Further enhancing the work on household distributional data: Techniques for bridging gaps between micro and macro results and nowcasting methodologies for compiling more timely results

Jorrit Zwijnenburg (OECD)

Paper prepared for the 34th General IARIW Conference

Dresden, Germany, August 21 – 27, 2016

Discussant: Sigita Grundiza (Eurostat)

The views expressed in this paper are those of the author, and should not be considered as representing the official views of the OECD or of its member countries.

The author would like to thank Mr. Roberto Astolfi and Mr. Matthew De Queljoe for their contributions to the work on the nowcasting.

The views expressed in this presentations are those of discussant, and should not be considered as representing the official views of the Eurostat.

Outline

1. **Overview of the paper**

2. **Conclusion and further steps**

3. **Discussion**

1. Overview of the paper

This paper focuses on two issues that pose challenges in the compilation of distributional results in the work by OECD expert group on Disparities in National Accounts (DNA) :

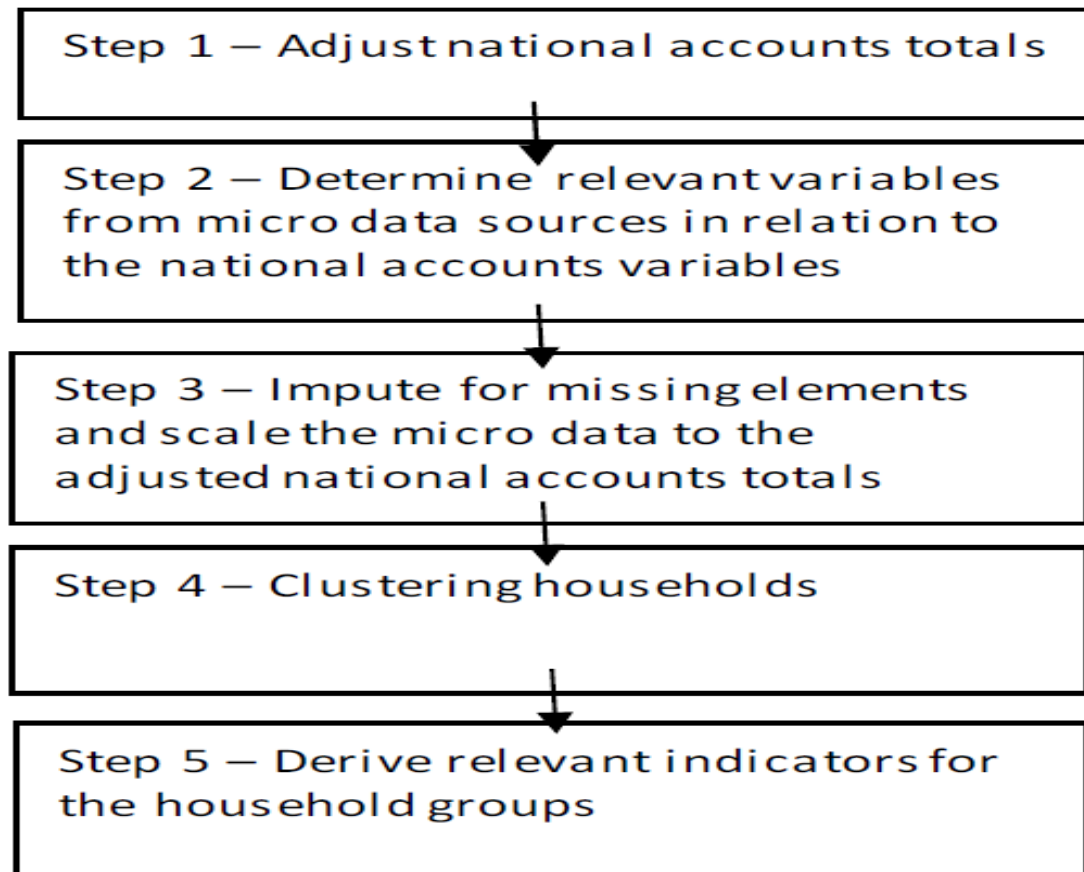**1.1 exsistent data gaps between micro and macro aggregates**

paper disucusses the possible reasons for these gaps and presents a framework is presented for allocating these gaps

