



# THE ROLE OF IMPUTATIONS IN COMPILING DISTRIBUTIONAL RESULTS

ESCoE WORKSHOP:  
IMPUTATION OF DATA INTO HOUSEHOLD SURVEYS

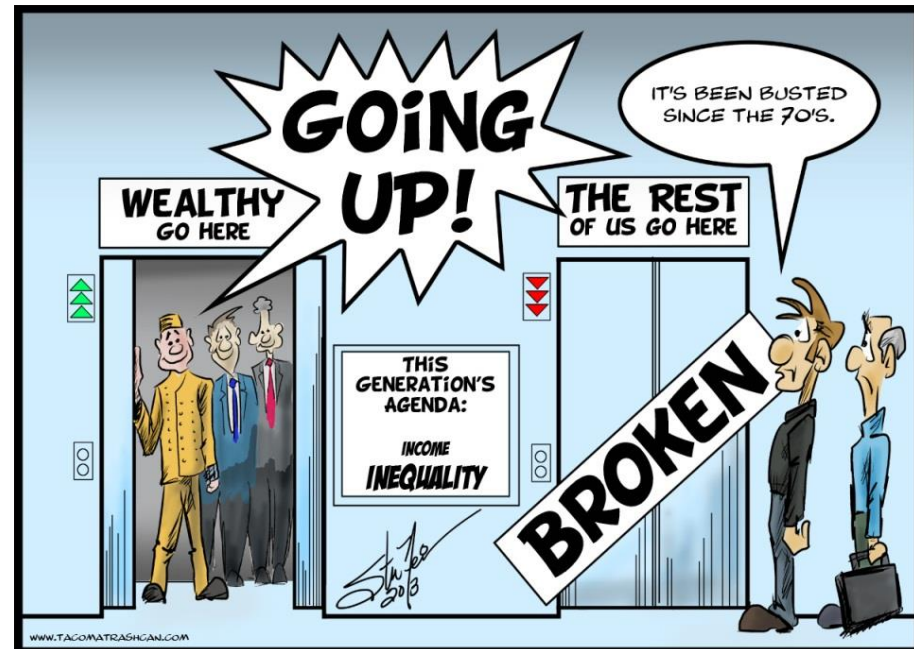
*LONDON, 2 OCTOBER 2017*

Presented by Jorrit Zwijnenburg (OECD)



# Contents

- Background of the EGDNA project
- Role and impact of imputations in compilation process
- Overview of items for which countries rely on imputations
- How to deal with imputed items?
- Conclusions





# Background

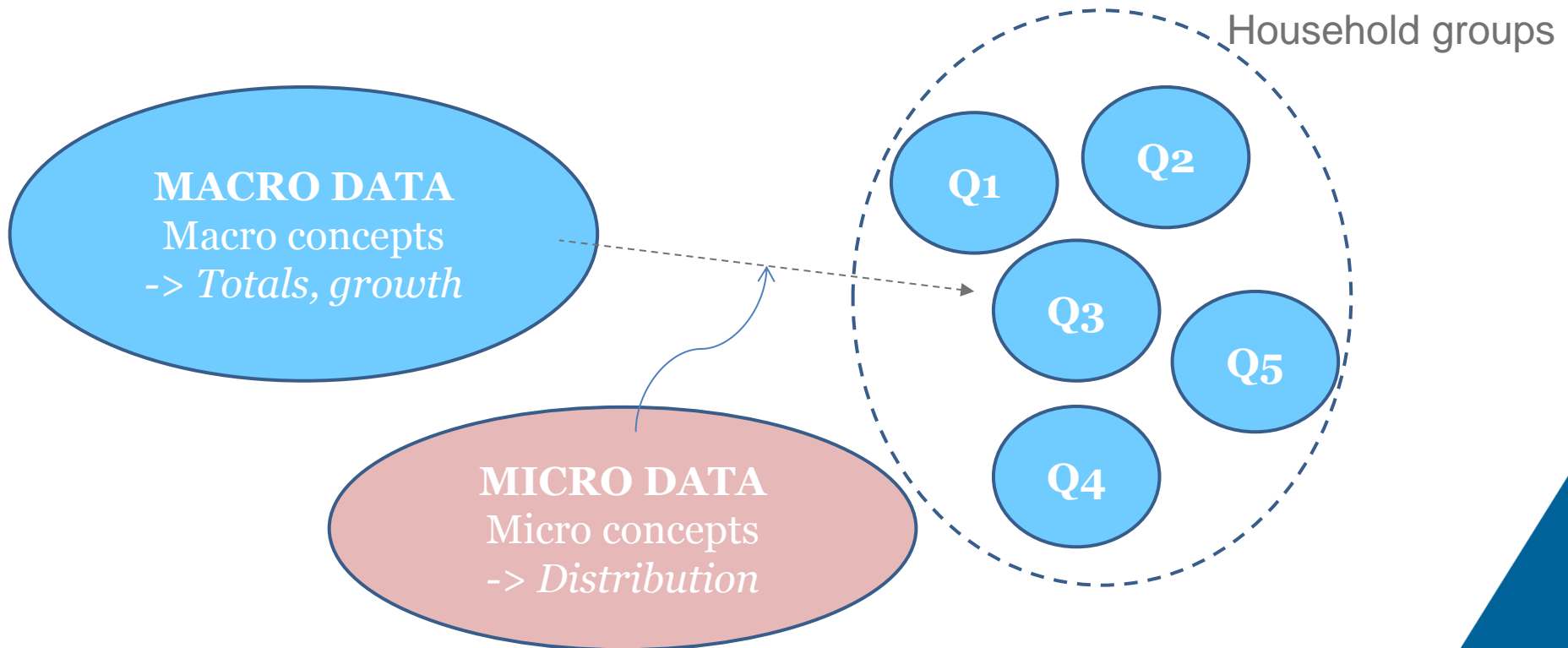
---

- 2009: Stiglitz, Sen, Fitoussi report stresses the importance of distributional results
- 2011: Launch of OECD/Eurostat Expert Group to develop methodology for compilation of distributional results within national accounts framework
- 2013: Publication of results from a first exercise
- 2014: Continuation of the work by an OECD EG DNA to further improve methodology and timeliness
- 2015: Second exercise on basis of improved methodology
- 2017: Publication of results from this second exercise



# Aim of the project (1)

Aim: Develop methodology to produce distributional results for household income, consumption and saving consistent with national accounts concepts using micro data sources





## Aim of the project (2)

---

- Distributional results for three household groupings:
  - Disposable income quintile (5 groups)
  - Main source of income (4 groups)
  - Household type (8 groups)
- The unit of analysis is the household
- OECD-modified equivalence scale is used to allocate households to quintile groups
- Calculations performed by members of the EG DNA: AUS, AUT, CAN, CHE, FRA, GBR, ISR, JPN, MEX, NLD, NZL, PRT, SVN, SWE, USA



# Aim of the project (3)

## HOUSEHOLD INCOME

### Income resources (received):

+ Self-employment income  
Imputed rent from dwellings  
Compensation of employees  
Property income

= **Primary Incomes (PI)**

+ Social benefits in cash  
Other transfers

= **Disposable Income (DI)**

+ Social transfers in kind

= **Adjusted Disposable Income (ADI)**

### Income uses (paid):

- Property income  
(e.g. interests paid on loans)

- Taxes  
Social contributions  
Other transfers

## HH. CONSUMPTION

### Expenditure:

+ Food  
Clothing  
Housing  
Health  
Education  
Transportation...

