

**TURKSTAT September 2021** 



#### Overview:

- Imputation Methods
- Data Control and Analysis
- Quality Adjustments
- Calculation of Elementary Indices
- Higher-level indices
- Calculation of Main Industrial Groupings (MIGs)
- Chain-linking
- Index Change Rates



#### Imputation Methods

Basic idea of imputation is to take the last available price and impute prices for the missing periods that trend with another index.

This other index could be an index of available prices for the general category of product or higher-level components of the PPI. (Based on CPA class)



#### Imputation Methods

Imputation methods can be partitioned in four broad classes:

- → 1. rule-based imputation
- 2. deductive imputation
- 3. model based imputation
- → 4. donor-based imputation



#### **Imputation Methods**

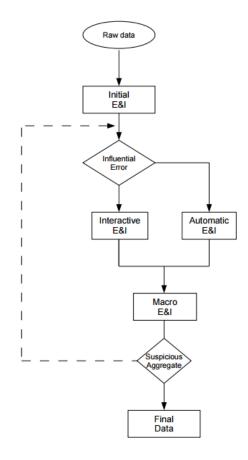
- Non-response rate is very low: ~1%
- Unit non-response and item non-response may occur
- ➡ Imputation at item level



#### Imputation Methods

- Apply growth rate for actual month for
  - 1. items under same company-CPA level industry and size group
  - If not available items under same CPA
  - If not available under same 4 and 3 digit level growth rate is applied
- Advantages:
  - **♦** Easy to apply
  - → Take into account trend changes among similar enterprises/products







- Some general rules should be followed when building an E&I strategy:
  - ◆ 1. identify and eliminate errors that are evident and easy to treat with sufficient reliability;
  - ◆ 2. select and treat with great care influential errors, carefully inspect influential observations; automatically treat the remaining non influential errors;
  - → 3. check the final output to see if there are influential errors undetected in the previous phases or introduced by the procedure itself.



- Four principles of the E&I process
  - 1. The data quality at the beginning and at the end of the E&I process must be assessed.
  - 2. The E&I process has to be designed and executed in a way that allows for control of the process.
  - 3. The data quality at the end of the process should satisfy the needs of the users.
  - 4. The process should be as simple, cheap and fast as possible.



- ➡ The error detection phase consists of identifying values that are not acceptable with respect to some pre-defined logical, mathematical or statistical criteria.
- ⇒ Error detection is often based on the use of edit rules. (CPA level based change rate, general change rate etc.)



- → 1. Missing values
- 2. Systematic errors
- ⇒ 3. Influential errors
- 4. Outliers
- ⇒ 5. Random errors



- ⇒ 1. Interactive Treatment
  - ➡ Contact with regional office responsible for data to review values
  - Contact with firm
- 2. Imputation
  - Missing values
  - Quality change



#### Data Control and Analysis Recommendations

- ➡ 1. Interactive treatment should be restricted to problems which cannot be solved by automatic treatment. Interactive treatment should be limited to the most relevant errors (e.g. influential errors and outliers).
- ⇒ 2. Interactive treatment should be performed by specialized and trained reviewers. Interactive treatment, and in particular manual imputation, must follow strict guidelines, which are designed, tested, tuned and monitored closely.
- ⇒ 3. Information from previous surveys or administrative data and access to the original questionnaire should be used when available.



#### Data Control and Analysis Recommendations

- ♣ 4. Number and length of call-backs to respondents should be kept minimal in order to reduce the burden on respondents and to avoid a negative impact on response rates. When doing call-backs, information on problems encountered by respondents should be collected to improve the survey.
- ⇒ 5. Interactive treatment and, in particular, changes to the data should be fully documented.



#### Data Control and Analysis Turkish Practice:

- Web based questionnaire (edit while doing the survey via program)
- Analyses of the Central Office



#### Data Control and Analysis Turkish Practice:

Web based questionnaire

- The questionnaire is filled by the enterprises via a web based format.
- Questionnaire is designed to check the consistency at several points and not to allow entering wrong or missing data.



#### Data Control and Analysis Turkish Practice:

Analyses of the Regional Offices

- Automated data analysis procedures
  - ⇒ SAS reports
  - Possible erroneous data is queried
  - Correction or validation of data



#### Data Control and Analysis Turkish Practice:

#### Analyses of the Central Office

- Automated and manual data analysis procedures
  - ➡ Harzemli analyze perspective
  - SAS reports
  - Deeper analyses for more important data
  - Analyses at index level
  - Possible erroneous data is queried and reported to Regional Offices
  - Correction or validation of data



#### **Outlier Identification**

▶ Lower and Upper Outer Fences (LOF and UOF, respectively) with k=3 are calculated:

$$LOF = Q_1 - k(Q_3 - Q_1)$$

$$UOF = Q_3 + k(Q_3 - Q_2)$$
with
$$k = 3$$

♠ Any observation outside of the outer fences is marked as problematic outlier.



