, 1986
Outline

1. Triangulation of data in measuring the distribution of economic well-being
2. Triangulation of concepts in measuring household inequality: $Y, C, W$ on uses side; and $K, L, NT$ on sources side
3. Triangulation and the family income “package” – interfamily transfers as key
4. Summary and conclusions
1. Triangulation in measuring the distribution of economic well-being

Three main sources of data:

1a. **Surveys** (cross-sections and panels)

1b. **Macro National Accounts** (economic and financial) and ‘international’

1c. **Administrative data** – public sector mainly, but also “big online data” for some purposes (Google, Forbes lists)
Emphases, approach and major points to be made

- Cross walking from one measure to the other is the future of economic inequality measurement
- Start at any one point: micros start with surveys or admin data—like tax files; macros take surveys and ‘gross up’ to get to DNAs (OECD) / DINAs (WID)
- Signal and noise are both important—but quality and harmonization are key elements
- Focus on income (flows) and components, but also stocks (wealth in next section of talk)
- Focus on top and bottom ends of distributions
1a. Survey data

- Basic bread and butter of income inequality since 1948 – eg USA CPS for cross sections & later, with panels (links across multiple generations; three or more now in PSID and other long running panels like GSOEP and BHPS)

- Demography, ownership, occupation, family/household structure, income types (especially earnings, home ownership, and social and private transfers—income package), outcomes (health) and investments (education/ human capital), can all be harmonized ex-ante (Canberra, EU-SILC, OECD) or ex-post (LIS, CNEF)

- Weaknesses -- sampling and non-sampling error (attrition in panels); poor on capital income and less so on government benefits; special concern, non response or under sampling of the rich—top 1-2 % in income or wealth
1b. National Accounts / Flow of Funds—aggregate “macro” data

- **Strengths**—covers larger economy and sources not on surveys or administrative data—e.g., corporate retained earnings, business wealth or other non-taxable flows.

- **Use** to “gross up” administrative and/or or survey amounts to reach domestic totals—methods now simplistic in the absence of other information (‘proportional imputation’ -taxable dividends are grossed up to match SNA totals for dividends and retained earnings); or more nuanced with other information.

- **Weaknesses**—need to separate household sector; misses assets owned in other nations (but see Zucman); misses many income or investment transfers across households, misses realized capital gains.
1c. Administrative Record Data

- **Strengths**—‘registers’- often more accurate sources of data for some components, especially capital income that is taxable (vs assets that do generate taxable income flows) and also realized capital gains; earnings reported to social security; income transfer support systems; institutional experiences (schooling; incarceration; immigration)

- **Weaknesses**—*eg tax files --units are tax filers, not households; cover only those who must pay tax and taxable income* (so miss poor and many income supports in most nations); *earnings files, miss unreported income* (casual labor, off the book earnings); *non-public transfer are missed and these can be very important* (in-vivos transfers across generations); *tax files miss nontaxable income flows like retained earnings and unrealized capital gains; mysterious ownership rules for tax minimization (US “chapter S” corporations)*; follow across generations?
Where should you start to depends on the tail you are interested in

- **Poverty and lower end of distribution**: start with surveys and build up; ‘direct’ matches with administrative data best for missing transfers— but not always available

- **Inequality and upper end of distribution**: start with fiscal (tax) administrative data, add survey information

- **Then both end with SNA help on missing income at the very end**
Figure: Triangulation of Data Sources to Improves Inequality and Poverty Measurement

Surveys

Inequality

Poverty

Administrative Records
Fiscal (Inequality)
Programmatic (Poverty)

Aggregate Accounts
Flow of Funds, SNA (Inequality)
Earnings, Benefits, Program Outlays (Poverty)
Poverty measurement: survey to administrative data to SNA

• Start with sample survey and go to administrative data for direct match
• Challenges and quality issues abound (below)
• Very important if links can be established over longer periods, for panel datasets (“generational accounting “)
• SNA shows missing earnings, other items
Inequality measurement: fiscal (tax)records to surveys to SNA