= **Consumption expenditure (CE)**

+ Social transfers in kind

= **Actual Consumption (AC)**

$$\text{Saving} = \text{DI} - \text{CE} = \text{ADI} - \text{AC}$$

## HH. SAVING



# Methodology: Step-by-step procedure

## **Step 1 – Adjust national accounts totals**

(exclude NPISHs, expenditures of non-resident hh's and people living in non-private dwellings)

## **Step 2 – Identify relevant variables from micro data sources that could be matched to NA variables**

(different data sources may be used for the various income and consumption items; concepts and classifications may deviate from national accounts)

## **Step 3 – Impute missing elements and scale the micro data to the adjusted national accounts totals**

(e.g. imputation for STiK, FISIM, income attributable to policy holders)

## **Step 4 – Cluster households into groups**

(e.g. on the basis of equivalized disposable income)

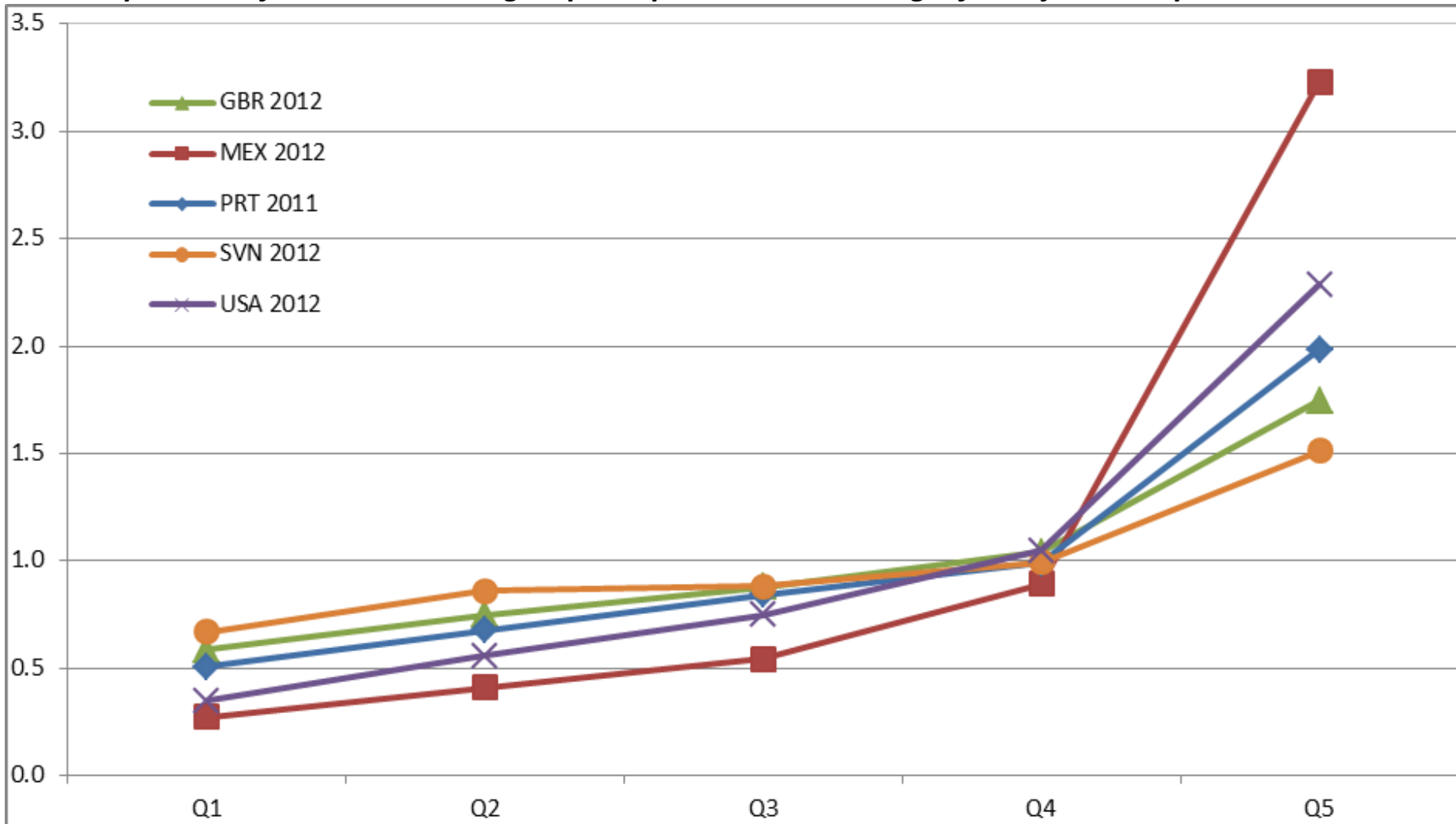
## **Step 5 – Derive relevant indicators for household groups**

(e.g. ratio to the average, highest to lowest)



# Results from 2015 exercise: Ratio to average

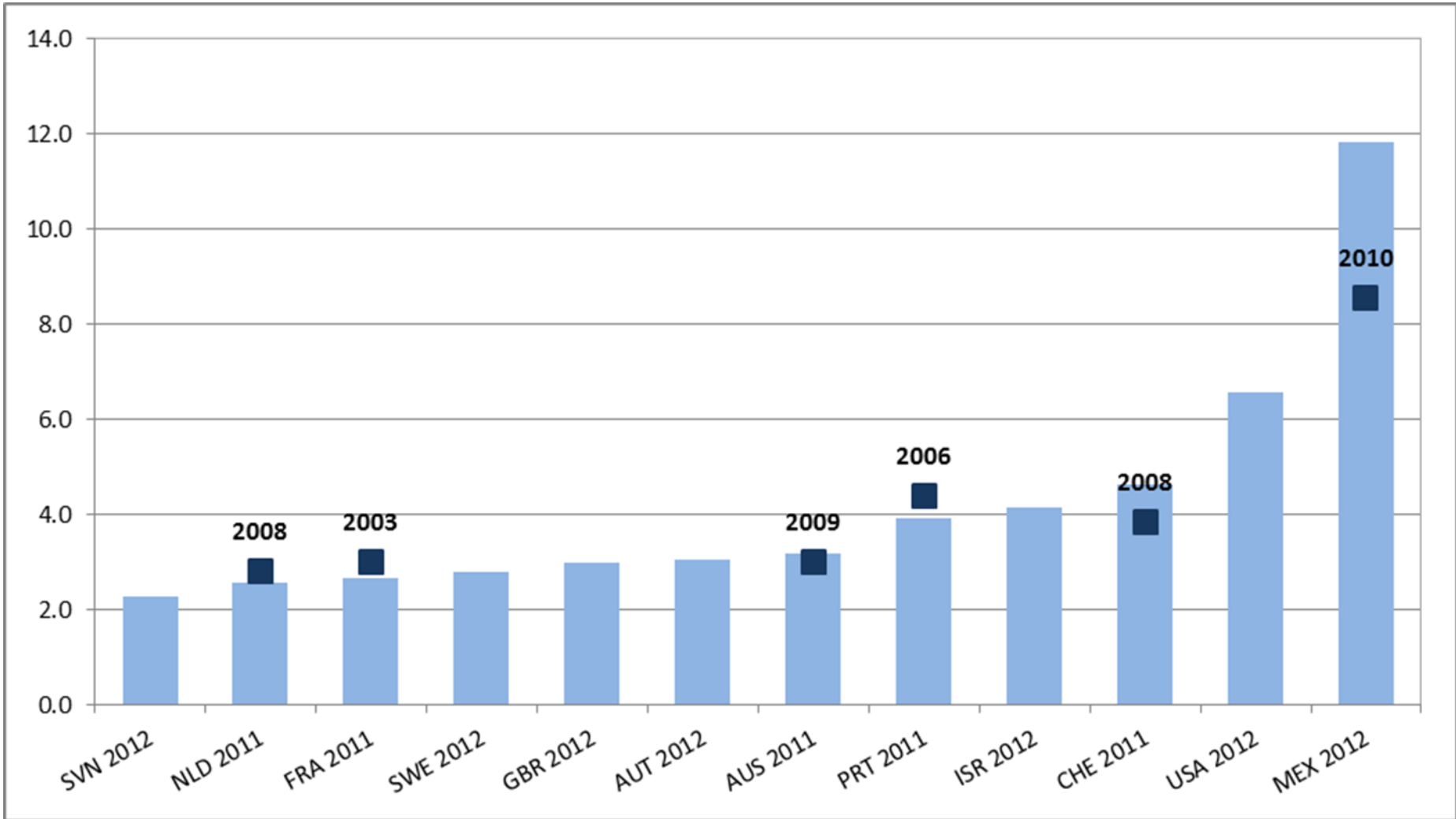
*Relative position of each household group compared to the average, for adjusted disposable income*







