Topic 3:Index Numbers and Times Series

Part I: Index Numbers:

- What is an _____ number?
- Constructing Indices
- Choosing between Indices
- Using Price Indices
- The Canadian ___
- Price Indices and the "Cost of Living"

Reference: Chapter 16

⇨We need to develop a measure that summarizes the characteristics of _____ data sets and is useful for period to period comparisons.

⇨An **I_____number** accomplishes this by aggregating information into a single measure that permits easy comparisons.

Definition: An INDEX NUMBER is a number which summarizes pieces of information in a ________ way, and forms the basis for comparative judgements.

⌘ An index number is a ratio of 2 numbers expressed as a percentage.

Used if we want to keep track of "_______" price changes over time or need to get an **aggregate** measure of output– different items have the own units.

Examples: CPI (Consumer Price Index);
PPI (Producer Price Index)

Basic Considerations for a "good" Index Number:

(i) Include information on___ items.

(ii) Combine in way which recognizes **r______** importance of each item.

(iii) Need a "**bench mark**" value and period to provide reference point for comparisons. ***Base period required***

Price and Quantity Indices

Price Indices

Two broad types: (and there are several variants of each)

(i) A__________ Indices

(ii) R________ Indices

Looking at (i): Aggregative Indices

Example: Bundle of Two Goods

Item	1990			2010	
	Price ($) (P)	Quantity (Q)		Price ($) (P)	Quantity (Q)
Good 1					
Good 2					

We want to summarize overall price change from 1990 to 2010. (Base Year =1990.)

Must be the **same goods** in each year! (Assumption)

Type of Price Index

Price of the "i^th" item in the year of interest.

$$P_{0t}^{S} = \left[\frac{\sum_{i=1}^{n} P_{it}}{\sum_{i=1}^{n} P_{i0}} \right] \times 100$$

The 100 indicates that we have a base value of 100 in the base year.

Base Year

Year of interest

Price of the "i^th" item in the base year.

<u>**S_____ Aggregative Index:**</u>

Let P_{it}= Price of the i^{th} item at time "t".
(i= 1, 2,....,n; t=0, 1, 2,....,T)

Construction of the <u>**Simple Price Index**</u>:

$$P_{0t}^{S} = \left[\frac{\sum_{i=1}^{n} P_{it}}{\sum_{i=1}^{n} P_{i0}} \right] * 100$$

The simple price index from base year '0' to year 't' for items 1 to n.

If t=0, index=100 (Base year value).

<u>Example</u>: See previous data for 1990 and 2010:
Base Year =1990

$$P_{0t}^{S} = \left[\left[\frac{\sum_{i=1}^{n} P_{it}}{\sum_{i=1}^{n} P_{i0}}\right]\right] \times 100$$

$$P_{0=1980,t=2000}^{S} = \left[\frac{1.80 + 2.70}{0.75 + 0.50}\right] \times 100 = 360 \Leftarrow {}^{\$}\!/_{\$} \text{ Unitless}$$

<u>Prices rose, overall, to be ___ times higher in 2010 than in 1990.</u>

⇨This type of index ***only*** tells about changes in ______.

⇨Individual ____ falls can be offset by other individual rises.

⇨With the simple index, all goods are given **equal ______**.

So......

<u>W_________ Aggregative Index:</u>

$$P_{0t}^{W} = \left[\frac{\sum_{i=1}^{n} P_{it} W_i}{\sum_{i=1}^{n} P_{i0} W_i} \right] \times 100$$

where W_i=(fixed) weight given to the i^{th} price.

(Limitation of these indices: quality changes: **Assumption of <u>identical</u> items.)**

How to Choose the Weights:

- Q________ of goods purchased. (Which quantity?)
- Percentage of _____________ on good. (p_iq_i)

Both measure "importance", but may *differ* over time.

So, when should they be measured?

Possibilities:

(I) L_________' Index:

$$P_{0t}^{L} = \left[\frac{\sum_{i=1}^{n} p_{it} q_{i0}}{\sum_{i=1}^{n} p_{i0} q_{i0}} \right] \times 100$$

Base period quantities= weights

Disadvantage:

W______ get ***outdated*** over time.

Tends to **overstate** changes in price for the price index.

(II) P________s Index:

$$P_{0t}^{P} = \left[\frac{\sum_{i=1}^{n} p_{it} q_{it}}{\sum_{i=1}^{n} p_{i0} q_{it}} \right] \times 100$$

_______ period quantities=weights

Disadvantage:

Weights constantly changing, so index change may be partly due to quantity changes. (Tends to __________ changes.)