1.2 the timeliness of the underlying micro data used in the compilation process is generally not suffiecient

the possible nowcasting techniques are disscussed to arrive at more timely distributional estimates
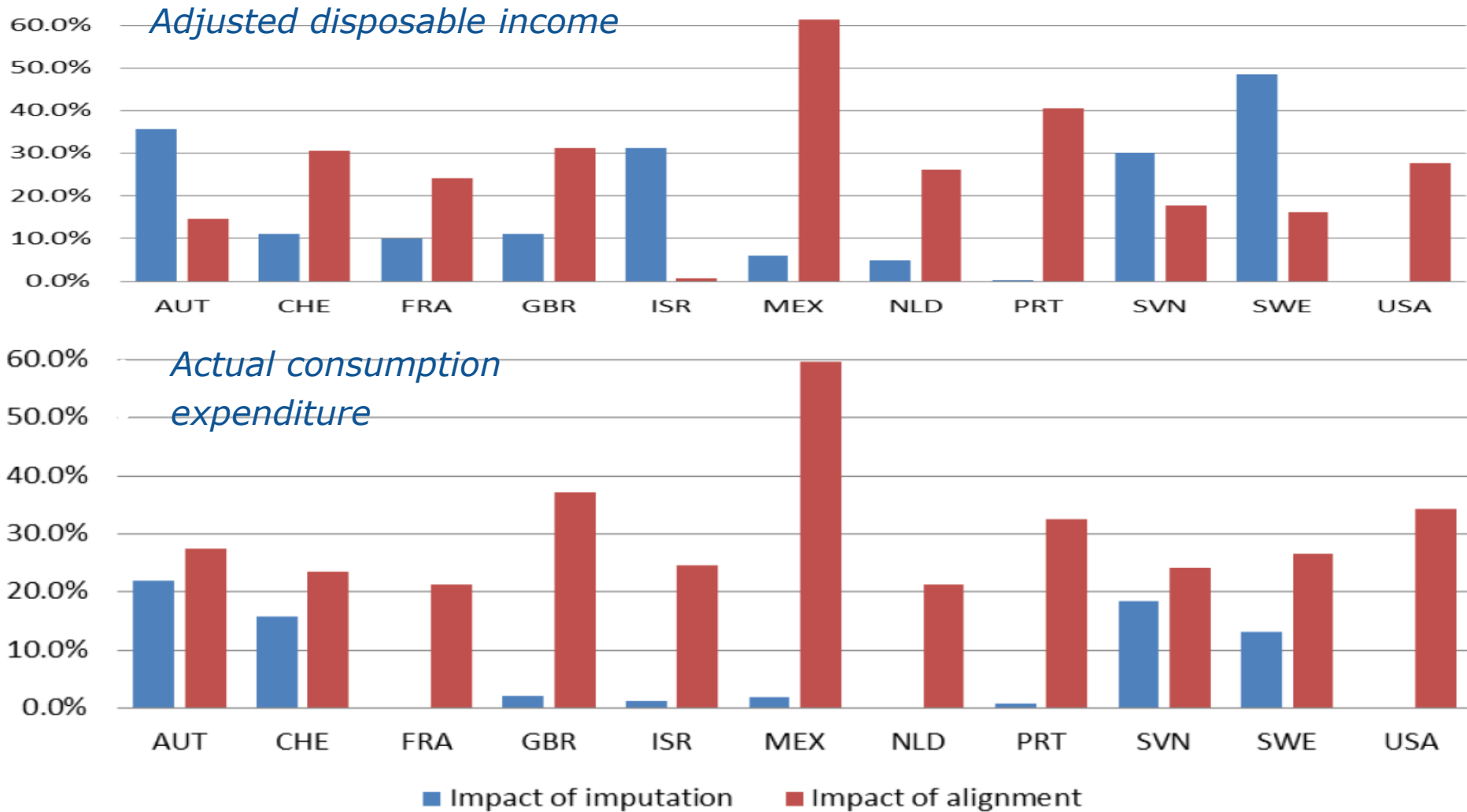
1. Methodology of the EG DNA

A step-by-step approach for the estimation of distributional information

Step 1 — Adjust national accounts totals

Step 2 — Determine relevant variables from micro data sources in relation to the national accounts variables

Step 3 — Impute for missing elements and scale the micro data to the adjusted national accounts totals

Step 4 — Clustering households

Step 5 — Derive relevant indicators for the household groups

1. Data gaps between micro and macro
Size of absolute alignments and imputations as % for HH

Adjusted disposable income

Actual consumption expenditure

■ Impact of imputation　■ Impact of alignment

1.1 Data gaps between micro and macro
Adjustment coefficient for main income and consumption components

Code	Instrument	Number of countries	Average	
			most recent year	second most recent year
B2	Operating surplus	6	1.47	1.27
B3	Mixed income	9	2.69	1.79
D1R	Compensation of employees	9	1.15	...
D41R	Interest (not adjusted for FISIM), received	8	2.08	1.90
D42R	Distributed income of corporations	7	5.06	10.67
D41P	Interest (not adjusted for FISIM), paid	9	3.58	2.47
D5P	Current taxes on income and wealth	10	1.18	1.19
D61P	Net social contributions	2	1.23	2.01
D62R	Social benefits other than STiK	10	1.22	1.30
D63R1	Social Transfers in Kind - Education	3	0.94	0.88
D63R2	Social Transfers in Kind - Health	3	1.36	1.37