- Start with fiscal/tax records (for 100 years!)
- Get “demography” from surveys – or try to match in same, add in non-filers
- Compare totals to SNA/Flow of Funds
- Mix and match (signal and noise)
- Do again—using better matching techniques
- Again challenges and quality issues abound
Errors and data quality

- **Survey errors** (sampling and non-sampling error; unit and item specific non-response)
- **Administrative data errors** (consent error; matching error; other administrative data errors)
- **SNA errors** (limiting universe to households; ability to adjust administrative and population totals)
- **All three** — methods to match across types

*BUT what do you learn?*  **LOTS of very strong signal despite the noise** & progress to reduce noise
What are we missing? (USA SNA/ NIPA):
Poverty (missing government transfers) vs Inequality (missing property, business income)

Figure 7
Ratio of NIPA Income to CPS Aggregate Income
Adjustment Factors $a_i$

Notes: See text for description of NIPA adjustment.
Source: Own calculations from BEA and public use CPS.
Lets look at some distributions–

- **Income Inequality** (before tax and transfers) fiscal data, combined data – strong signal
- **World Poverty** just surveys—strong signal
- **World Income Inequality**— how big is the elephants nose?? (Pinocchio cometh)
- **Role of SNA (and FoF)** —crucial here in getting the aggregates right
USA-- top 10 percent income share fiscal (tax) records only

Note 2015 exceeds 1927 for all time high

(Source http://elsa.berkeley.edu/~saez/TabFig2015prel.xls)
Top 10% income shares across the world, 1980–2016

Source: WID [https://wid.world/](https://wid.world/) can find for data series and notes.
Bottom 50% income shares across the world, 1980–2016

Source: WID, see https://wid.world/ for data series and notes.
EXTREME POVERTY
($2 per person per day Surveys only)

Source --Max Roser
https://ourworldindata.org/extreme-poverty
No need to know lots more—except China's war on poverty worked.

World population living in extreme poverty, 1820 to 2015

Extreme poverty is defined as living on less than 1.90 "international-$" per day.
International-$ are adjusted for price differences between countries and for price changes over time (inflation).

Data: Bourguignon and Morrisson (2002) – Inequality among World Citizens and from 1981 onwards PovcalNet
The visualization is available at OurWorldInData.org, where you find more visualizations and research on human development.
Licensed under CC-BY-SA by the author Max Roser.

Robert M. La Follette School of Public Affairs
UNIVERSITY OF WISCONSIN–MADISON

700 million poor in 2015; SDG target 300m in 2030
And if we solve a Africa– boom

Projection of the number of people in extreme poverty until 2030

Extreme poverty is defined as living with less than $1.90/day. This is measured by adjusting for price changes over time and for price differences between countries (PPP adjustment).

The projection is based on the SSP2 scenario used in the IPCC reports which assumes the continuation of current global socioeconomic trends at the global level.


The data visualization is available at OurWorldInData.org. There you find research and more visualizations on this topic.
Combinations- who gets the growth?
“elephants noses”--“Original paste up*” (red) and Surveys** (blue)

* Milanovic, 2015; Milanovic and Lakner, 2014
** Brookings, 2018
Now Pinocchio strikes: from Surveys* (blue) to Tax Files plus SNA** (orange)

Figure 7. Growth Incidence Curve Using Survey Data versus Distributed National Accounts Data

- Updated Data with Consistent Sample, 2011 PPP, Filled In If Any Survey
- WID Distributional National Accounts Data

* Brookings, 2018
** WID 2018
How did that happen?

