



ΙΔΡΥΜΑ ΟΙΚΟΝΟΜΙΚΩΝ & ΒΙΟΜΗΧΑΝΙΚΩΝ ΕΡΕΥΝΩΝ
FOUNDATION FOR ECONOMIC & INDUSTRIAL RESEARCH

Weighting of responses in the Consumer Survey: Effects on the Consumer Confidence Indicator

**In the context of:
TASK FORCE ON THE QUALITY OF BCS DATA**

*Michail Vassileiadis, Head of Macroeconomic
Analysis and Policy Unit, IOBE (EL)*

Joint EU/OECD BCS Workshop, 14-15 November 2013

Contents

1. Review of the weighting approaches in voluntary surveys

- Introduction
- Features of the weighting approaches
- Sampling methods – weighting approaches

2. Weighting approaches followed in the DG ECFIN Consumer Survey

- Main characteristics of the weighting approaches
- Update frequency of the weights

3. Impact of weighting approaches on variance / tracking performance of the Consumer Confidence Indicator

- Introduction
- Effects on volatility
- Effects on tracking performance
- Impact of weighting features on volatility-tracking performance of the CCI

4. Main Findings - Suggestions

1. Review of the weighting approaches in voluntary surveys

Introduction

- The application of a **weighting approach** in a voluntary survey is **usually linked** to the **sampling method** that is followed in an earlier stage
- Even **non-weighting** of the sample responses is linked to **certain sampling methods**
- However, since it is not very unusual to follow a **combination of sampling approaches**, the weighting methods are **not attached** to them

Features of the weighting approaches

- **Population size weighting**

- Usually applied in case of examination of sample data from two or more countries
 - Used for the **harmonisation of the data**, so as to reflect the **relative significance of each country's population size**

- **Design weighting**

- Used for smoothing the differences in the selection probabilities of the respondents. Design weighting helps to:

- a) **Make the sample statistics representative** of the underlying population or
- b) **Compensate for over- or under-sampling** of specific cases or for disproportionate stratification

- The design weights are computed by the normalised inverse of the inclusion probabilities

Features of the weighting approaches

- **Post –stratification or non – response weighting**
 - **Non-responsiveness:** a source of bias if the non-respondents are a non-random fraction of the surveyed population (usually the case)
 - **Non-response weighting** compensates for this bias
 - In this scope, non-response weighting is used for **handling unit non-response in surveys**

Features of the weighting approaches

- Most non-response weighting schemes involve “**post-stratification**” which is in essence a two-step procedure:
 - (i) **Identify a set of “control totals”** of the population that the survey ought to match;
 - (ii) Calculate **weights** to adjust the **sample totals** to the control totals
- Other non-response weighting schemes: **Raking or iterative solutions / proportional fitting, logistic regression, calibration weighting**

Sampling methods – weighting approaches

- **Design weighting** is usually used when a **non-probability sampling method** is implemented (ex. quota sampling)
- The surveyed units are selected according to previously made assumptions regarding the population of interest →
 - Selection is nonrandom, a potential source of exclusion bias
 - Thus, design weighting is applied
- In case where the drawn sample is “**nationally representative**” weighting is not needed

Sampling methods – weighting approaches

- **Non-response weighting** is mainly followed in **probability-based sampling** (ex. random sampling)
- In probability – based sampling the formation of the sample does not follow some selection criteria →
 - Room for **non-responsiveness** by certain population groups, another **source of bias** for the sample based estimations
 - To avoid such an event, application of **non-response weighting**

2. Weighting approaches followed in the DG ECFIN Consumer Survey

Main characteristics of the weighting approaches

- ❖ Information extracted from **IOBE's questionnaire** in the context of the Task Force on the quality of the BCS data and from the **DG ECFIN Consumer Survey metadata**
- ❖ **24 out of 30 countries** participating to the **Consumer Survey** replied to the questionnaire (**Thank you for your assistance!**)