Code	Instrument	Number of countries	Average	
			most recent year	second most recent year
CP010	Food and non-alcoholic beverages	10	1.48	1.53
CP020	Alcoholic beverages, tobacco and narcotics	9	3.60	5.37
CP030	Clothing and footwear	10	1.57	1.70
CP040	Housing, water, electricity, gas and other fuels	9	1.23	1.16
CP050	Furnishings, households equipment & house maint.	10	1.60	1.71
CP060	Health	9	2.47	2.72
CP070	Transport	8	1.56	1.59
CP080	Communications	10	1.25	1.53
CP090	Recreation and culture	10	1.90	1.85
CP100	Education	10	1.09	0.92
CP110	Restaurants and hotels	10	1.54	1.32
CP120	Miscellaneous goods and services	7	1.89	1.88

1.1 Data gaps between micro and macro: possible reasons

Step 1: Adjustment of the national accounts totals

i. The quality of the national accounts totals

ii. The quality of the adjustments to the national account totals

Step 2: Linking micro data source variables to the national accounts variables

iii. Assumptions regarding the conceptual and classification differences

Step 3: Imputation for missing elements and aligning data to national accounts totals

iv. The quality of the correction for the underground economy and illegal activities

v. The quality of the micro data – Estimation errors

vi. The quality of the micro data – Measurement errors

1.1 Framework to allocate the micro-macro gaps (1/4)

	Item xx.	Original estimate	Correction	Ultimate Estimate
I	National account total (A)	…	…	…
	- Adjustment for NPISH (B1)	…	…	…
	- Adjustment for non-private households (B2)	…	…	…
	- Adjustment for expenditures of non-resident households on the territory (B3)	…	…	…
	= Adjusted NA total (C=A-B1-B2-B3)	…	…	…
II	Micro total (D)	…		…
	= Macro-Micro gap (E=C-D)	…	…	…
III	Conceptual or classification issues (F)	…	…	…
	Underground and illegal activities (G)	…	…	…
	Other elements missing in micro data (H)	…	…	…
	Estimation errors (under-/overcoverage) (I)			…
	Measurement errors (under-/overreporting) (J)			…
	Reasons n.e.c. (K)			…
	= Remaining gap (L=E-F-G-H-I-J-K)			…

1.1 Framework to allocate the micro-macro gaps (2/4)

	Item xx.	Estimate	Q1	Q2	Q3	Q4	Q5
IV	*Allocation on the basis of micro data*						
	Original micro aggregate (P)	…	…	…	…	…	…
	Revised micro aggregate (Q)	…	…	…	…	…	…
V	*Allocation on the basis of meso corrections*						
	Conceptual or classification issues (R)	…	…	…	…	…	…
	Underground and illegal activities (S)	…	…	…	…	…	…
	Other elements missing in micro data (T)	…	…	…	…	…	…
	Estimation errors (under-/overcoverage) (U)	…	…	…	…	…	…
	Measurement errors (under-/overreporting) (V)	…	…	…	…	…	…
	Reasons n.e.c. (W)	…	…	…	…	…	…
VI	Alignment of remaining gap (X=C-Q-R-S-T-U-V-W)	…	…	…	…	…	…
	Final estimate (Y=Q+R+S+T+U+V+W+X)	…	…	…	…	…	…

1.1 Framework to allocate the micro-macro gaps (3/4)

Reasons for gaps between micro and macro aggregates	Average	Maximum
As percentage of original gap		
Correction of NA total	-1.1%	0.0%
Correction of adjustment for NPISH	0.2%	3.9%
Correction of adjustment for non-private hh's	0.0%	0.0%
Correction of adjustment for expenditure of non-residents	1.0%	9.1%
Correction of micro total	4.6%	82.5%
As percentage of gap after revision of adjusted NA totals		
Conceptual and classification issues	12.5%	100.0%
Underground and illegal activities	3.4%	55.9%
Other elements missing in micro data	11.2%	69.7%
Estimation errors (under-/overcoverage)	0.0%	0.0%
Measurement errors (under-/overreporting)	25.0%	100.0%
Reasons n.e.c.	1.9%	23.7%
Alignment of the remaining gap	46.1%	153.6%

1.1 Framework to allocate the micro-macro gaps (4/4)

Conceptual issues

Legend:
- Country A - B3 Mixed Income
- Country A - D11R Wages and salaries
- Country A - CP010 Food and non-alcoholic beverages
- Country A - CP020 Alcoholic beverages, tobacco and narcotics
- Country B - CP090 Recreation and culture
- Country B - CP110 Restaurants and hotels
- Country C - P31DC Final domestic consumption expenditure
- Country D - B3C Mixed income excl. underground economy and own production