#### **Outlier Identification**

- TERROR (TRAMO for errors) is an application to quality control of data
- ➡ In particular, to the detection of errors in reported indices data
- ➡ For each series, the program automatically identifies an ARIMA model and detects several types of outliers



#### **Quality Adjustments**

- The producer price index is intended to reflect pure price change over time and to do so by measuring prices for a constant quality.
- If there has been a change in either the representative product's specification or the sales conditions, the resulting observed price change (the price difference between the transactions before and after the change) may include a combination of a pure price change and a change in quality.
- As a price index must reflect pure price change, adjustments must be made for changes in the characteristics (or changes in quality) or sales conditions of individual goods and services. These adjustments are called quality adjustments.



#### **Quality Adjustments**

• The reported or observed price change ( $\Delta p$ ) can be sub-divided into two elements, that is a real (pure) price change ( $\Delta p_p$ ) and a quality change ( $\Delta p_q$ ), and can be portrayed by the following formula.

$$\Delta p = \Delta p_p + \Delta p_q$$

Where

 $\Delta p$  = observed/reported price change between two periods

 $\Delta p_p$  = real price change between two periods with comparable pricedetermining characteristics

 $\Delta p_q$  = apparent price change caused by a quality change between two periods



#### **Quality Adjustments**

- When selecting the representative products that will be observed, the products should be produced continuously over time to the same level of quality, as well as having a high degree of representativeness.
- The aim therefore is to identify products for which, where possible, no changes to the product specifications are anticipated in the future.
- In reality, markets are dynamic and continuously undergo structural changes for a variety of reasons.
- Products change or disappear from the market and completely new products are introduced. Selling conditions change and there are changes in the population of enterprises (such as insolvencies, start-ups, mergers and demergers) which may lead to different products entering the market.



#### **Quality Adjustments**

- These continuous quality changes must be identified and taken into account when calculating the index.
- Reporting units must note changes in the product specifications or sales conditions when reporting prices.
- When substantial price changes are observed, the reasons for the change should be investigated with the reporting units. Depending on the nature of the change, various procedures and possibilities to adjust for the quality change can be applied.
- Within the statistical office a decision should be taken, case by case, which quality adjustment procedure to implement. The objective for the staff of the statistical office is to determine the pure price change  $(\Delta p_p)$  between the old and new representative products.



#### **Quality Adjustments**

- Procedures if the old and new representative products exist in the market at the same time:
  - unadjusted price comparison (or direct price comparison or comparable replacement),
  - chaining (or link-to-show-no-price-change),
  - overlap pricing,
  - option prices.
- Procedures if the new and old representative products do not exist in the market at the same time:
  - expert estimation,
  - matched model approach,
  - hedonic approach.



#### **Overlap Pricing**

- In the method of overlap pricing, the old and new representative products overlap
  in the market, meaning that there is a price for both representative products in a
  period.
- As such, the price difference observed in the market can be assumed to be the quality difference. A comparable previous month's price can therefore be directly observed.
- The advantage of this method lies in the fact that the price difference can be subdivided by simple calculation into the elements of the quality change and pure price change.
- An attempt should therefore always be made to overlap the price reports in the event of a change in the quality. Frequently this may require further contact with the reporting unit as generally only the current price for the representative product being observed is reported.



#### **Overlap Pricing**

Example of overlap pricing

	Date of observation			
	t-2	t – 1	t	t + 1
Product m (old)	10	$10 $ $\Delta p_q = 9$	_	_
Product m (new)	18	19 APq	20	21
		$\Delta p_p$	= 1	
Index	110.0	110.0 (110.0)	→ 115.8	121.6

The calculation of  $\Delta p_p$  for the product representative m is the following:  $\Delta p_p = P_{m(new)}^t - P_{m(new)}^{t-1}$ 

Δpp= real price change

 $P_{m(new)}^{t}$  = price of new product representative in the current month

 $P_{m(new)}^{t-1}$  = price of new product representative in the previous month

The calculation of  $\Delta p_q$  is the following:  $\Delta p_q = P_{m(new)}^{t-1} - P_{m(old)}^{t-1}$ 

 $\Delta p_{\text{q}}\text{=}$  quality change between old and new representative products

 $P_{m(new)}^{t-1}$  = price of new product representative in the previous month

 $P_{m(old)}^{t-1}$  = price of old product representative in the previous month

27



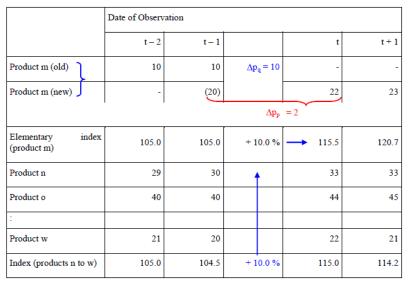
#### Matched model - approach

- If the old product representative is replaced by one with different quality and it is not possible to observe both products in the same period, a matched model approach may be used. In order to calculate the pure price change of the new product representative, the price developments of other, similar representative products are used.
- Depending on the type of product and the number of price observations in the elementary aggregate, these may be similar representative products within the same detailed level of the product classification or from a higher aggregate.
- It is presumed that the average percentage change of the selected other representative products can serve as a reference for the price change of the new representative product. This presupposes an ideal market in which all producers experience similar price changes.