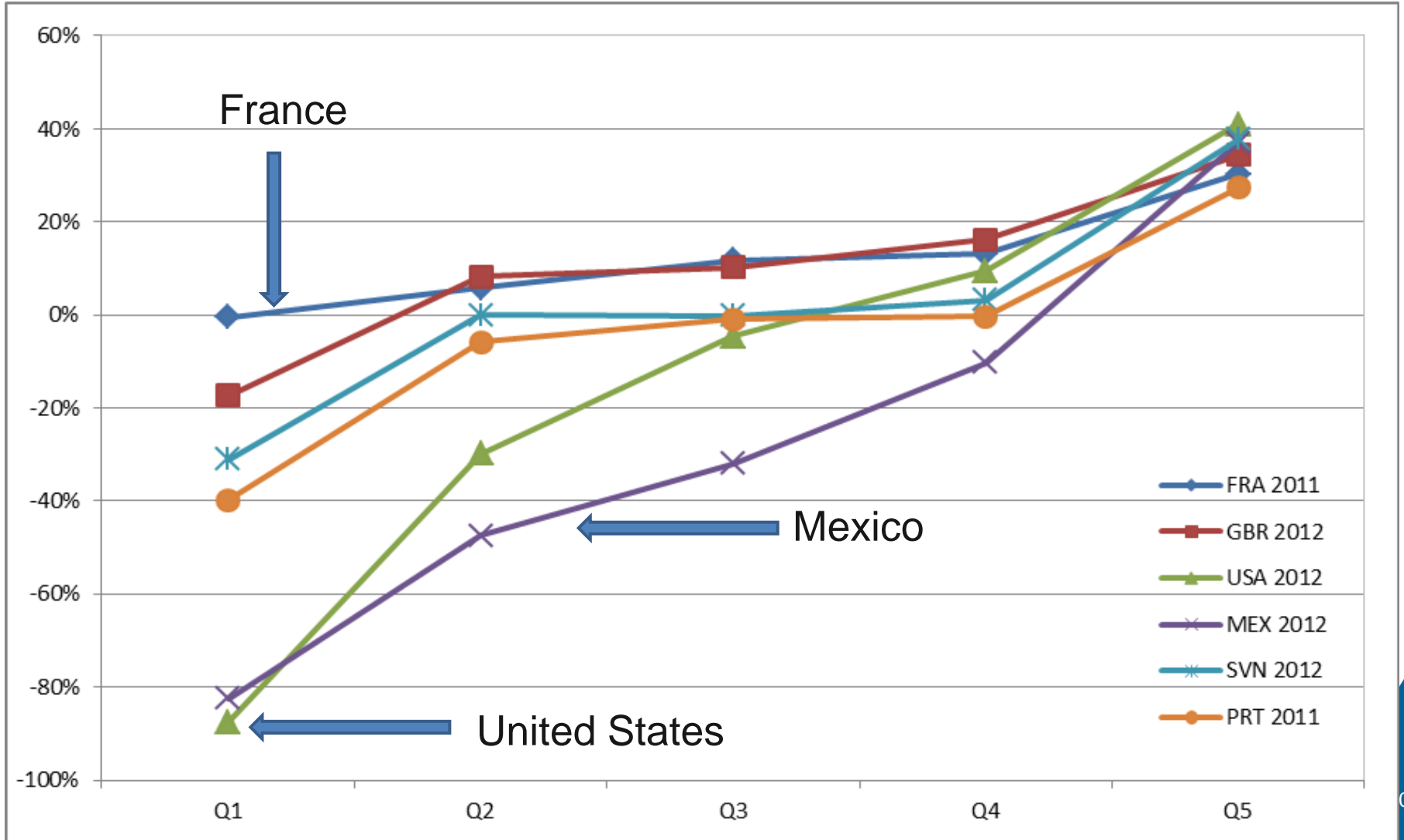
# Results from 2015 exercise: Ratio of highest to lowest income





# Results from 2015 exercise: Savings ratios

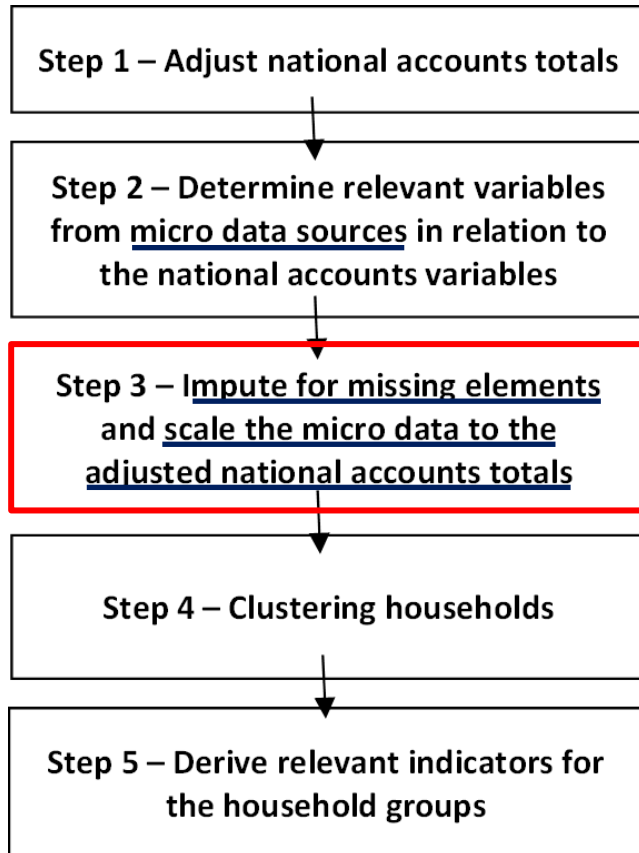
*Saving as percentage of disposable income by Equalized disposable income quintile*