Measurement errors

Legend:
- Country A - CP010 Food and non-alcoholic beverages
- Country A - CP020 Alcoholic beverages, tobacco and narcotics
- Country A - CP110 Restaurants and hotels
- Country E - D1R Compensation of employees
- Country E - CP010 Food and non-alcoholic beverages
- Country E - CP020 Alcoholic beverages, tobacco and narcotics
- Country E - CP070 Transport
- Country E - CP120 Miscellaneous goods and services

12

1.2 Compiling early estimates: possible approaches

A) A top-down approach starting from 'macro' distributional results for previous years

B) A bottom up approach starting from micro data

C) A meso approach starting from data on household groups

1.2 Compiling early estimates: A top-down approach starting from 'macro' distributional results for previous years

Step 1: Preliminary adjustment of national accounts totals

Step 2: Extrapolation of distributional results for the quintiles

*-The **direct approach** consists of directly estimating each transaction, regardless of their level of aggregation.*

*-In the **indirect approach**, 'child' transactions at the highest possible level of disaggregation are estimated directly and 'parent' transactions are obtained by summing up these estimated child transactions.*

- **A. Naïve method (t-1 or latest available shares)**
- **B. Average shares**
- **C. Linear trend Method**
- **D. Regression against the adjusted national accounts total**
- **E. Regression against auxiliary data**

