# STATISTICAL METHODS WITH APPLICATION TO FINANCE

a.y. 2023-2024

## Index numbers and Stock Market indices

R. PAPPADÀ
rpappada@units.it

Department of Economics, Business, Mathematics and Statistics "B. de Finetti"
University of Trieste

March 3, 2024

# Index numbers

One of the most important uses of statistics is to *summarize data* and to *make comparisons*.

A great attention is often paid to how business, economic, and financial variables change with time or place:

variation in gross national production, variation in the price of consumer goods, changes in stock market prices, salary increase for a company that wishes to transfer an executive to another city.

An **index number** is a summary measure that compares related items over time or place.

Cheng-Few Lee, John C. Lee, Alice C. Lee (2013) Statistics for Business and Financial Economics, Third Edition. Springer.

# Index numbers

Index numbers are used in business and finance to express the level of an activity or phenomenon in relation to its level at another time or place:

- one way to compare price changes in two different periods is by constructing an **index of price**
- we can compare the performance of our stocks to the rest of the stock market in the same time period by using an **index of stock prices**

# Market indices

A stock market index is a statistical measure that shows how the prices of a group of stocks change over time. It encompasses either all or only a portion of stocks in its market

- Examples are the Dow Jones Industrial Average (DJIA), the NASDAQ, the Dax, the FTSE 100, Nikkei, etc.
- Stock market indices are often viewed as short-run indicators of the changing economic and political conditions affecting the market (e.g., the Standard & Poor 500 for the US market)
- Most widely known indices are related to a geographic area, or a given industry sector, but some indices, like MSCI (Morgan Stanley Capital International), refer to world markets

# Market indices

Stock market indices are constructed as **index numbers** with different weighting schemes resulting into

- Price-Weighted Indices
  - *Just like a simple arithmetic average*
  - *example: the Dow Jones Industrial Average (DJIA) Index*
  - *The constituent stocks are weighted based on their price*
- Market-Value-Weighted Indices
  - *weighted aggregative price indices*
  - *example: the Standard & Poor's 500 Composite Index*
  - *the portfolio reflects the actual market capitalization of each firm*

# Table of Contents

# Simple aggregative price index

The **simple aggregative price index** is defined as

$$I_t = \frac{\sum_{i=1}^{n} P_{ti}}{\sum_{i=1}^{n} P_{0i}} \times 100 \quad (1)$$

where

$P_{ti}$ : price of the $i$th commodity in the year $t$
$P_{0i}$ : price of the $i$th commodity in the base year $0$

## Limitations

- The units used can affect the price index (to see this, compute the same indices when the price of eggs is stated in half dozens!)
- The relative importance of the commodities is not taken into account (great influence of price variation of high-priced commodities)

# Simple aggregative price index: example

**Example 1:** Consider a combination of several goods. Assume that the quantity for the $i$-th good is $Q_{0i} = 1$, for all $i$.

| Commodity | Quantity | 1989 $P_0$ | 1990 $P_1$ | 1991 $P_2$ |
|---|---|---|---|---|
| Eggs | One dozen | 1.00 | 1.20 | 1.50 |
| Milk | One gallon | 1.50 | 1.75 | 2.00 |
| Butter | One pound | 1.10 | 1.35 | 1.60 |
| Bread | One loaf | 0.40 | 0.70 | 0.90 |
| | | 4.00 | 5.00 | 6.00 |

$$\text{base year } \rightarrow 1989 : I_{89} = 4.00/4.00 \times 100 = 100$$
$$1990 : I_{90} = 5.00/4.00 \times 100 = 125$$
$$1991 : I_{91} = 6.00/4.00 \times 100 = 150$$

The cost of this specific list of goods increased $25\%$ between 1989 and 1990 and $50\%$ between 1989 and 1991.

# Simple relative price index

We can improve on our index by taking the average of the price relatives:

$$I_t = \frac{1}{n} \sum_{i=1}^{n} \frac{P_{ti}}{P_{0i}} \times 100 \tag{2}$$

For the data in Example 1 we have:

| Commodity | 1989 $P_0/P_0$ | 1990 $P_1/P_0$ | 1991 $P_2/P_0$ |
|---|---|---|---|
| Eggs | 1 | 1.20/1.00=1.20 | 1.50/1.00=1.50 |
| Milk | 1 | 1.75/1.50=1.17 | 2.00/1.50=1.33 |
| Butter | 1 | 1.35/1.10=1.23 | 1.60/1.10=1.45 |
| Bread | 1 | 0.70/0.40=1.75 | 0.90/0.40=2.25 |
| *average* | | 5.35/4=1.34 | 6.53/4=1.63 |
| | $I_{89} = 100$ | $I_{90} = 134$ | $I_{91} = 163$ |

However, each item in the index is weighted by $1/P_0$, which makes all items equally important!