# Role of imputations in compilation process (1)

Step-by-step approach:



Step 3: Impute for missing elements:

- In case no micro data is available
- In case micro data is not comprehensive (e.g. due to underground economy, illegal activities, thresholds, exemptions)

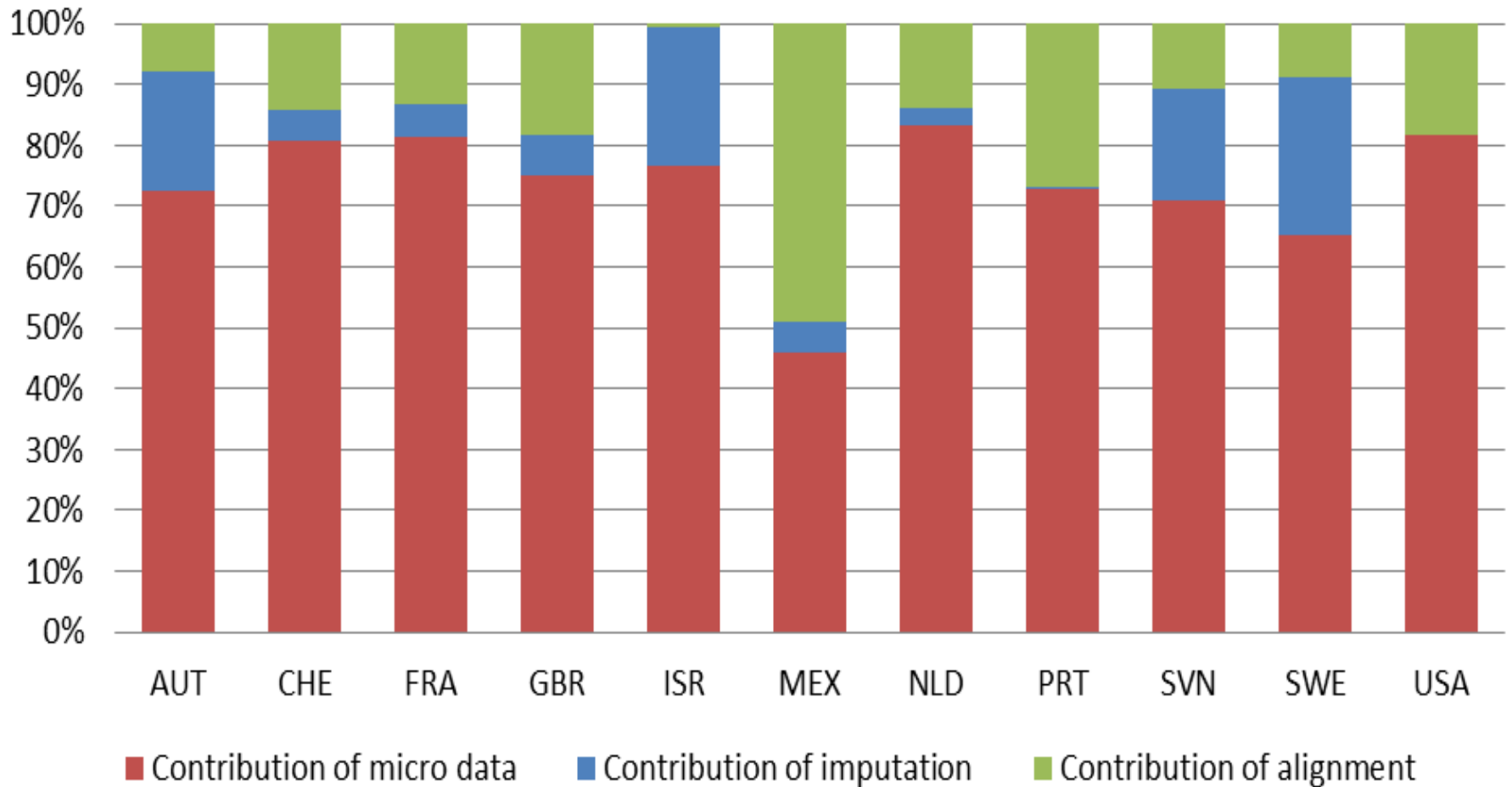
Distributional data are result of:

- Micro data
- Alignment to NA aggregates
- Imputations for missing elements



# Role of imputations in compilation process (2)

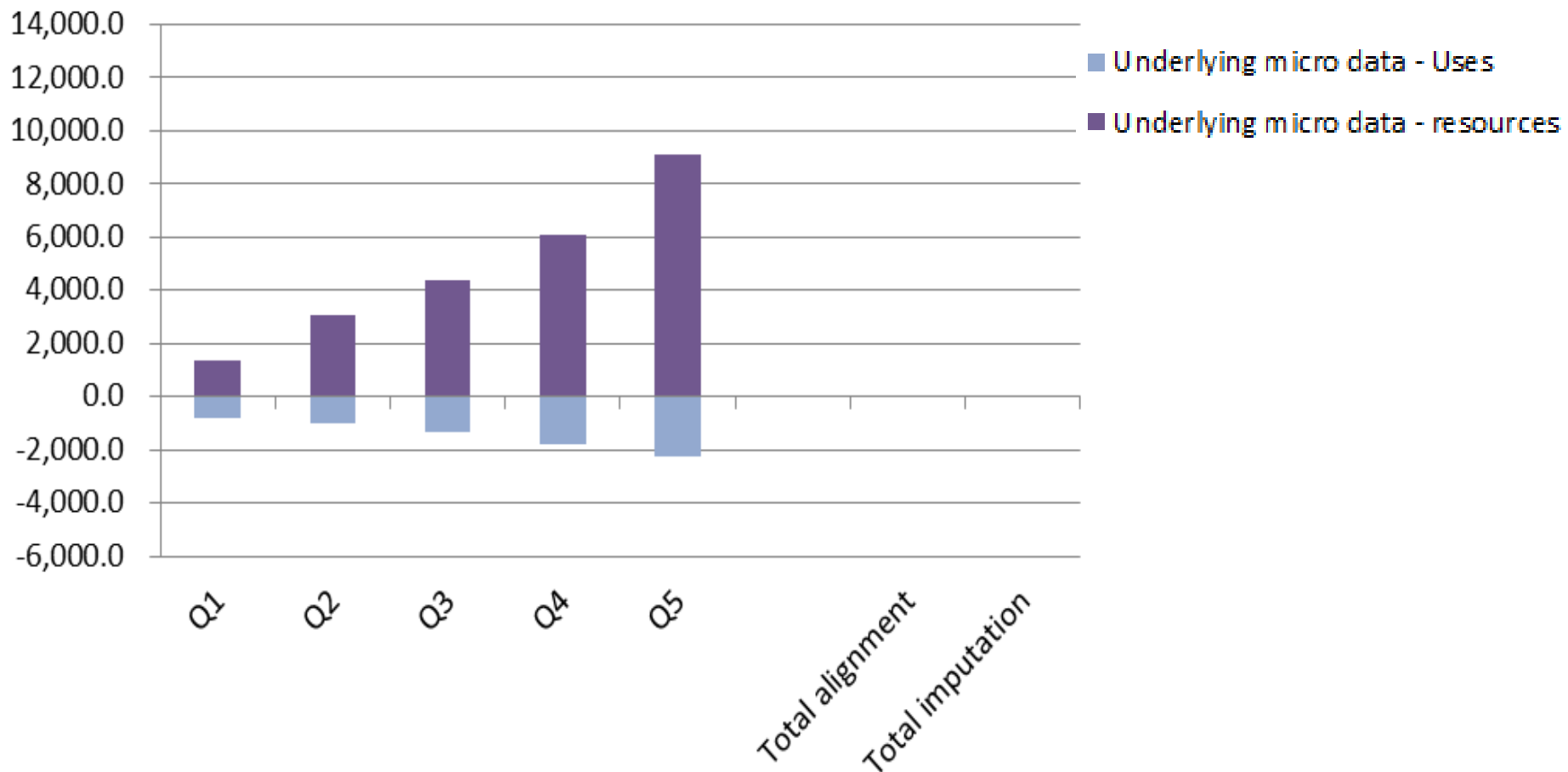
## Contributions to adjusted disposable income flows