Step 3: Alignment of the results to the adjusted national accounts totals

Step 4: Derivation of indicators

14

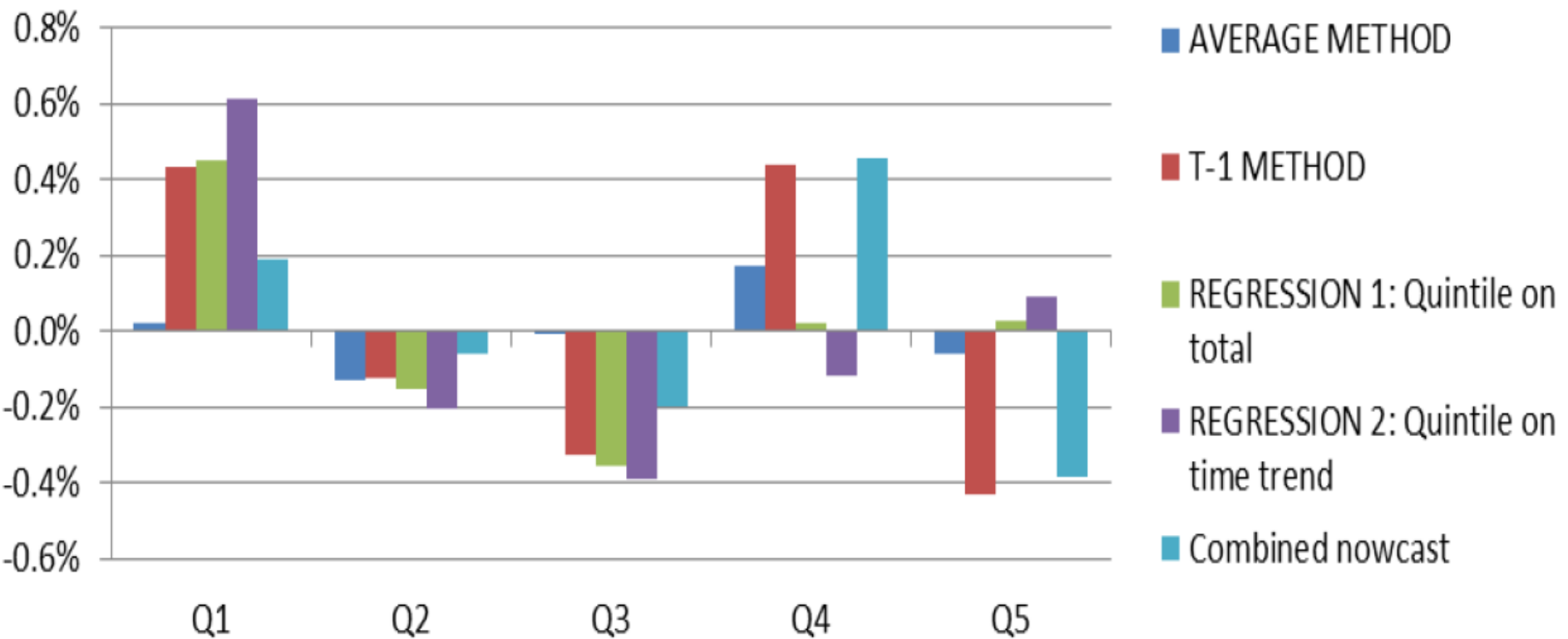# 1.2 Compiling early estimates: A top-down approach - Australia
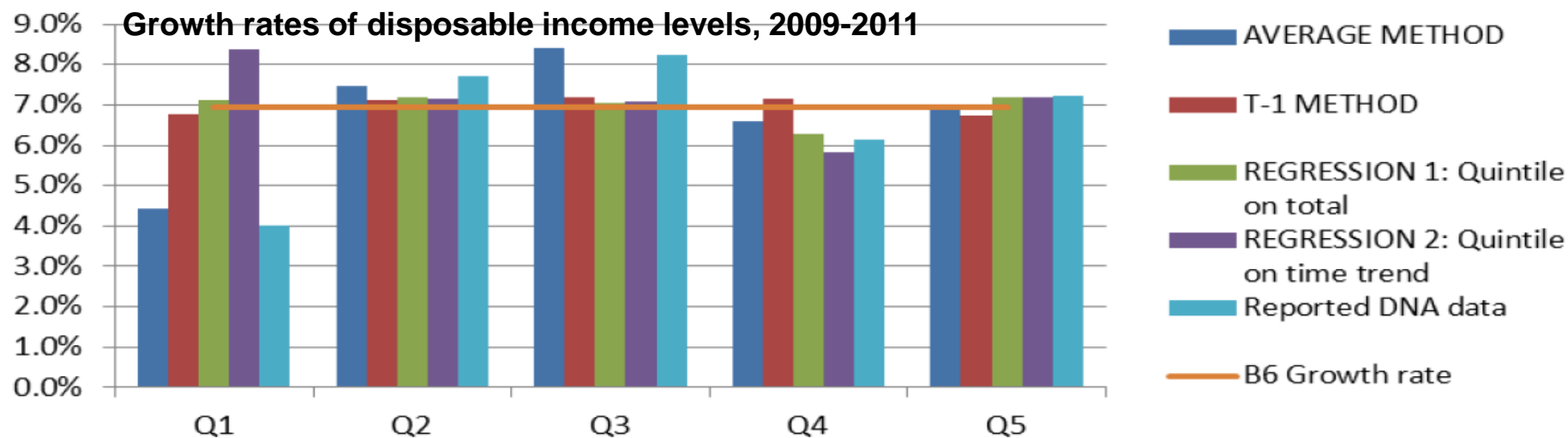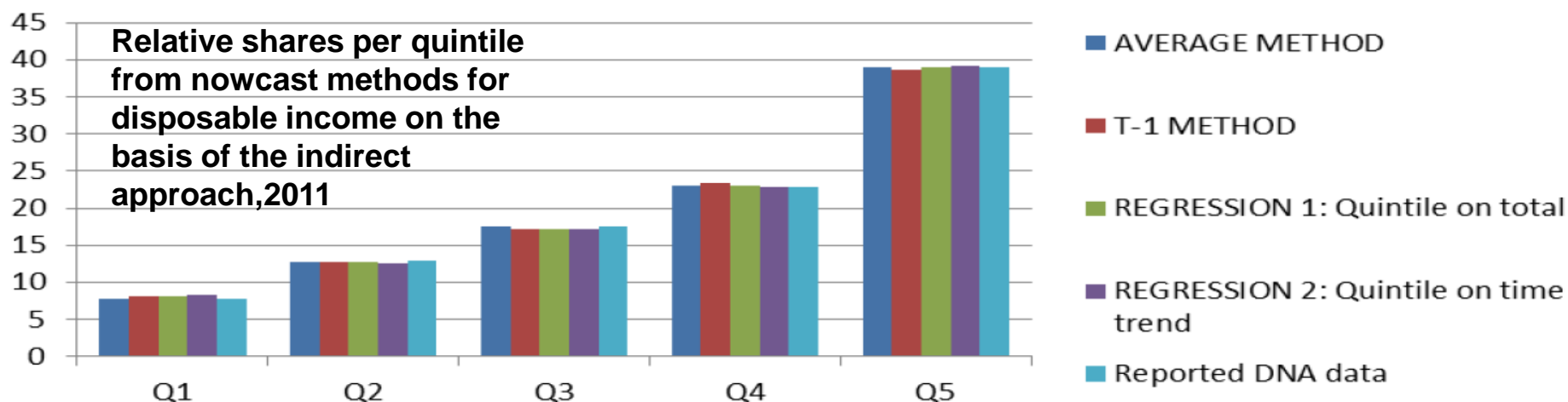
Relative shares per quintile from nowcast methods for disposable income on the basis of the indirect approach, 2011