# Weighted aggregative price index

**Weighted aggregative price indices** look at both the prices and quantities, in order to reflect the value of some commodities in relation to the value of others:

- the *value* of a commodity is determined by its **price and the quantity** purchased; for the $i$-th commodity in the base year

$$W_{0i} = P_{0i}\, Q_{0i}, \quad i = 1, \ldots, n$$

  where $P_{0i}$ is the price per unit and $Q_{0i}$ is quantity;
- the weighted relative for the $i$-th commodity in period $t$ is

$$\left(\frac{P_{ti}}{P_{0i}}\right) W_{0i} = \left(\frac{P_{ti}}{P_{0i}}\right) P_{0i}\, Q_{0i} = P_{ti}\, Q_{0i}$$

# Weighted aggregative price index: Example

For the data in Example 1, consider prices and quantities for the year 1989 (base year) and compute the weights $W_{0i}$ for each commodity.

| $i$ | $P_{0i}$ | $Q_{0i}$ | $W_{0i} = P_{0i}Q_{0i}$ | $P_{1i}/P_{0i}$ (1990) | $P_{2i}/P_{0i}$ (1991) |
|---|---|---|---|---|---|
| 1 | 1.00 | 150 | 150 | 1.20 | 1.50 |
| 2 | 1.50 | 300 | 450 | 1.17 | 1.33 |
| 3 | 1.10 | 200 | 220 | 1.23 | 1.45 |
| 4 | 0.40 | 1100 | 440 | 1.75 | 2.25 |

Compute the weighted price index for 1990

$$I_1 = \frac{(1.20)150 + (1.17)450 + (1.23)220 + (1.75)440}{150 + 450 + 220 + 440} \times 100 = 138.7$$

and for 1991

$$I_2 = \frac{(1.50)150 + (1.33)450 + (1.45)220 + (2.25)440}{150 + 450 + 220 + 440} \times 100 = 169.2$$

# Weighted aggregative price index: Example

For the data in Example 1, consider prices and quantities for the year 1989 (base year) and compute the weights $W_{0i}$ for each commodity.

| $i$ | $P_{0i}$ | $Q_{0i}$ | $W_{0i} = P_{0i}Q_{0i}$ | $P_{1i}/P_{0i}$ (1990) | $P_{2i}/P_{0i}$ (1991) |
|---|---|---|---|---|---|
| 1 | 1.00 | 150 | 150 | 1.20 | 1.50 |
| 2 | 1.50 | 300 | 450 | 1.17 | 1.33 |
| 3 | 1.10 | 200 | 220 | 1.23 | 1.45 |
| 4 | 0.40 | 1100 | 440 | 1.75 | 2.25 |

Compute the weighted price index for 1990

$$I_1 = \frac{(1.20)150 + (1.17)450 + (1.23)220 + (1.75)440}{150 + 450 + 220 + 440} \times 100 = 138.7$$

and for 1991

$$I_2 = \frac{(1.50)150 + (1.33)450 + (1.45)220 + (2.25)440}{150 + 450 + 220 + 440} \times 100 = 169.2$$

# Weighted aggregative price index: Example

For the data in Example 1, consider prices and quantities for the year 1989 (base year) and compute the weights $W_{0i}$ for each commodity.

| $i$ | $P_{0i}$ | $Q_{0i}$ | $W_{0i} = P_{0i}Q_{0i}$ | $P_{1i}/P_{0i}$ (1990) | $P_{2i}/P_{0i}$ (1991) |
|---|---|---|---|---|---|
| 1 | 1.00 | 150 | 150 | 1.20 | 1.50 |
| 2 | 1.50 | 300 | 450 | 1.17 | 1.33 |
| 3 | 1.10 | 200 | 220 | 1.23 | 1.45 |
| 4 | 0.40 | 1100 | 440 | 1.75 | 2.25 |

Compute the weighted price index for 1990

$$I_1 = \frac{(1.20)150 + (1.17)450 + (1.23)220 + (1.75)440}{150 + 450 + 220 + 440} \times 100 = 138.7$$

and for 1991

$$I_2 = \frac{(1.50)150 + (1.33)450 + (1.45)220 + (2.25)440}{150 + 450 + 220 + 440} \times 100 = 169.2$$

# Weighted aggregative price index

In the example, we found that the *weighted aggregative price indices* computed from the four commodities for 1990 and 1991 are about 39% and 69% higher than those of 1989, respectively (because 1989 is the base year, $I_0$ must equal 100).