Main characteristics of the weighting approaches

Sampling methods

- ✓ The majority of the countries (19 out of 31) use a **random sampling method** (simple, stratified or systematic)
- ✓ In 10 countries **quota sampling** is applied
- ✓ Only two Consumer Surveys (Italy, Spain) use a combination of random sampling-quota sampling

Use of a weighting approach

- ✓ The vast majority of the countries (24 out of 31) apply a weighting method to the sample responses for the Consumer Survey
- ✓ Almost all the Consumer Surveys following a probability sampling method (random sampling) weight their answers (excl. the Netherlands)
- ✓ Half of the countries implementing quota sampling apply design weighting

Main characteristics of the weighting approaches

Applying weights for...

✓ **Selection bias reduction: the most significant reason for weighting**

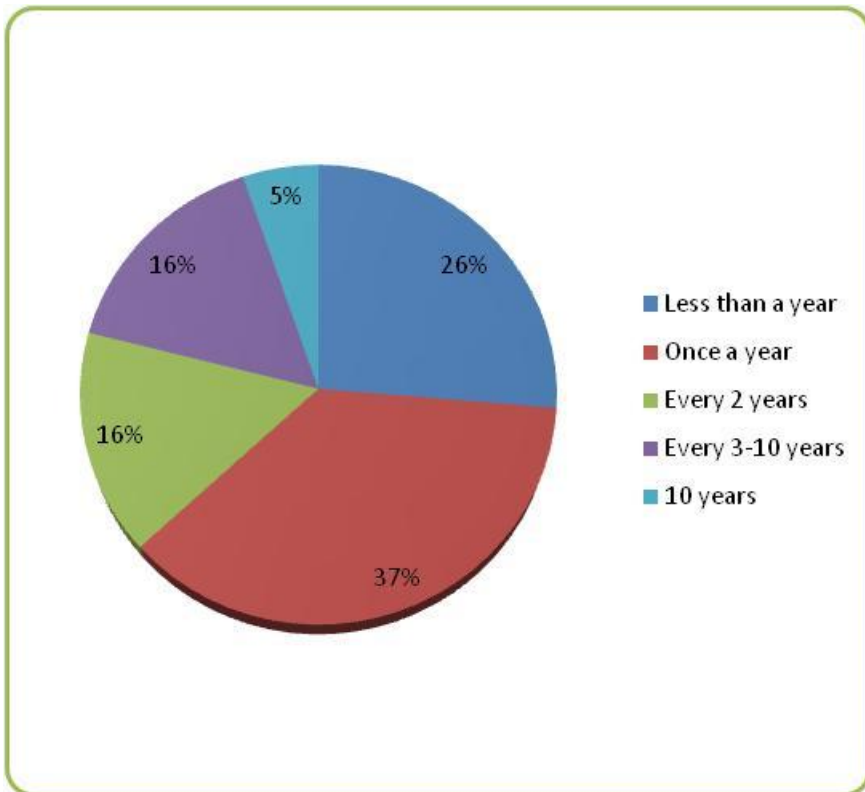
➤ 15 out of 24 countries that use weights stated that it is one of the reasons for weighting

✓ Weighting is used to a smaller extent for **variance reduction** (9 countries)

✓ The **change of the underlying population** seems to be a less crucial factor (referred by 3 countries)

Update frequency of weights

Figure 1: Update frequency distribution



Source: IOBE

✓ More than 1/3 of the countries applying weights (7 countries) update the weight coefficients once a year*

✓ 5 countries adjust them more frequently (each month or every quarter)

✓ Thus, the majority of the countries that follow a weighting approach (63%) update the weights at least once a year

*Information extracted from 19 countries applying weights that replied to IOBE's questionnaire in the context of the Task Force on the quality of the BCS data

3. Impact of weighting approaches on variance / tracking performance of the CCI

Introduction

Based on the definitions of the weighting approaches, the answers of the institutes to the IOBE questionnaire and DG ECFIN metadata, the following categorisation w.r.t. weighting approaches was applied:

- **Countries that weight responses** (non-response weighting, design weighting)
 - Only five countries implementing non-probability sampling and design weighting → difficult to handle for statistical inference
- **Countries that do not use a weighting approach** → small fraction of the partner countries (7 out of 31)
 - The vast majority of them uses non-probability sampling (excl. Italy, the Netherlands)

Introduction

- **Criteria to assess the impact of each weighting approach on volatility – compare their effects:**
 - Months for Cyclical Dominance index
 - Statistical testing for the significance of the difference in the volatility of the CCI among the different groups w.r.t. the weighting approach
- **Tracing of the potential linkages between different weighting approaches-CCI tracking performance:**
 - Correlation coefficient between the CCI and Household-NPISH Consumption expenditure (various time spans)

Effects on volatility – Evidence from the MCD index

- **Data from 29 countries were used (EU-28 & Turkey)**
- **The volatility of the Consumer Confidence Indicator is not high in all the countries that do not use weights, according to their MCD index values ($MCD \leq 3$)**
- **On the contrary, in more than 1/3 of the countries that assign weights to the survey responses, the variance of the CCI is high ($MCD > 3$)**

Table 1: Country distribution w.r.t. weighting approach – MCD level

	MCD		
	MCD= 1 or 2	MCD=3	MCD>3
No weighting	HU, SE	EE, IT, NL, SK, TR	
Weighting	DE, LT	BE, EL, LV, ES, FR, PT, FI, MT, AT, LU, RO, UK	DK, CY, PL, SI, BG, HR, CZ, IE

Effects on volatility – Evidence from the MCD index

- The **volatility** of the Consumer Confidence Indicator cannot be assessed as **either high or low in the majority of both** weighting and non-weighting countries. The respective **proportion** is **considerably higher** among the latter country group (72% Vs 55%)
- **However, MCD index median: 3 months in both country groups**
- **Overall, evidence** from the distribution of the MCD index **of lower volatility of the CCI in countries that do not apply weights**

Effects on volatility – Statistical testing

- **1st STEP:** Carry out a sample normality test, to determine which statistical test on the difference in the volatility of the CCI between the two country groups can be implemented
 - **Result:** The volatility data of the countries participating to the DG ECFIN Consumer Survey are drawn from a population that is not normally distributed (**Shapiro-Wilk test**)
- **2nd STEP:** Test for the statistical significance of the difference in the volatility → a non-parametric statistical test was used, the **Mann-Whitney U test**
 - The **P-value** of the test for the MCD level in the two country groups is **lower than 0.05** →
 - ❖ **The difference in the volatility of the Consumer Confidence Indicator between countries that use weights and these that do not is statistically significant**

Effects on volatility – Statistical testing

Based on the mean rank of each group, the volatility of the CCI in the non-weighting countries is lower

This result is in line with the standard error increase due to non-response weighting found in theory in case where:

- *The variables used as control totals are unrelated with the survey variables*
- *A small number of extremely large weights exists*

However, **this result must be treated with cautiousness** as the **sample of the non-weighting countries is very small (7 countries)**

The test results on potential sources of higher variance in countries applying weights will be presented after the effects of weighting on tracking performance