# Possible impact of imputations – An example

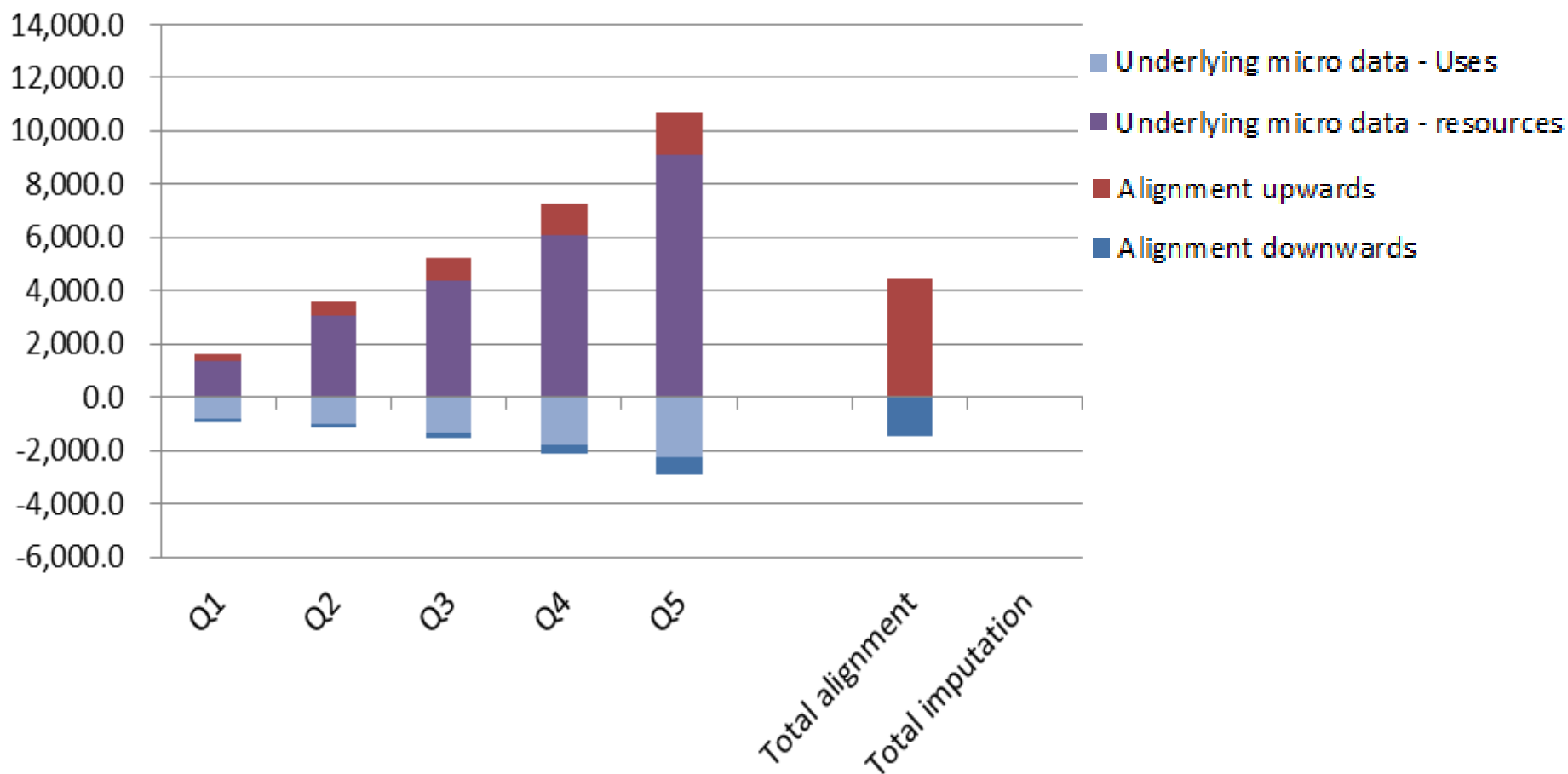
## 1. Allocation of micro data across quintiles