For a basket of $n$ commodities and base year $t = 0$, the **weighted aggregative price index** in period $t$ is

$$I_t = \frac{\sum_{i=1}^{n} \left(\frac{P_{ti}}{P_{0i}}\right) W_{0i}}{\sum_i W_{0i}} \times 100 \qquad (3)$$

Such indices differ from the ones computed using Eq. (2) because the relative importance of each good in the basket is now taken into account.

# Table of Contents

# Introduction

Weighted aggregative price indices look at both the prices and quantities:

- The Laspeyres price index uses base-year quantities.
- The Paasche price index uses current-year quantities.
- Fisher's ideal price index is a geometric average of the Laspeyres and Paasche indices.

We can construct analogous quantity indices to express how the quantity of goods purchased has changed, looking at both the prices and quantities.

# The Laspeyres Price Index

A weighted aggregative price index has the form

$$I_t = \frac{\sum_{i=1}^{n} \left(\frac{P_{ti}}{P_{0i}}\right) W_{0i}}{\sum_i W_{0i}} = \frac{\sum_{i=1}^{n} \left(\frac{P_{ti}}{P_{0i}}\right) P_{0i}Q_{0i}}{\sum_i P_{0i}Q_{0i}}$$

Thus, the index can be rewritten as

$$I_t^L = \frac{\sum_{i=1}^{n} P_{ti}Q_{0i}}{\sum_{i=1}^{n} P_{0i}Q_{0i}} \times 100 \tag{4}$$

This formula is referred to as the **Laspeyres price index**.

- the numerator is the total amount of money in year $t$ that is required to buy the same quantity purchased in the base year;
- the denominator is the total cost of purchasing from the base year

# The Laspeyres Price Index: pros and cons

The Laspeyres Price Index:

- can be used to reflect price changes between two periods
- is easy to apply since you would only need current prices for updating

One disadvantage is that the Laspeyres index tends to give more weight to those items that show a dramatic price increase, usually followed by a decrease in the demand (measured by $Q$)

The *Consumer Price Index* (CPI) is computed as a Laspeyres index.

# The Paasche Price Index

An alternative is to use *current-year* quantities ($Q_t$). The **Paasche price index** is computed as

$$I_t^P = \frac{\sum_{i=1}^{n} P_{ti}\, Q_{ti}}{\sum_{i=1}^{n} P_{0i}\, Q_{ti}} \times 100 \tag{5}$$

It expresses *what today's basket would have cost at yesterday's prices*. It provides a realistic and up-to-date estimate of total expense. On the other hand

- it is difficult to apply due to the complexity of updating the reference-year quantities
- we cannot use it to reflect price changes between two periods

# Fisher's ideal price index

**Fisher's ideal price index** represents a compromise between the Laspeyres price index and the Paasche price index. Formally, it is defined as the geometric average of the Laspeyres and Paasche indices:

$$I_t^F = \sqrt{I_t^L \times I_t^P} \tag{6}$$

Note that the Fisher index is larger than the Paasche index and smaller than the Laspeyres index.

# Laspeyres, Paasche and Fisher indices: Example

**Example 2** *Prices and quantities for three commodities in 2012 and 2013.*

| Commodity | 2012 | | 2013 | |
|---|---|---|---|---|
| | Price (€) | Quantity | Price (€) | Quantity |
| Shirts | 22.00 | 6 | 19.00 | 8 |
| Shoes | 71.00 | 2 | 78.00 | 3 |
| Dresses | 53.00 | 8 | 62.00 | 7 |

Compute the Laspeyres price index, the Paasche price index and the Fisher index using 2012 as the base year.

