



# **NTA data: 1981-2014**

## **Report of the Taiwan Team**

Kevin Yu-Ching Hsieh ([actung-2@econ.sinica.edu.tw](mailto:actung-2@econ.sinica.edu.tw))

An-Chi Tung ([actung@econ.sinica.edu.tw](mailto:actung@econ.sinica.edu.tw))

# Outline

1. Data introduction
2. Problems we have
3. A work schedule
4. Some applications

# 1. Data introduction

- Basics
  - ✓ Availability: annual data
  - ✓ Length: 1981~2014
- Problems (to be elaborated in Sec 2)
  - ✓ Some data are under close government control
  - ✓ Some data are inadequate
  - ✓ Some data are still under revision

# Data sources

- Micro
  - ✓ **Family Income and Expenditure Survey, FIES**
- Macro
  - ✓ **National Income, NI**
  - ✓ Education Statistical Yearbook
  - ✓ Public Financial Statistical Yearbook
  - ✓ Labor Insurance Statistical Yearbook (incl. FI, NP)
  - ✓ Government Employee Insurance Statistical Yearbook
  - ✓ National Health Insurance Statistical Yearbook
  - ✓ ...

# Main micro data source: FIES

- Format
  - ✓ Questionnaire
  - ✓ 1981-2013 (2014 forthcoming)
- Size
  - ✓ around 15 thousand households (universal sampling rate 0.2%), 40 thousand individuals
- Content
  - ✓ Individual data (income, some transfers)  
+ household data (private consumption)

# Main macro data source: NI

- Some data are in 2008 SNA
  - ✓ All data in 2007~2014
  - ✓ Highly aggregated variables in 1981-2006
- Other data are in 93 SNA (or 68 SNA)
  - ✓ Disaggregated variables of 1981-2006, including those data in “Income and outlay accounts in matrix format”
- How 2008 SNA differs from earlier systems
  - ✓ GDP increases because R&D is reclassified as investment
  - ✓ social insurances are reclassified into the public sector

## 2. Problems we have

- A. Some (micro) data are closely controlled by the government (*Issue 1*)
- B. Some data (e.g., social insurances) are inadequate (*Issues 2-4*)
- C. Some (macro) data are still under revision and not always consistent across years (*Issue 5*)

# 2A. Micro data availability

## Issue 1: data access restriction

- Facts

- ✓ All data in 1981-2006 are OK
- ✓ Individual-level data for 2007-2013, not accessible before 2013, are conditionally released after 2014

- Problem

- ✓ All computations have to be done inside the gov building

- Solution

- ✓ We use the data on site
- ✓ Inconvenient, but we get our results



# 2B. Various data problems

## Issue 2: over-aggregation

- Fact
  - ✓ For example, all cash social benefits (LI, GEI, FI,... ), except National Pension, are combined into one number
- Problem
  - ✓ We need to separate the total by social insurance program
- Solution
  - ✓ The “social insurance type” of each individual in FIES helps us to decide the type of social insurance

# 2B. Various data problems

## Issue 3: under-identification

- Fact
  - ✓ In 33% of all households (or 14% by amount), some beneficiaries of social benefits are “un-identified”
- Problem
  - ✓ To whom to allocate the benefit?
- Solution
  - ✓ Find out who is eligible and has no benefit reported,
  - ✓ if there is only one such person in the HH, allocate to him
  - ✓ In few cases when there are two or more such persons, we divide equally among them

# 2B. Various data problems

## Issue 4: meaning of data

- Facts about NHI contribution
  - ✓ Who is covered? Both the insurants and dependents
  - ✓ Who pays? The insurant, those who have wage income
  - ✓ How much to pay? It varies by wage level, type of occupation, number of dependents
- Problem
  - ✓ Do we attribute the premium to the insurant, or to both insurants + dependents?