# Possible impact of imputations – An example

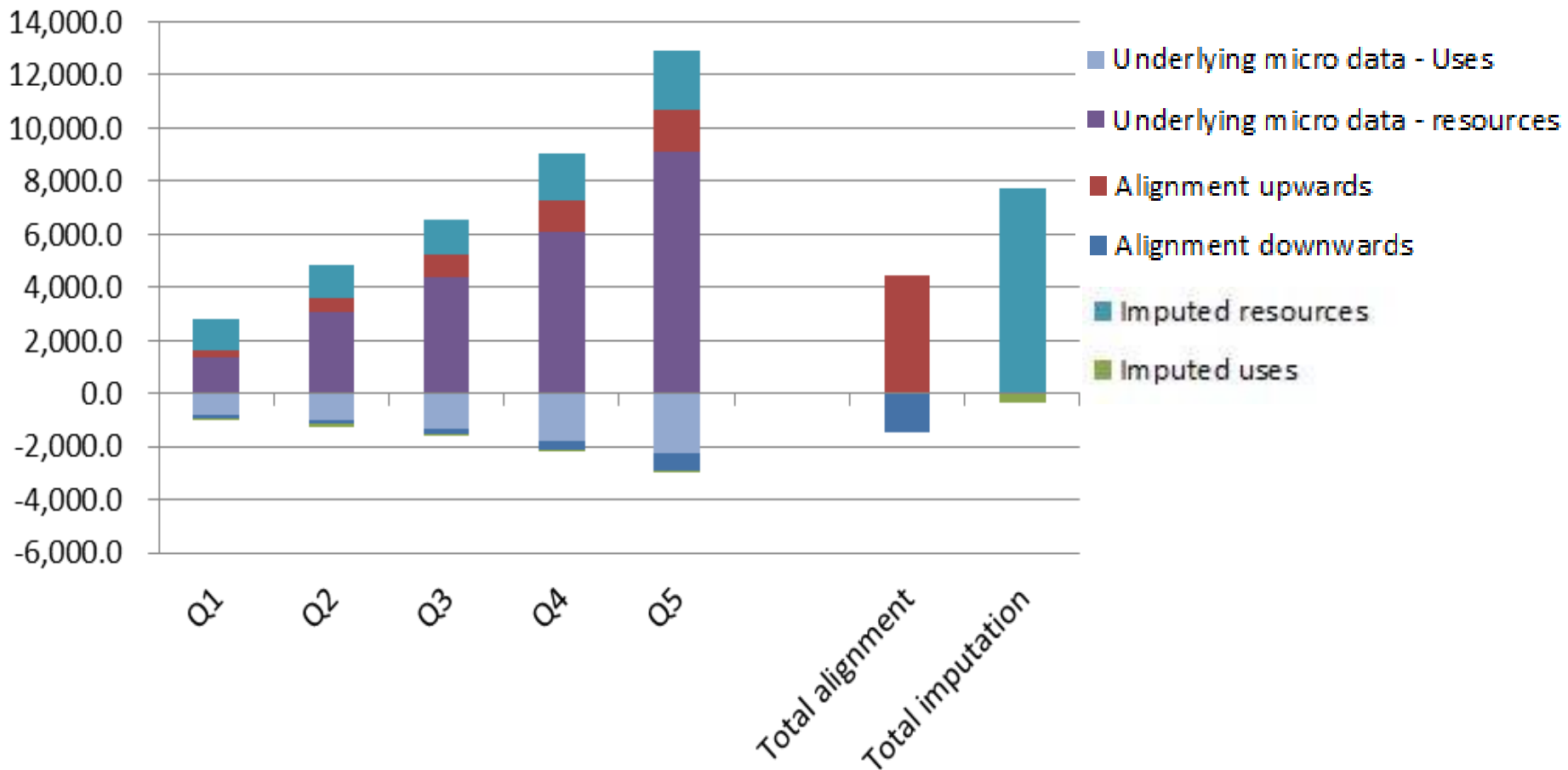
## 2. Allocation of micro-macro gaps across quintiles assuming proportional allocation





# Possible impact of imputations – An example

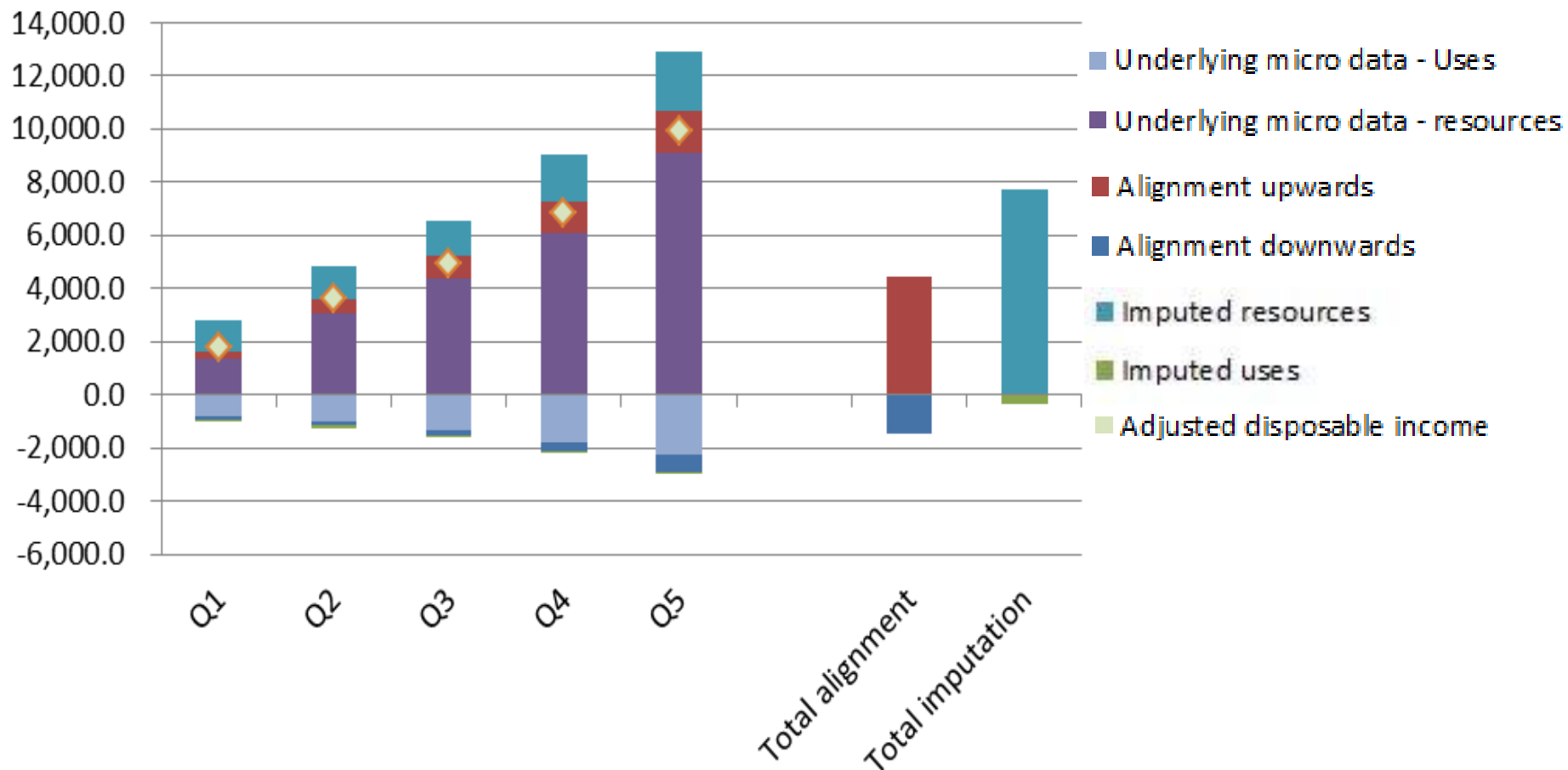
## 3. Allocation of imputations across quintiles