1. Laspeyres price index

$$\frac{\sum_{i=1}^{3} P_{ti}Q_{0i}}{\sum_{i=1}^{3} P_{0i}Q_{0i}} = \frac{19 \times 6 + 78 \times 2 + 62 \times 8}{22 \times 6 + 71 \times 2 + 53 \times 8} = 1.097 \rightarrow I_t^L = 109.7$$

# Laspeyres, Paasche and Fisher indices: Example / 2

2. Paasche price index

$$\frac{\sum_{i=1}^{3} P_{ti} Q_{ti}}{\sum_{i=1}^{3} P_{0i} Q_{ti}} = \frac{19 \times 8 + 78 \times 3 + 62 \times 7}{22 \times 8 + 71 \times 3 + 53 \times 7} = 1.079 \rightarrow I_t^P = 107.9$$

3. Fisher's Ideal Price Index

$$I_t^F = \sqrt{I_t^L \times I_t^P} = \sqrt{109.7 \times 107.9} = 108.79$$

# Table of Contents

# Quantity Indices

A **quantity index** measures a change in quantity from a base year to a particular year (e.g. the volume of industrial production, the physical volume of imports and exports)

For instance, the Laspeyres quantity index is

$$I_t^{LQ} = \frac{\sum_{i=1}^{n} Q_{ti} P_{0i}}{\sum_{i=1}^{n} Q_{0i} P_{0i}} \times 100 \tag{7}$$

- it represents the total cost of the quantities in period $t$ at base-year prices as a percentage of the total cost of the base-year quantities
- any change in the index is due to the change in quantities between the base year and the year of interest

# Quantity Indices: Example

For the data in Example 2, compute the Laspeyres quantity index, the Paasche quantity index, and Fisher's ideal quantity index, using 2012 as the base year.

1. Laspeyres quantity index

$$\frac{\sum_{i=1}^{3} Q_{ti}P_{0i}}{\sum_{i=1}^{3} Q_{0i}P_{0i}} = \frac{8 \times 22 + 3 \times 71 + 7 \times 53}{22 \times 6 + 71 \times 2 + 53 \times 8} = 1.0888 \rightarrow I_t^L = 108.88$$

2. Paasche quantity index

$$\frac{\sum_{i=1}^{3} Q_{ti}P_{ti}}{\sum_{i=1}^{3} Q_{0i}P_{ti}} = \frac{8 \times 19 + 3 \times 78 + 7 \times 62}{6 \times 19 + 2 \times 78 + 8 \times 62} = 1.0705 \rightarrow I_t^P = 107.05$$

3. Fisher's Ideal quantity index

$$I_t^F = \sqrt{I_t^L \times I_t^P} = \sqrt{108.88 \times 107.05} = 107.96$$

# Table of Contents

# Price-weighted indices

A price-weighted index shows the change in the average price of the stocks that are included in the index.

Price weighting itself causes high-priced stocks to dominate the series.

Hence, such indices will rarely reflect the evolution of underlying market values because the $100 stock might be that of a small company and the $10 stock that of a large company.

# Market-Value-Weighted indices

In a market-value-weighted index, the importance of each stock in the index depends on the market-value of its company, which is

$$\textit{share price} \times \textit{number of shares outstanding}$$

- securities with the highest capital value have the greatest influence over the values of price-weighted indices
- the portfolio reflects the actual market capitalization of each firm

# Example

**Example 3** *Prices ($P_0$, $P_1$) for three stocks A, B, and C and the number of shares ($Q_0$, $Q_1$) at base year $t_0$ and in next period.*

| Stock | $P_0$ | $Q_0$ | $P_1$ | $Q_1$ |
|---|---|---|---|---|
| A | 27 | 47 | 32 | 50 |
| B | 18 | 100 | 16 | 105 |
| C | 92 | 72 | 90 | 88 |

We can compute the price-weighted index as a simple aggregative price index:

$$I_{t_1} = \frac{32+16+90}{27+18+92} \times 100 = 100.73$$

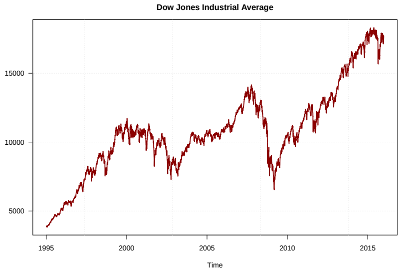
We calculate the market-value-weighted index as follows:

$$\begin{aligned} I_{t_1} =& \frac{P_{1A}Q_{1A} + P_{1B}Q_{1B} + P_{1C}Q_{1C}}{P_{0A}Q_{0A} + P_{0B}Q_{0B} + P_{0C}Q_{0C}} \times 100 \\ =& \frac{32(50) + 16(105) + 90(88)}{27(47) + 18(100) + 92(72)} \times 100 = 115.55 \end{aligned}$$

# Table of Contents

# The Dow Jones Industrial Average

The *Dow Jones Industrial Average* is probably the best known example of a price-weighted stock index, including 30 large, well-known, companies traded on the NASDAQ or the NYSE




However, the DJIA is not the best indicator of the overall market performance since it includes only 30 stocks, and the use of a price-weighted index gives an advantage to some DJIA components over others.