- Crucial role of SNAs, especially for income from property, retained earnings, businesses, and so on
- The crucial role of DINA and DNA projects
- The need for better identification of capital (property, business, closely held corporation) incomes, and other sources like pension funds, as demonstrated below
Reconciling national labor income and labor income reported on tax returns

From taxable to total labor income

- Tax evasion & other
- Employer fringe benefits & payroll taxes
- Wages and self-employment income on tax returns

Source: Appendix Table I-S.A8b

Source: Alvaredo, et al 2018

https://eml.berkeley.edu/~saez/ACPSZ2018AER_Slides.pdf
A growing fraction of labor income is missed by tax data

Source: Alvaredo, et al 2018

https://eml.berkeley.edu/~saez/ACPSZ2018AER_Slides.pdf
Reconciling national capital income and capital income reported on tax returns

From taxable to total capital income

- Non-filers & other
- Retained earnings
- Corporate income tax
- Income paid to pensions & insurance
- Imputed rents + property tax
- Dividends, interest, rents & profits reported on tax returns

Source: Appendix Table I-S.A8

Source: Alvaredo, et al 2018

https://eml.berkeley.edu/~saez/ACPSZ2018AER_Slides.pdf
Most capital income is missed by tax data

From taxable to total capital income

Source: Alvaredo, et al 2018
https://eml.berkeley.edu/~saez/ACPSZ2018AER_Slides.pdf
Who is where in world income distribution
Some examples of what we get from good combinations --

- Distributions- national and world ( & elephants nose)
- Progress and regress in reducing poverty and lower end inequality ( backbone of “shared prosperity” & “inclusive growth” efforts)
- Better estimates of mobility within and across generations
- Better measures of effects of public transfers - but missing strategic and irregular transfers, eg remittances and cross-generational transfers
- The ‘answers’ all depend on where you focus the question —lower end-transfers vs. upper end -capital income (who owns the robots ?)
Policy lesson from country data - ever growing top 1% share is not inevitable

Share of Total Income going to the Top 1% since 1900

The evolution of inequality in English speaking countries followed a U-shape.

The evolution of inequality in continental Europe and Japan followed an L-shape.

Data source: World Wealth and Income Database (2018). This is income before taxes and transfers. This data visualisation is available at OurWorldinData.org. There you find the raw data and more visualisations on inequality and how the world is changing. Licensed under CC-BY-SA by the author Max Roser.

Figure 27: Rising inequality and income growth: China vs others

Average annual growth rate of real per adult pre-tax national income, 1978-2015

Source, WTID, http://wid.world/
Frontiers to Explore

• **Better units**—who shares and how?—bridging and matching to other data (e.g., Censuses)

• **Matching more data**—linking surveys and administrative records, and the “new” science of adjustments for missing matches

• **Better distributors** than proportionate income inflation to get to SNA totals

• **Income hoarding and hiding** (Zucman and SNA/FoF—“FATS”)—uncover “hidden incomes”
2. Triangulation in measuring sources of inequality, Y,C,W

- “the most pertinent measures of the distribution of material living standards are probably based on jointly considering the income, consumption, and wealth position of households or individuals.”

Commission on the Measurement of Economic Performance and Social Progress (Fitoussi, Stiglitz et al., 2009):

- Income(Y), consumption(C), and wealth (W, NW) we need all three together for same units

- Start with aggregate accounting this time
Flows and stocks: Income (Y); Consumption (C); Net Worth (NW)

• Haig and Simons definition, income (Y) is equal to consumption (C) plus the change in net worth ($\Delta NW$) realized over an income accounting period.

• So defined, H-S income is a measure of potential consumption: amount one could consume or transfer without changing total net worth (one’s stock of assets or debts)

• Thus according to a “uses “ of income definition:

$$Y = C + \Delta NW$$
Sources of income

• Functional “sources” side of income \( (Y) \), arrive at the same measure adding together income from earnings \( (E) \), income from capital \( (K_I, \text{ including capital gains plus other income from wealth}) \), plus net transfers \( (N_T, \text{ which includes those received minus those paid out}) \)

\[
Y = E + K_I + N_T
\]

• If we ignore \( N_T \) (for now—but not for long!), divide self-employment income, into income from labor and capital, we are left with the macroeconomists’ functional distribution of income.