#### Matched model - approach

Example of matched model approach



The calculation of  $\Delta p_p$  is the following:

$$\Delta p_p = P_{m(new)}^t - P_{m(new)}^t \times \frac{I_k^{t-1}}{I_k^t}$$

 $\Delta p_p$ = real price change

 $P_{m(new)}^{t}$  = price of new product representative in the current month

t-1 = index of comparable representative products in the previous month

 $\int_{\mathbf{k}}^{\mathbf{t}}$  = index of comparable representative products in the current month



#### Calculation of Elementary Indices

A PPI typically proceeds in two or more stages. In the first stage, elementary price indices are estimated for the elementary aggregates of a PPI.

In the second stage, these elementary indices are combined to obtain higher-level indices using the elementary aggregate indices with revenue weights.

An elementary aggregate consists of the revenue for a small and relatively homogeneous set of products defined within the industrial classification used in the PPI. (CPA 2008)



#### Calculation of Elementary Indices

The aggregates covered by a CPI or a PPI usually are arranged in the form of a tree-like hierarchy

Weighted arithmetic aggregates are used in Turkish practice



#### Higher-level Indices

- These elementary indices are combined to obtain higher-level indices using the elementary aggregate indices with revenue weights.
- Weights based on CPA hierarchy
  - sections,
  - divisions,
  - groups,
  - classes,
  - categories
  - subcategories.



#### Calculation of Main Industrial Groupings (MIGs)

- The **main industrial groupings**, abbreviated as MIGs, provide an alternative statistical breakdown of the economic activities of **industry**, as compared to the sectoral breakdown of the Statistical classification of economic activities in the European Community (NACE).
- Intermediate Goods
- Durable Goods
- Non Durable Goods
- Capital Goods
- Energy



#### Chain-linking

Chain-linking means constructing long-run price or volume measures by cumulating movements in short-term indices with different base periods.

Chain-linked index can be constructed by multiplying a series of short-term indices measuring the change from one period to the next as follows:

$$CI_{0 \to t} = I_{0 \to 1} \times I_{1 \to 2} \times I_{2 \to 3} \times I_{3 \to 4} \times \cdots \times I_{t-1 \to t} = \prod_{\tau=1}^{t} I_{\tau-1 \to \tau}$$

Chain-linked indices do not have a particular base or weight period. Each link of the chain-linked index has a base period and weight period, and the base and weight period are changing from link to link



#### Chain-linking

Annual chain-linking of monthly/quarterly data implies that each link in the chain is constructed using the chosen index number formula with the average of the previous year (t-1) as base and reference period. The resulting short-term monthly/quarterly indices must subsequently be linked to form long, consistent time series expressed on a fixed reference period.

The chain-linked Laspeyres type volume measures should be derived by compiling monthly/quarterly estimates at the average prices of the previous year.



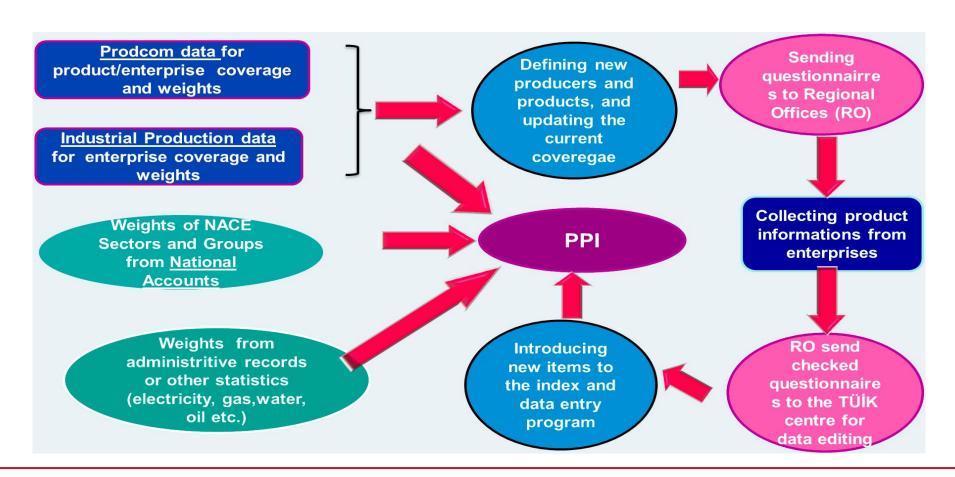
#### Chain-linking

Short-term link in annually chain-linked Laspeyres

$$LP(\overline{t-1}) \rightarrow (m,t) = \frac{\sum_{i} \overline{q}_{i,t-1} \times p_{i,m,t}}{\sum_{i} \overline{q}_{i,t-1} \times \overline{p}_{i,t-1}}$$



**Chain-linking Annual Update** 





#### **Index Change Rates**

- Monthly change (%)
- Rate of change on December of the previous year (%)
- Annual change (%)
- Rate of change in twelve months moving averages (%)



# Thank You.

39