- AVERAGE METHOD
- T-1 METHOD
- REGRESSION 1: Quintile on total
- REGRESSION 2: Quintile on time trend
- Reported DNA data

Growth rates of disposable income levels, 2009-2011

- AVERAGE METHOD
- T-1 METHOD
- REGRESSION 1: Quintile on total
- REGRESSION 2: Quintile on time trend
- Reported DNA data
- B6 Growth rate

1.2 Compiling early estimates: the top down approach

The gaps between the nowcast results and the actual values for the various underlying transactions of household disposable income for Australia, 2011
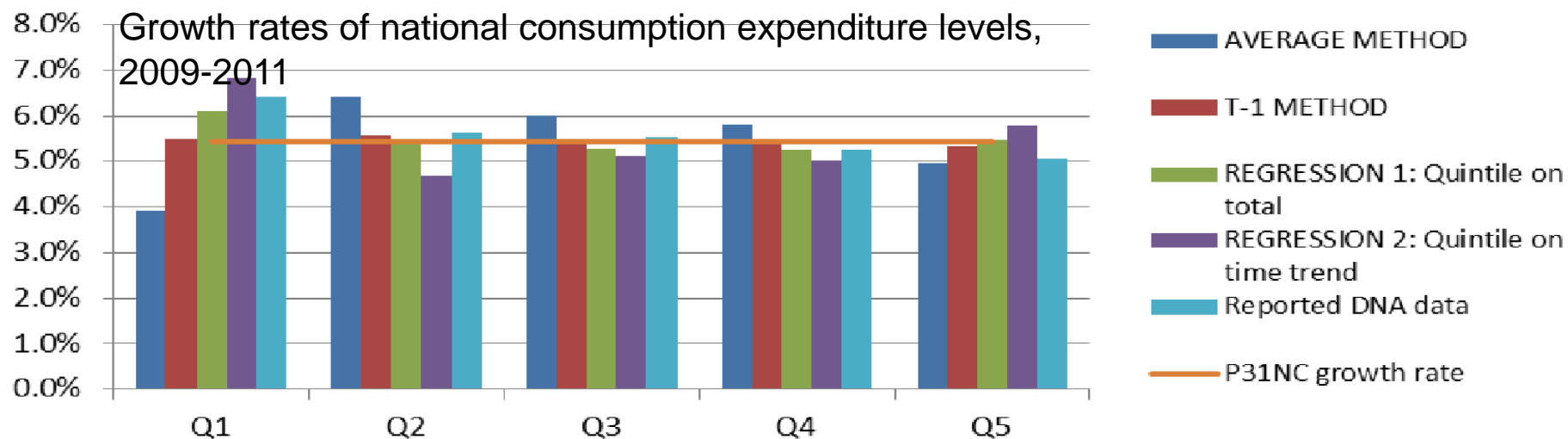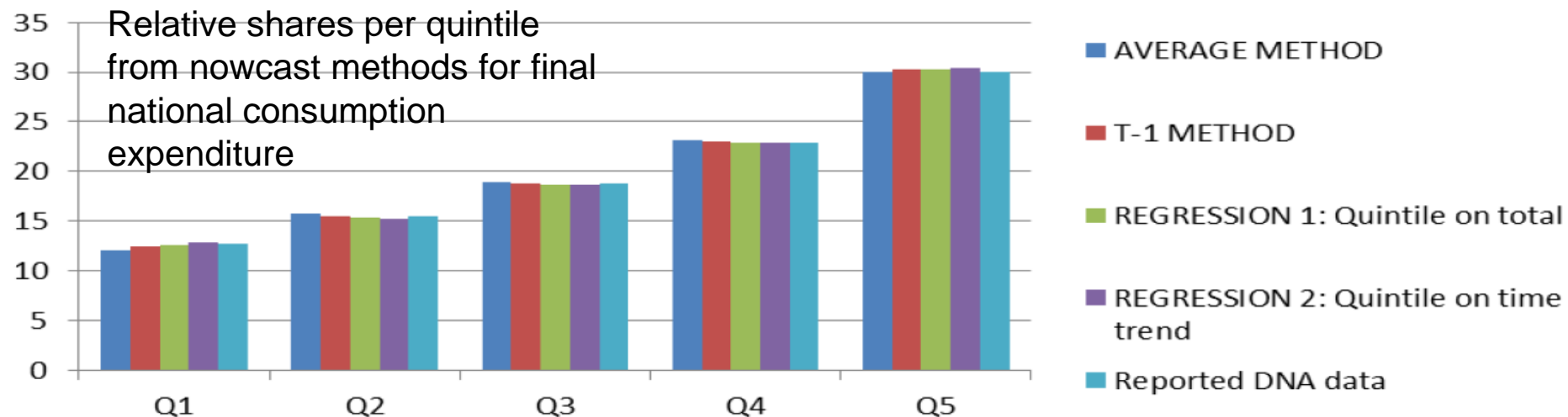
	Instrument	Average method	t-1 method	Regression 1	Regression 2	"Best" choice
B2	Operating surplus	7.34	49.04	50.85	50.19	Average method
B3	Mixed income	10.20	11.79	12.66	13.35	Average method
D1R	Compensation of employees	1.39	0.94	1.79	1.96	t-1 method
D4N	Net property income	18.35	13.32	16.07	17.82	t-1 method
D41'R	Interest received (excl. FISIM)	4.91	5.68	3.98	2.70	Regression 2
D42R	Distributed income of corporations	8.43	10.40	9.08	9.65	Average method
D45R	Rent received	10.33	9.14	11.69	9.26	t-1 method
D41'P	Interest paid (excl. FISIM)	4.05	5.25	4.21	6.89	Average method
D45P	Rent paid	10.33	9.15	8.97	9.27	Regression 1
D5P	Current taxes on income and wealth	3.57	1.82	3.74	2.88	t-1 method
D61P	Net social contributions	2.09	0.44	1.41	1.28	t-1 method
D62R	Social benefits other than STiK	4.30	3.07	3.41	4.79	t-1 method
D71P	Non-life insurance premiums	1.42	0.29	0.38	0.97	t-1 method
D72R	Non-life insurance claims	1.42	0.29	0.46	0.97	t-1 method
D75R	Misc. current transfers received	14.95	20.08	13.67	25.78	Regression 1
D75P	Misc. current transfers paid	6.29	1.95	2.30	3.69	t-1 method