# Possible impact of imputations – An example

## 4. Distributional results across quintiles

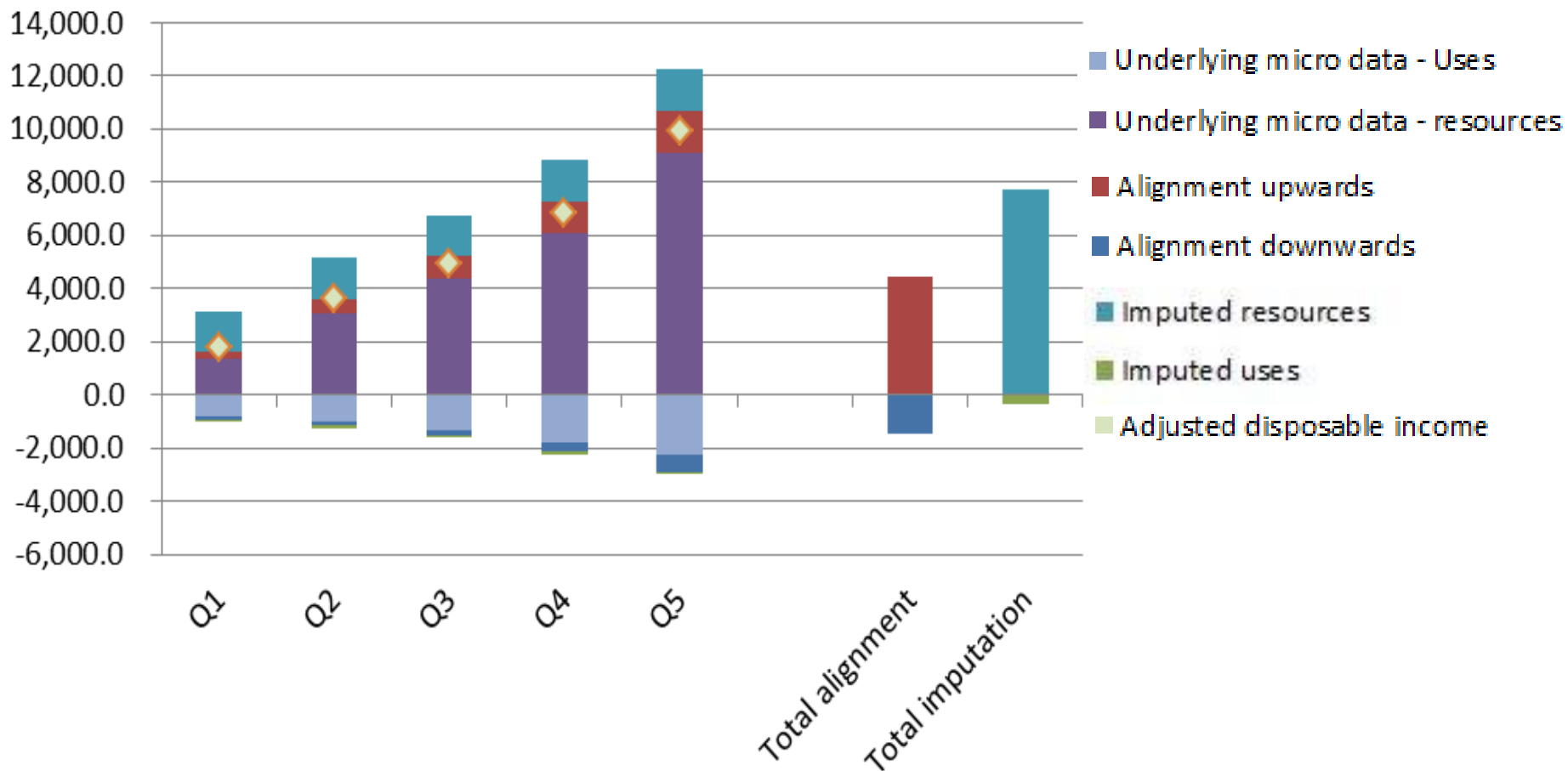






# Possible impact of imputations – An example

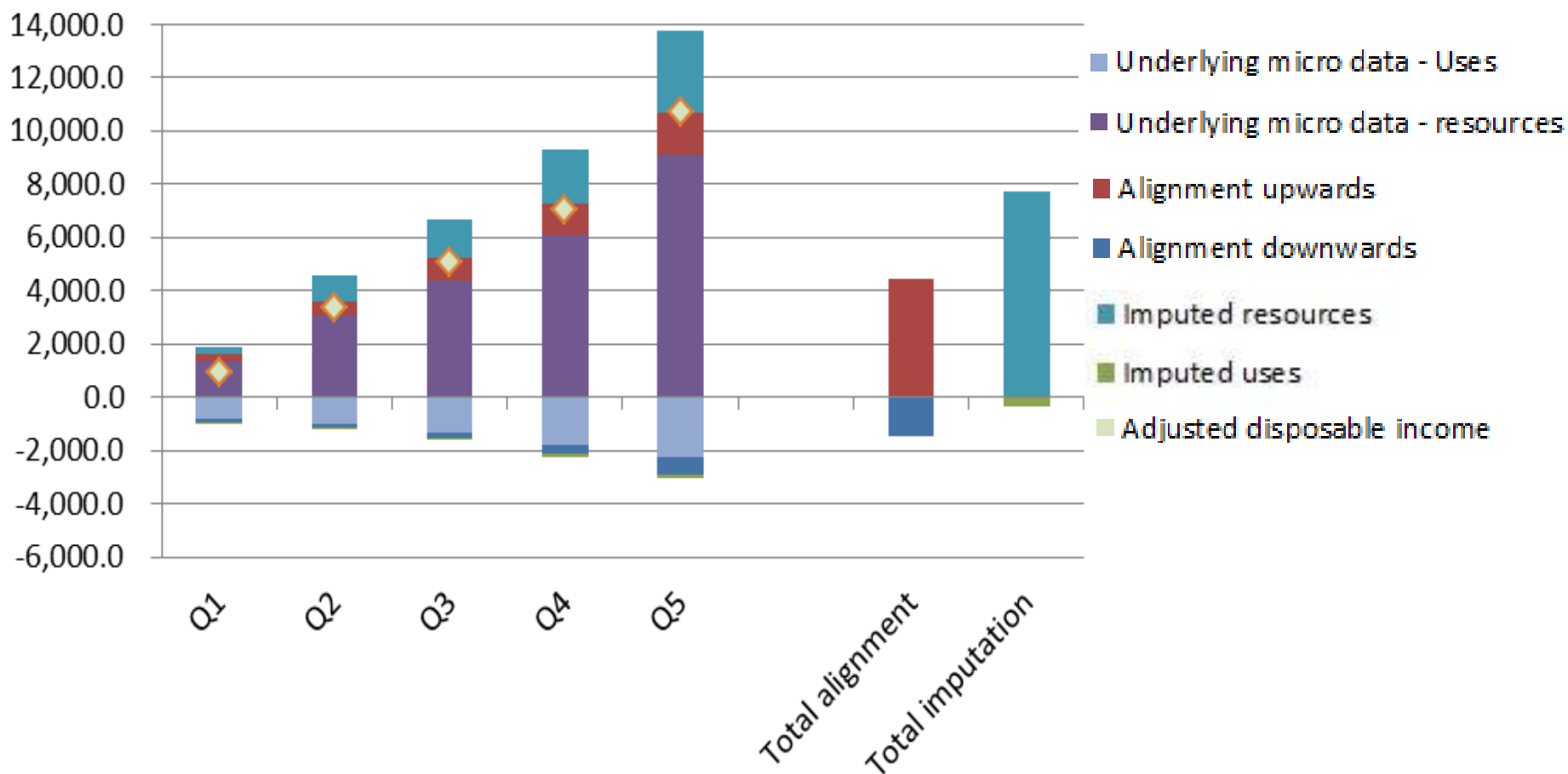
## 5. Alternative 1: Equal distribution across quintiles