Effects on tracking performance

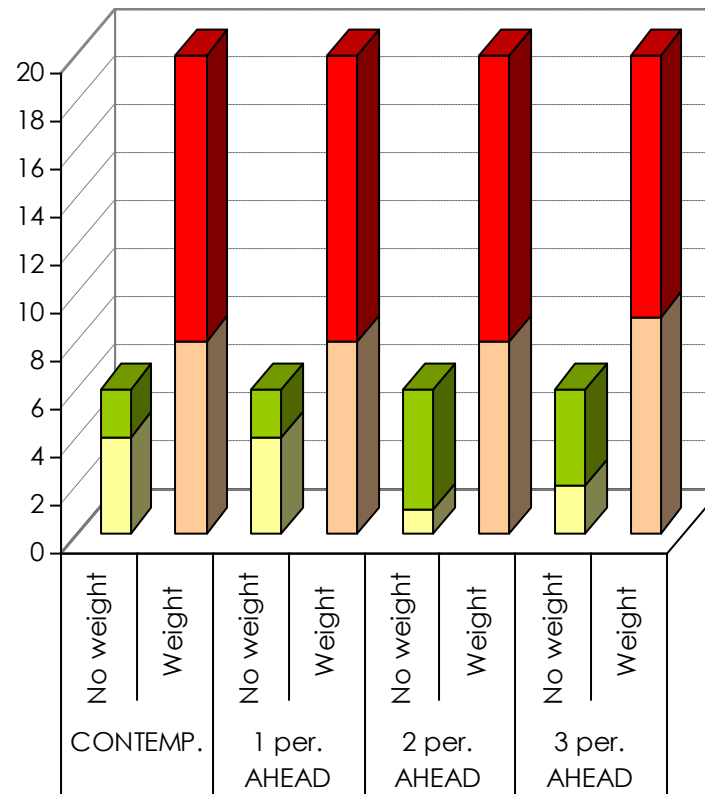
- **Data from 27 countries (EU-28 excl. Croatia)** on the correlation between the CCI and household-NPISH consumption for various time spans were used
- **Different critical levels of correlation for each time span:** As the CCI is mainly defined by expectations about the next quarter, higher critical values were set for one - two months ahead
- **Contemporary – One period ahead tracking performance:** Higher and stable proportion of countries with a high CCI-private consumption correlation among these that do not follow a weighting approach (66% Vs. 38% in countries applying weights)

Effects on tracking performance

- **Changes in the tracking performance two-three periods ahead:**
 - **Significant decrease** of the **non-weighting countries** with high correlation coefficient (\cong 20-30%)
 - **No signs of weakening tracking performance** from the **countries that assign weights** (two periods ahead). On the contrary, **as the time interval increases, the outcome is slightly better** (three periods ahead: 43% with a high correlation)
- **Conclusively, none of the two main weighting approaches constantly achieves a better predictability of the fluctuations** in the household-NPISH consumption
- These findings are summarised in the following figure:

Effects on tracking performance

Figure 2: Tracking performance of the CCI relative to the weighting approach (various time spans)



Lower strata: Countries with high correlation
Upper Strata: Countries with low Correlation

Source: IOBE

Impact of weighting features on the volatility of the Consumer Confidence Indicator

In order to **interpret the high volatility** in **countries using weights** we focused on the potential effects of some weighting features:

- a) **Population characteristics / strata** included in non-response and design weighting
- b) **Update frequency of the weights**

❖ Population characteristics used in the formation of weights

- **No common pattern among countries with a high MCD** as to how many and which characteristics are taken into consideration (table 2)
- As usual in population sample surveys, **gender, age group and region of residence** are mainly used. Thus:

A relation between **certain weighting factors and high volatility** cannot be established

Impact of weighting features on the volatility of the Consumer Confidence Indicator

Table 2: Population characteristics affecting the weight coefficients
(weighting countries with high volatility)*

Countries	Population Characteristics							
	Age	Gender	City size	Family size	Income	Education	Region	Other
Bulgaria	✓	✓	✓				✓	
Czech Republic	✓	✓	✓	✓		✓	✓	Economically active/ inactive
Croatia		✓				✓	✓	
Cyprus	✓	✓						
Denmark	✓	✓		✓	✓	✓	✓	Ethnicity & dwelling type
Poland	✓	✓	✓	✓			✓	
Slovenia	✓	✓	✓			✓	✓	

* No data availability for Ireland

Source: IOBE

Impact of weighting features on the volatility of the Consumer Confidence Indicator

❖ Updating frequency of the weights

- The **median of the update frequency** is **lower in countries** with high volatility (2 yrs. instead of 1 yr. in weighting countries with a small MCD index value)
- The **P-value** of the **Mann-Whitney U test** for the difference in the update frequencies of the two country subgroups is **lower than 0.10**
➔ The difference is statistically significant
- The **mean rank** of the **weighting countries with low volatility** does not exceed that of those with a high MCD ➔ **The update frequency is significantly higher among the former country group**