1.2 Compiling early estimates: the top down approach

Difference between nowcast results and reference DNA data for the quintile shares of disposable income on the basis of the indirect approach, including the best choice ('combined') nowcast for Australia, 2011

Relative shares per quintile from nowcast methods for final national consumption expenditure

Growth rates of national consumption expenditure levels, 2009-2011

1.2 Compiling early estimates: the top down approach

The gaps between the nowcast results and the actual values for the various underlying transactions of consumption expenditure for Australia, 2011

		Average method	t-1 method	Regression 1	Regression 2	"Best" choice
CP010	Food and non-alcoholic beverages	1.65	0.60	0.23	0.24	Regression 1
CP020	Alc. beverages, tob. and narcotics	5.59	2.43	0.64	0.69	Regression 1
CP030	Clothing and footwear	6.42	2.15	2.32	1.26	Regression 2
CP040	Housing, water, electricity, …	0.52	0.15	0.17	0.16	t-1 method
CP050	Furnishings, households equip., …	5.12	1.77	0.80	0.75	Regression 2
CP060	Health	11.60	3.52	0.36	3.13	Regression 1
CP070	Transport	2.59	0.87	1.30	0.47	Regression 2
CP080	Communications	2.75	0.91	0.89	0.51	Regression 2
CP090	Recreation and culture	4.79	1.37	0.71	1.62	Regression 1
CP100	Education	6.80	1.92	1.04	2.42	Regression 1
CP110	Restaurants and hotels	3.71	0.93	1.19	1.56	t-1 method
CP120	Misc. goods and services	2.57	2.63	3.23	4.52	Average method