# Possible impact of imputations – An example

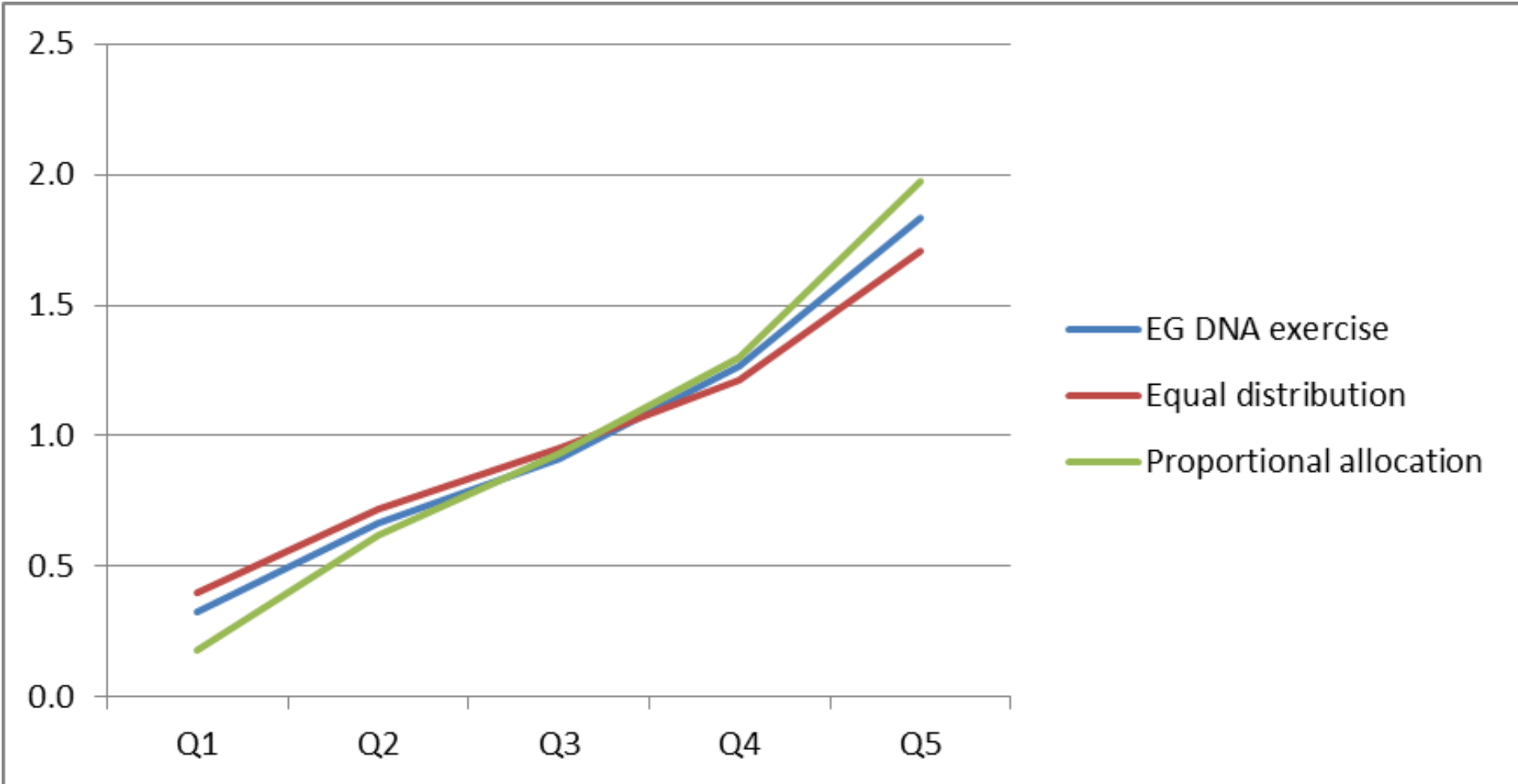
## 6. Alternative 2: Proportional allocation across quintiles





# Possible impact of imputations – An example

***'Ratio to the average' on the basis of the three approaches to allocate imputed items***





# For which items do countries rely on imputations?

SNA code	Item	# of countries	Share in Adj. disp. inc.
B2A	Owner occupied dwellings	3	5%
B3A	Own account production	4	1%
B3B	Underground production	3	2%
D122/D612	Employers' imputed social contributions	6	1%
D44	Investment income disbursements	6	4%
D45	Rent	6	0%
FISIM	Financial intermediation services indirectly measured	8	2%
D613+D614	Households' social contributions (actual and imputed)	3	2%
D72-D71	Net non-life insurance premiums	4	0%
D75	Net miscellaneous current transfers	6	0%
D63	Social transfers in kind	5	14%
P33	Resident households' consumption expenditure abroad	4	2%
D8	Change in net equity of households in pension funds	7	3%



# Issue of micro-macro gaps in relation to imputations

---

Gaps between micro aggregates and NA totals may relate to missing information:

- Conceptual differences: e.g. micro data on wages and salaries may not include expenses for food and transport to and from work.
- Underground economy: e.g. income from illegal activities or activities that are not required to register or fall below a certain threshold.
- Measurement errors: e.g. undercoverage of specific household groups or underreporting by specific household groups.
- .....

This may also need to be solved via specific imputations.



# How to allocate imputed items? (1)

---

Method 1: Indirect allocation based on the distribution of a related component.

## Examples:

- *Employers' imputed social contributions (D122 and D612) on the basis of Wages and salaries (D11)*
- *FISIM on the basis of Interest received (D41R) and Interest paid (D41P)*
- *Investment income payable on pension entitlements (D442) and Change in net equity of hh's in pension funds (D8) on the basis of pension contributions and benefits (parts of D61 and D62)*



## How to allocate imputed items? (2)

---

Method 2: Indirect allocation based on external data, such as socio-demographic information.

### Examples:

- *Owner occupied dwellings (B2A)* on the basis of property information from the land registry
- *Social transfers in kind (D63)* on the basis of actual value approach or insurance value approach (insurance premium equivalence based on socio-demographic information)
- *Investment income disbursements (D44)* on the basis of balance sheet information (insurance technical reserves, pension entitlements, investment fund shares)



## How to allocate imputed items? (3)

---

Method 3: Indirect allocation at the end of the process in a way that does not affect the overall distribution.

### Examples:

- *Resident households' consumption abroad (P33)* on the basis of *Final domestic consumption expenditure (P31DC)*
- *Non-life insurance premiums and benefits (D71 and D72)* on the basis of *Disposable income (B6)*

This method has to be regarded a last resort. Preferably, a proxy can be found to obtain a 'better' allocation.





## How to allocate imputed items? (4)

---

The framework of the National Accounts may also provide insight on how imputed amounts may best be allocated to relevant households or household groups.

Confrontation of income, consumption and wealth may show discrepancies which may relate to missing information for specific groups.

E.g. France: If consumption  $>$  income without households stating that they have to sell assets or incur liabilities, this could be a sign that some income elements are missing.



# Conclusions

---

- Imputations may have a large impact on distributional results
- In addition to specific imputed items, micro-macro gaps may also point to missing information and the need for imputations
- For each imputed item one needs to find the best proxy
- It is best to allocate imputations at the most detailed level of items
- The framework of the National Accounts may assist in allocating imputed amounts to underlying households
- Further exchange of best practices may benefit countries in dealing with imputations



Thank you for your attention

For more information please contact:  
[Jorrit.Zwijnenburg@oecd.org](mailto:Jorrit.Zwijnenburg@oecd.org)