# How the DJIA works?

Consider a set of $k$ stock share prices $P_{ti}$ ($i = 1, \dots, k$) and define an index

$$I_t = \frac{1}{d_t} \sum_{i=1}^{k} w_i P_{ti} \tag{8}$$

for a given set of **weights** $w_i$, and a **divisor** $d_t$.

- The divisor is initially chosen in such a way that the resulting index assumes a ‘nice’ value, say, 100 or 1000.
- The divisor is changed when the index composition is changed to reflect new market conditions (e.g., substitutions, *splits*)

For the DJIA, weights in Eq.(8) are all $w_i = 1$. The DJIA divisor has changed significantly over the years (e.g., it was at 16.67 back in 1928, 0.147 in Sept. 2019).

# Index adjustments: example

Let us consider how to manage an index for a stock market on which four stocks (A, B, C, D) are traded. The 2-for-1 stock split by stock A is shown.

| | Before split | After 2-for-1 split | Adjustment |
|---|---|---|---|
| Stock A | 60 | 30 | 30 |
| Sum of prices (A+B+C+D) | 120 | 90 | 90 |
| Value of the index | 120/4=30 | 22.5 | 30 |
| Divisor $d$ | 4 | 4 | 3 |

1. The initial divisor for the price-weighted index is $d = 4$
2. The new divisor is changed in order to reflect the split without introducing a discontinuity in the index

$$\frac{90}{d'} = 30 \rightarrow d' = 3$$

# Criticisms of the Dow Jones Industrial Average

Two major shortcomings are:

- A less than 1% representation of the total stock market may be misleading and may not represent the actual state of the economy
- Its construction implies that, for example, a component with a share price of $\$120$ would influence the DJIA more than four times a company with a stock price of $\$30$ even though the $\$30$ stock price company may be more important to the economy

# Table of Contents

# The Standard and Poor's 500

The Standard & Poor's (S&P 500) index, which comprises industrial firms, utilities, transportation firms, and financial firms, is an example of a market-value-weighted index, also called a capitalization-weighted index


Figure 1: The Standard and Poor's 500, weekly closing values (1995–2015). *Source: Yahoo! Finance*

# How is the S&P 500 Index calculated?

The S&P 500 Index is also said to be *float-adjusted*, that is, the number of shares used for calculation is the number of shares *floating*, rather than outstanding. It is computed as

$$I_t = \frac{\sum_{i=1}^{k} P_{ti} Q_{ti}}{d_t} \tag{9}$$

where the *divisor* $d_t$ is adjusted to ensure the continuity of the index after changes in the composition of the stocks

Remark: If a 2-for-1 stock split occurs, a market-value-weighted index would not be affected

# Example

Consider the following scenario with 5 stocks, for given prices and outstanding shares at base period (0) and on the next day:

| Stock | $P_0$ | $Q_0$ | $P_0Q_0$ | $P_1$ | $Q_1$ | $P_1Q_1$ |
|---|---|---|---|---|---|---|
| A | \$3 | 50 | 150 | \$4 | 50 | 200 |
| B | \$1 | 50 | 50 | \$1 | 50 | 50 |
| C | \$7 | 70 | 490 | \$7 | 70 | 490 |
| D | \$9 | 20 | 180 | \$9 | 20 | 180 |
| E | \$10 | 10 | 100 | \$9 | 10 | 90 |
| total market value | | | 970 | | | 1010 |

Let's assume we want our index's starting value to be 100:

$$\frac{970}{d} = 100 \rightarrow d = 9.7$$

The current index value (t=1) is then

$$\text{Index value} = \text{total market value/d} = \frac{1010}{9.7} = 104.1$$

The index has risen in value by $4.1\%$ in a day.

# Index Return

The return on a stock index is calculated as follows:

$$R_t = \frac{I_t - I_{t-1}}{I_{t-1}} \tag{10}$$

where

- $R_t$ is the index return for time $t$
- $I_t$ is the index level at time $t$
- $I_{t-1}$ is the index level at end of the previous period (time $t-1$)