1.2 Compiling early estimates: the top down approach

Difference between nowcast results and reference DNA data for the quintile shares of consumption expenditure on the basis of the indirect approach, including the best choice ('combined') nowcast for Australia, 2011
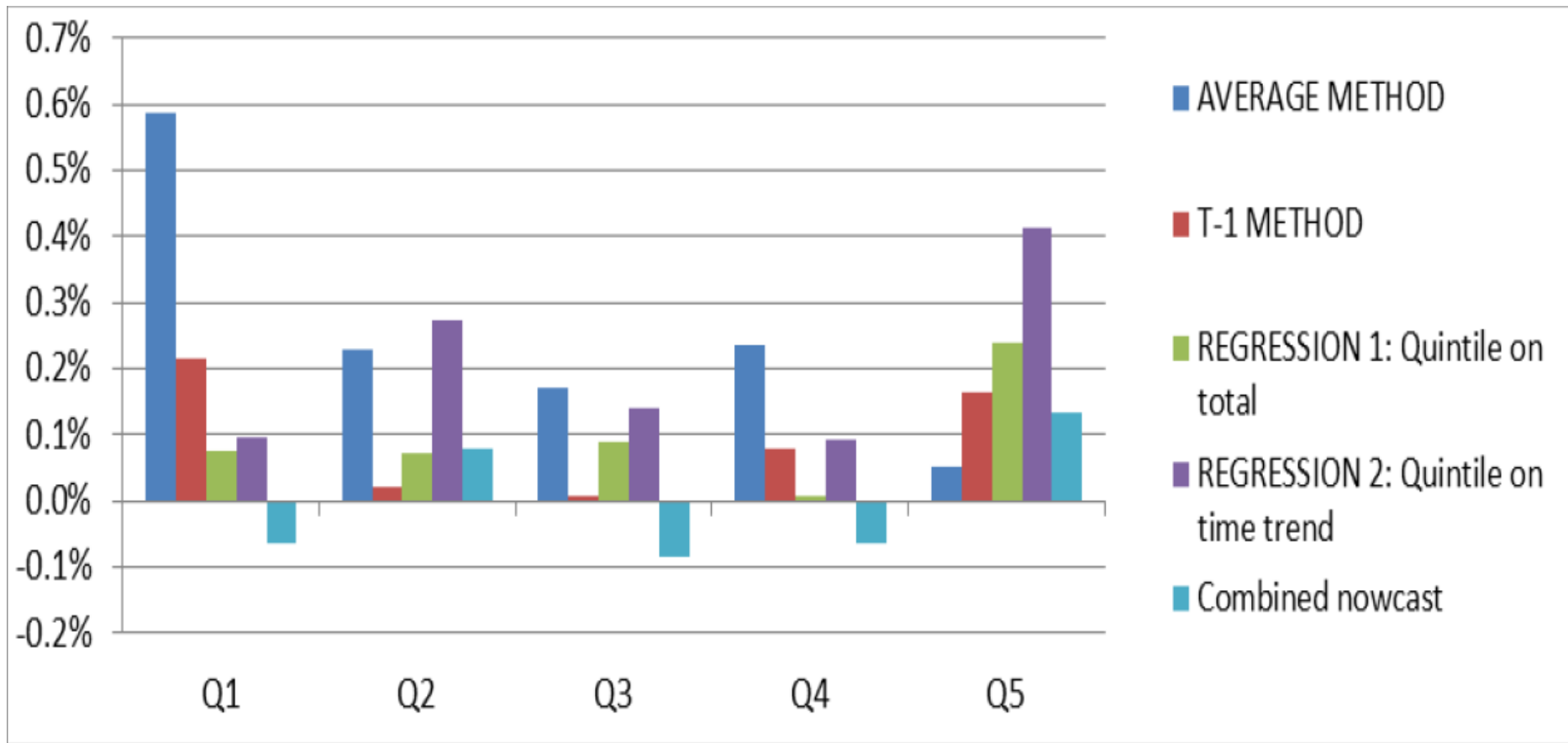
2. Conclusions and further work-
Data gaps between micro and macro aggregates

-Distributional results is very dependent on the quality and the alignment of micro and macro data used in the process;

-As currently not a lot of information is available on the reasons for these gaps most countries apply a proportional allocation to the underlying households;

-There may be various reasons underlying these gaps, that the allocation to households may vary across these reasons and that this may differ from the distributions derived from the micro data;

-A framework has been presented via which micro and macro experts can allocate the gaps to underlying reasons;

- More research is needed, amongst others via comparisons of different data sources and via linkages of data across different data sources.

2. Conclusions and further work – Compiling early estimates

-Preliminary results on the basis of these techniques within the top-down approach showed that nowcasts come relatively close to the reported shares of income and consumption by quintile. Larger differences can be are observed when analysing the growth rates. An appropriate combination of methods across all the underlying components may in that case lead to more accurate results;

-More research will be needed in this domain. The research should be broadened to test these techniques for a broader range of countries and on the basis of longer time series. Furthermore, on the basis of additional data sets, the applicability of the micro and meso approach should also be explored;

- In addition to some other issues that still need attention, the Expert Group on Disparities in National Accounts will continue its work on both the micro-macro gaps and the nowcasting. Furthermore, together with the European Central Bank, the Expert Group wants to start exploring possibilities to also include the wealth dimension in the work on distributional results.

3. Discussion

- Relevance of the quality of the micro data used for the distribution of NA aggregates;

- The NA aggregates for HH sector in some countries include the quasi corporations (up to 50 employees), this issue could be analyzed in details for adjustment for NA aggregates (similarly as for NPISH);

-Macro nowcasting methodology was presented for Australia, that is known to have high quality micro data for ICW from the survey, this is also the country where the distributional information for the most years is available – analysis of the results involving more countries is needed;

- Macro approach should be further tested together with the meso/micro approach; macro approach is based on very strong assumptions, that could turn to be misleading in the times of the rapid changes in the economic growth or implementation of fiscal/social policy in the country; the development quality framework could be considered.

- The use of the distributional NA information: the consistency of these indicators with the currently used indicators for measuring inequalities