• So what is here for distributional analyses from the sources side?
Sources side: $Y = E + KI (+ NT)$

- Factor Shares—$E$ (labor share of national income) falling in USA: more than 50% in 1970’s, now 42%
Falling labor share around the world

source: IMF World Economic Outlook, 2017

Labor is losing out
The share of national income paid to workers has been declining in many countries.
(evolution of the labor share of income, percent)

Source: IMF, World Economic Outlook, April 2017.
Why? Capital, up

- Technological change, global trade--- and policy --
  - 'Regulatory' policy: rising concentration of industry, less competition & more profit
  - Pro-capital tax policy (not all, but USA!)
  - 'Rent capture': sheltered markets, limited enforcement, protected market niches, and political power
Why? Labor, down

• Rising monopsony power and policy, global competition from cheap labor, insecurity of work—*not* just decline of unions
  - “non–compete clauses”;
  - workplace inflexibility;
  - spatial immobility of workers;
  - rise of “gig” economy
  - declining real minimum wages
Tax policy in USA!

The Trump Economy
Corporate Profits vs. Wages

Percent change from one year ago

-2.5  0  2.5  5.0  7.5  10.0  12.5  15.0

Profits
GOP tax bill passed

Wages

2017 Q1 2017 Q2 2017 Q3 2017 Q4 2018 Q1 2018 Q2

Proﬁts: Real corporate proﬁts after tax with inventory valuation adjustment and capital consumption adjustment

Wages: Real average hourly earnings of production and nonsupervisory employees: total private


Back to uses side: \[ Y = C + \Delta NW \]

- Stock (\( W \)) and flow (above) issues abound.
- The hardest thing to measure is the real change in net worth (\( \Delta NW \)) as much of it is not realized or distributed and hence not captured in surveys or registers — but it is behaviorally VERY important.
- The thing we can measure much better is the stock \(-W\) (NW) alone using proper samples (SCF), and administrative data (tax, flow of funds, SNA).
Why care about \( \Delta NW \)?

- Changes in housing values vary enormously – spatially and cyclically.
- Changes in financial wealth also have cyclical and idiosyncratic changes.
- But most stocks and financial wealth including defined contribution pension plans are owned by the top decile (about 75% in USA) in the period when capital is winning on the sources side.
- E.g., 1998 and 2017, two “very good years” for top decile wealth and pension holders in USA.
Wealth as key

• The stock, NW can replace the flows, Y and C, multiple times over
• from OECD DNA, to WidWorld DINA and the better surveys distribution of wealth
• from panel data -- dynasty and mobility across three generations or more now
• Key: role of intergenerational transfers in maintaining off-spring economic position
The distribution of family wealth: USA 1963-2016


$12 million

$9

$6

$3

$0

10th: $1.457,201
99th percentile: $1.457,201
95th: $409,182
90th: $238,860
10th: $724
50th: $82,746
90th: $520,133
10th: $950
50th: $97,300
90th: $1,186,570
99th percentile: $10,400,000
95th: $2,387,250
90th: $1,186,570

Note: 2016 dollars.

Source: SCF at http://apps.urban.org/features/wealth-inequality-charts/
### Table 1: 2016 USA only-SCF NW/Y/C Combined File Descriptives

<table>
<thead>
<tr>
<th></th>
<th>NW (000)</th>
<th>Y (000)</th>
<th>C (000)</th>
<th>NW/Y</th>
<th>NW/C</th>
</tr>
</thead>
<tbody>
<tr>
<td>P95</td>
<td>$2.400</td>
<td>.197</td>
<td>.135</td>
<td>12.2</td>
<td>17.7</td>
</tr>
<tr>
<td>P50</td>
<td>$ .097</td>
<td>.047</td>
<td>.044</td>
<td>2.1</td>
<td>2.2</td>
</tr>
</tbody>
</table>

In fact in 2016 in USA --

P95 NW could finance 51 years of P50 (median) income;
P50 NW could finance .5 years of income at P95

Note:
NW – From SCF for March 2016
Y – Disposable income from SCF for calendar year 2015
C – Total consumption from imputed/enhanced SCF totals for calendar year 2015

Source: Authors’ calculations from 2016 SCF and related work (Fisher et al., 2018)
C,Y & W together for same families

• Question:
What fraction of all households that were in the top 5% of the income (Y) distribution, were also in the top 5% of the consumption (C) distribution and the top 5% of the wealth (NW) distribution year by year?