# 2B. Various data problems

## Issue 4: meaning of data (continued)

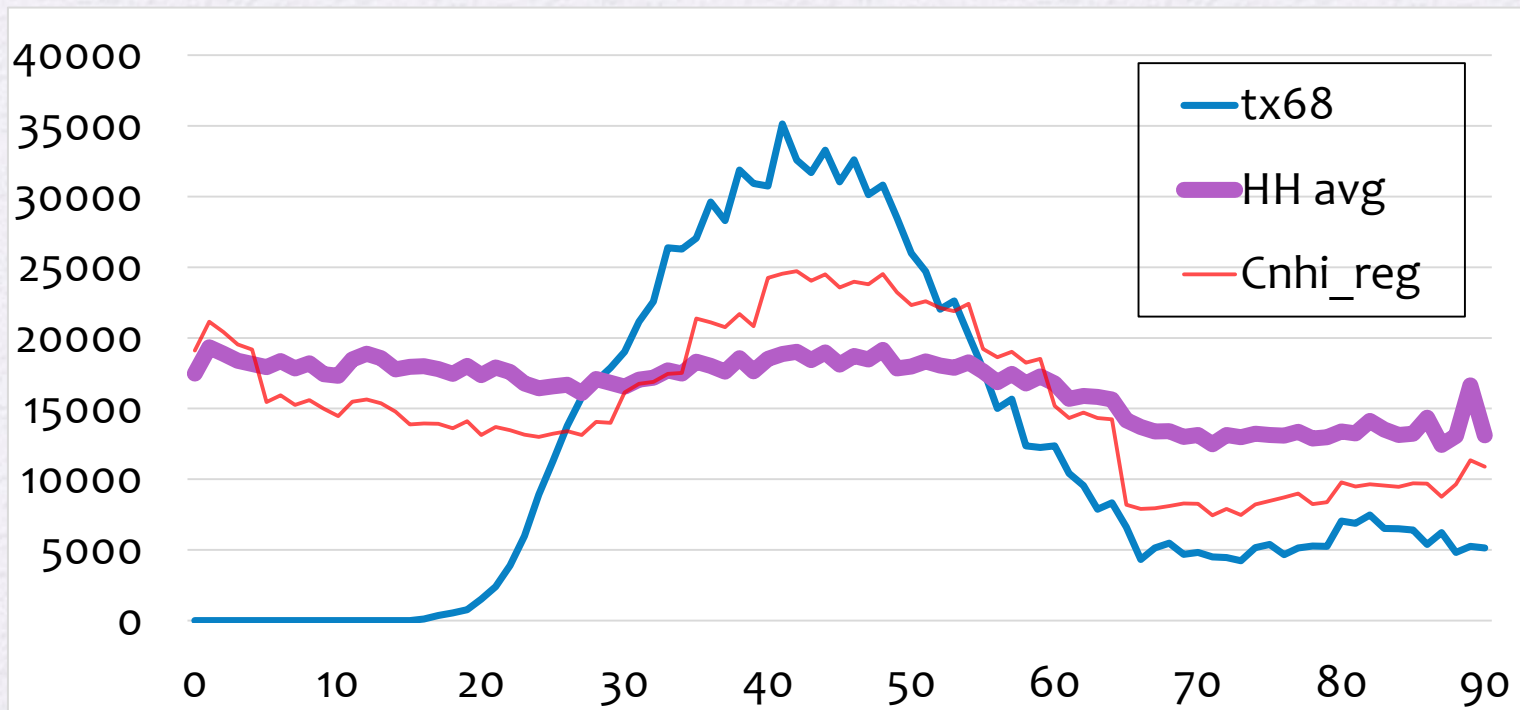
- Three possible solutions
  - ✓ Use the amount reported by each insurant
  - ✓ Allocate HH total to all insurants + dependents
  - ✓ Regression by age of insurants + dependents
- Problem
  - ✓ Which one to choose (we use the 1<sup>st</sup> method now)?

# NHI contribution by age

(a) amount reported by insurants

(b) HH average, by insurants and dependents

(c) Regression by age, by insurants and dependents



# 2C. Macro data inconsistency

## Issue 5: time consistency

- Facts

- ✓ All 2007-2014 data are in 2008 SNA
- ✓ For 1981-2006:
  - ✓ the most aggregated data (eg., total gov C) are revised into 2008 SNA
  - ✓ medium-level aggregated data (eg., gov C in edu) are mostly in 1993 SNA
  - ✓ most disaggregate data (eg., gov operating surplus) are in 1993 SNA or 1968 SNA

# 2C. Macro data inconsistency

## Issue 5: time consistency

- Problem

- ✓ The time series data are not time consistent

- Solution

- ✓ We use data in 2008 SNA whenever available.
- ✓ To accommodate the differences between aggregate and dis-aggregate variables, some variables are taken as given, and others are estimated as residuals accordingly
- ✓ any change in the macro total requires the re-calculation of the entire NTA accounts. Not a small work!

- Issue 1: data access restriction **solved!**
- Issue 2: over-aggregation **solved!**
- Issue 3: under-identification **solved!**
- Issue 4: meaning of data **searching for solution!**
- Issue 5: macro data consistency **solved!**



# 3. A work schedule

- By end of 2015
  - ✓ complete NTA: 2010, 2014
  - ✓ LHS: times series of YL, CF (tax adjusted), 1981-2014
- By end of 2016/2017
  - ✓ complete NTA:
    - ✓ 1981-2015 (so that earlier estimations are all revised)
  - ✓ An NTA operation manual in Chinese
  - ✓ Gender NTA: 2014

# 4. Some applications

- Policy evaluation:
  - ✓ National Pension Program: Generational inequity using Generational Accounting (Hsieh and Tung, 2015)
  - ✓ Long-term Care Insurance (Hsieh and Tung, ongoing)
- Times series or cohort analysis
  - ✓ Cohort analysis: Decomposition (Hsieh and Tung, ongoing)
- Others
  - ✓ Cross-country comparison (Under planning)

# Example 1: NTA and GA

- GA offers a clear and concrete indicator, *GI*, of public finance solvency and generational equity
- Yet GA hinges on a good set of age-specific data, such as NTA
- NTA framework helps to clarify the obscurity in GA, by not overlooking most of the in-kind transfers
- Application: We use NTA and GA to compare pre-NP and after-NP *GI*'s.

# Example 2: Long-term Care

- Long-term Care Insurance is underway
- Can we measure its financial deficits over time?
- How to predict future demand pattern, based on prevalence rising/declining of disability or dementia?
- Defined Contribution vs. Defined Benefit

- Thanks for your attention.
- Comments and suggestions are mostly welcome.