Impact of weighting features on the volatility of the Consumer Confidence Indicator

- Extension of the analysis: **Comparison** of the volatility of the **countries updating frequently their weights** (at least once a year) & of the **countries not using weights**
- **Difference in volatility not statistically significant** (P-value: 0.22)
- **Interpretation** of the previously found **higher variance in countries using weights**: Owed to the high MCD level of those that update the weights infrequently
- **In conclusion:**
 - **If the weight coefficients are frequently updated** (at least twice a year=average in the frequently updating countries) **then the weights are not a source of higher CCI volatility**
 - **Countries with high variance can limit it by updating the weights more regularly**

Impact of weighting features on the tracking performance of the Consumer Confidence Indicator

The examination of the impact of the **weighting features** was **extended** to the **tracking performance of the CCI**

❖ Population characteristics used in the formation of weights

- Almost **all the countries with a high CCI-private consumption correlation coefficient use weights based on the same characteristics** (gender, age group, region of residence, size of city of residence)
- These are the **most commonly used weighting factors** in population sample surveys. **Thus:**

No special combination of population characteristics for the formation of weights **in countries with a good tracking performance**

Impact of weighting features on the tracking performance of the Consumer Confidence Indicator

Table 3: Population characteristics affecting the weight coefficients
(weighting countries with good tracking performance)*

Countries	Population Characteristics							
	Age	Gender	City size	Family size	Income	Education	Region	Other
Bulgaria	✓	✓	✓				✓	
Greece	✓	✓	✓				✓	
Latvia	✓	✓					✓	Nationality
Lithuania	✓	✓						Living area
Portugal				✓	✓			
Romania	✓	✓	✓				✓	
Spain	✓	✓	✓				✓	

* No data availability for Ireland

Source: IOBE

Impact of weighting features on the tracking performance of the Consumer Confidence Indicator

❖ Updating frequency of the weights

- **Same update frequency median** in countries with a good tracking performance and with a low CCI-private consumption correlation (1 year)
- The **difference** in the update frequency among the two country subgroups is **not statistically significant**
- A **good tracking performance** of the fluctuations in the household-NPISH consumption by the CCI is **not a result of frequent update of the weights**

4. Main Findings-Suggestions

Main Findings-Suggestions

- **Weights** are used in the **vast majority of countries** participating to the DG ECFIN Consumer Survey (**24 out of 31**)
 - **19 countries** apply **non-response weighting** and **5 countries** implement **design weighting**
- Based on the **Months for Cyclical Dominance** index, the **volatility of the Consumer Confidence Indicator** was found to be **lower in countries that do not assign weights**
- **None** of the **two main country groups** w.r.t. the use of weights achieves a better **tracking performance of the CCI**:
 - **Non-weighting countries** perform better in terms of **one period ahead & contemporary CCI-consumption correlation**
 - The **correlation of the CCI** with changes in private consumption **two or three months ahead** is **higher in the countries that use weights**

Main Findings-Suggestions

Weighting features with a potential negative impact on the volatility of the Consumer Confidence Indicator

- A) **Population characteristics/strata** taken into account for the formation of the weights:
- **No common pattern** among the **countries with high MCD** as to **how many - which are used**
- B) **Update frequency of weights:**
- **Higher update frequency** in countries with a **low volatility using weights (median: 1 yr. Vs 2 yrs.)**. The **difference** from countries with high MCD is **statistically significant**
 - **Insignificant difference** in the volatility between countries not using weights and countries frequently updating the weights →

Frequent update of the weights(at least twice a year) will help to reduce the variance of the CCI

Both weighting features do not have an effect on the tracking performance in countries with high CCI – private consumption correlation

Thank you!

Michail Vassileiadis

+30-210-9211270, mvassiliadis@iobe.gr

my research associates

Fotini Thomaidou (thomaidou@iobe.gr)

Evangelia Valavanioti(valavanioti@iobe.gr)