• Answers:
1989 -- 32 %
2007 -- 49 %
2016 -- 44% *

* March 2016-summer 2018, stock markets rose more than 30% in USA, suggesting that the answer in USA is now more than 50%
Why should we care--corrosive effect of $W$ on mobility

- Wealth is passed generation to generation in two forms:
  - **Inheritance** -- only at death of oldest parent, so late in life
  - **In-vivos** -- at key stages in life course, early on through key periods of human and physical capital formation
    - note the “glass floor” at the top: child neighborhood; education; co-sign mortgage; free rent; subsidized internships; and often lifetime job in family firm

- Suggesting another triangle ---
3. Triangulation and the family income package

- Households rely on “packages” of income from three main sources as another key triangle:

1. **own efforts** (and partners) -- the labor market, capital markets, savings (private pensions)

2. ‘**Extended**’ family members --- private transfers from those living in other households

3. the **state/public sector** --- (i.e., transfers in, net of taxes out).
The middle element “extended family” as key measurement issue

• Much has been written about own earnings and even “pre-distribution” (the rules governing the labor and capital markets like minimum wages, work hours rules)

• Even more has been written about “redistribution” (the way taxes and benefits affect households)

• But consider interfamily transfers
  
a. increasing inequality and reducing mobility—by means of one time strategic transfers in-vivos

  b. reducing world poverty—by means of remittances from emigrated family members in search of jobs - key to African poverty reduction?
Increasing inequality and reducing mobility via in-vivos transfers

- In the United States, in the aggregate, regular private cash transfers pale in comparison with large, irregular private inter-vivos “strategic transfers”.
- These transfers are rarely recorded as consumption, or income data, or (except in some cases where ‘donors-only’ are queried) wealth surveys (typically known only to the private money managers).
- On donor side: households in the top wealth quartile of persons 50 or over who made a transfer, averaged gifts of over $40,000 in 2009-10 alone (Banerjee 2015).
- But the survey offers no information on the economic status of recipient children or grandchildren.
b. reducing poverty via remittances

- More work is sorely needed on transnational inter-household transfers; a growing body of work attempts to quantify and locate remittances, most of which flow from richer to poorer countries (World Bank 2017), but data are incomplete.
- Aggregate sources abound (next slide) but---
- Many surveys query whether households received “regular” transfers “in” but fail to specify if those transfers are domestic or transnational and amounts reported are underestimates --eg African receipt high and growing
Remittances

Global remittances a record level $613-billion (2017)
4. Conclusions--substantive matters

• Inequality is larger than we think using any one micro source in most nations
• But patterns differ across nations
• The ‘answers’ you find all depend on where you focus:
  -- transfers and the poor
  -- or inequality, wealth, capital income and transfers ('who owns the robots' ?)
4. Conclusions on measurement innovation

• Triangulation of data, economic resources, and family income packages are all growing in importance

• There is a lot of noise but also a lot of signal and combining macro-data and micro-data (surveys, administrative data) are helpful in each case

• Members of IARIW increasingly need to know, appreciate & understand both perspectives
The end

• Thank you
• Questions please
• Comments to smeeding@wisc.edu
Sources

Additional slides
US top 1% income shares (L, orange, and lots of K) and composition*


Excluding capital gains and negative observations. Income categories are defined below in Section 1.
The World as 100 People over the last two centuries

**Extreme Poverty**
- 6 not living in extreme poverty
- 94 living in extreme poverty
- 90 not living in extreme poverty
- 10 living in extreme poverty

**Basic Education**
- 83 have not attained any education
- 17 have basic education or more
- 14 have not attained any education
- 86 have basic education or more

**Literacy**
- 88 are not able to read
- 12 are able to read
- 15 are not able to read
- 85 are able to read

**Democracy**
- 59 not living in a democracy
- 41 living in a democracy
- 44 not living in a democracy
- 56 living in a democracy

**Vaccination**
- 14 not vaccinated
- 86 vaccinated

**Child Mortality**
- 57 survive the first 5 years of life
- 43 die before they are 5 years old
- 96 survive the first 5 years of life
- 4 die before they are 5 years old

Data sources:
- Vaccination: WHO Global data are available for 1980 to 2015 – the WHO’s vaccination was licensed in 1999
- Literacy: Literacy data for the period 1820 to 1950, UNICEF for 1990 and later.

All these visualizations are from OurWorldInData.org an online publication that presents the empirical evidence on how the world is changing.

Licensed under CC-BY-5A by the author Max Roser.
Total income growth by percentile across all world regions, 1980-2016

Source: WID, see https://wid.world/ website for more details.
Growth in Household Wealth, 1950-2016

Notes: Lines show growth rates for different wealth groups, with blue for the bottom 50 percent, for the middle class (50th percentile to 90th percentile), and orange for the top 10 percent. All time series are indexed to 1 in 1971. Vertical line indicates financial crisis.
Source: Authors’ calculations

Robert M. La Follette
School of Public Affairs
UNIVERSITY OF WISCONSIN–MADISON
Intergenerational transfers are frequent and large and make a difference

Consider the source— but see the numbers too

Fig. 7: Financial assistance to adult children
Parents aged 47–65 who have provided financial support to adult children

- Helped with college loans or tuition
- Allowed to move home rent free
- Helped to buy a car
- Helped with car insurance
- Helped with rent or utilities
- Co-signed a loan or lease
- Helped with medical insurance
- Helped with paying credit card debt
- Helped with house down payment
- Helped with a mortgage payment

Source: Ameriprise Financial
The outcomes are not inevitable: we can do better

Institutions matter: compare China, US, France—

• Invest in human capital, especially for kids (health, education, upward mobility) – how countries treat children is key
• Tax capital income (no K gains roll-over) same as labor income
• More widely shared profits –how owners treat valued workers will be important, esp. if scarce and highly productive
• Mandatory defined contribution pensions managed by third party
• Employer labor partnerships, post secondary education & training (eg German work sharing; Danish and EU ‘ALMPs’)
• Promote shared prosperity and inclusive growth, value firms for more than the bottom line (dignity of work, environment)
• Give labor a voice in political discourse
How to think of Decennial Population Censuses?

- Are they surveys (treat as such here) or administrative data?
- Example every 10 years in USA you are “compelled” to answer the short form and give family demography (relatedness, names, gender, race/ethnicity, co-residence, occupation, other)
- Since 1980s all US born children have unique identifier
- Incredibly useful to link generations, addresses, and so on
Figure. Linking Triangles Across Generations in the USA

- Decennial Census Data* 1940 and Before
- Census 1950
- Census 1990-2000
- Census 2010
- Census 2020

Surveys

Administrative Records

Transfers Across Generations

Aggregate Accounts

Youngest Generation

Middle Generation

Current Generation

Robert M. La Follette School of Public Affairs
UNIVERSITY OF WISCONSIN–MADISON
Y, C and W (NW) -- USA, 1989-2016

• Consider C, Y and NW, all three for the same persons
• Findings----
--measures of one-dimensional inequality understate the level of inequality and the growth in inequality since 1989
-inequality in income (Y), consumption (C) and wealth (or net worth, NW) all rising separately
-inequality in any two dimensions increased faster than in any one dimension
-inequality in all three dimensions together rose by the most

See more below
Comparison of share held by top 5%
C, Y, W  one dimension

Year

Top 5% Share (%)
0 10 20 30 40 50 60 70

Wealth—SCF
Wealth—Saez & Zucman
Income—SCF
Income—Piketty & Saez
Consumption—SCF
Consumption—CE
2-D inequality: Top 5% shares in two dimensions – share of wealth by ranking

Robert M. La Follette
School of Public Affairs
UNIVERSITY OF WISCONSIN–MADISON
2-D inequality: Top 5% shares in two dimensions (1989=100)
3-D inequality:
Percent of households in top 5% of income, consumption, and wealth