Some Compromises:

(III) M________-Edgeworth Index:

$$P_{0t}^{ME} = \left[\frac{\sum_{i=1}^{n} p_{it} \left(\frac{q_{i0} + q_{it}}{2} \right)}{\sum_{i=1}^{n} p_{i0} \left(\frac{q_{i0} + q_{it}}{2} \right)} \right] \times 100$$

Average of base period and current period ___________ are the weights.

(IV) F_____'s "Ideal" Index:

$$P_{0t}^{F} = \sqrt{P_{0t}^{L} \times P_{0t}^{P}}$$

(Geometric Mean of the L________' and P______ indices)

Example: Calculations from earlier data:

Year	P_{0t}^{S}	P_{0t}^{L}	P_{0t}^{P}	P_{0t}^{ME}	P_{0t}^{F}
1990					
2010	360	458.18	390.1	422.61	422.7

Calculations:(from data on page 2)

If the base year is 1990, with a value of 100, then for t=2010:

Simple Aggregative Index

$$P_{0t}^{S} = \frac{1.80 + 2.70}{0.75 + 0.50} \times 100 = 360$$

Laspeyres':

$$P_{0t}^{L} = \frac{1.80(1) + 2.70(4)}{0.75(1) + 0.50(4)} \times 100 = 458.18$$

Paasche: $$P_{0t}^{P} = \frac{1.80(2)+2.70(3)}{0.75(2)+0.50(3)} \times 100 = 390.1$$

Marshall-Edgeworth:

$$P_{0t}^{ME} = \left[\left[\frac{\sum_{i=1}^{n} p_{it}\left(\frac{q_{i0}+q_{it}}{2}\right)}{\sum_{i=1}^{n} p_{i0}\left(\frac{q_{i0}+q_{it}}{2}\right)}\right]\right] \times 100$$

$$= \left[\frac{1.80\left(\frac{1+2}{2}\right)+2.70\left(\frac{4+3}{2}\right)}{0.75\left(\frac{1+2}{2}\right)+0.50\left(\frac{4+3}{2}\right)}\right] \times 100 = 422.61$$

Fisher: $$P_{0t}^{F} = \sqrt{P_{0t}^{L} \times P_{0t}^{P}} = \sqrt{458 \times 390} = 422.7$$

Changing The Base:

Often one may want to change the base period and /or base value of an index after it has been constructed.

Suppose we have an index with a base value of 100 in 1970. But, we desire to have a base of 100 in 1972:

Year	Old Index		New Index
1970	100	Adjust	
1971	110	each	
1972	120	entry	
1973	130	by	
1974	142	dividing	
⋮	⋮	by	⋮
1994	296	1.2	

What numerical value multiplied by 120 equals 100, such that the relatives are left unaltered? (Or equivalently 120 divided by what numerical value equals 100?)

$$\frac{120}{X} = 100 \text{ solving for X: } \frac{120}{100} = 1.2$$

or

$$120 \times X = 100 \text{ solving for X: } \frac{100}{120} = 0.833$$

- Base period is changed by proportional scaling– ***r________ unaltered***.

▲To create the new index, divide each value of the "old" index by 1.2 (or multiply each value of the "old" index by 0.833).

Overall % changes do not change.

<u>Quantity Indices:</u> Same idea as price indices.

<u>Reverse</u> the roles of prices and quantities. Prices are now acting as the ________.

<u>S______ Quantity Index:</u>

$$\boxed{Q_{0t}^{\ S} = \frac{\sum q_{it}}{\sum q_{i0}} \times 100}$$

<u>Laspeyres' Quantity Index:</u>

$$\boxed{Q_{0t}^{L} = \left[\frac{\sum_{i=1}^{n} q_{it} p_{i0}}{\sum_{i=1}^{n} q_{i0} p_{i0}} \right] \times 100}$$

Base period prices = weight

<u>P_______ Quantity Index:</u>

$$Q_{0t}^{P} = \left[\frac{\sum_{i=1}^{n} q_{it} p_{it}}{\sum_{i=1}^{n} q_{i0} p_{it}} \right] \times 100$$

Current period prices are the weights.

<u>Marshall Edgeworth</u>:

$$Q_{0t}^{ME} = \left[\left[\frac{\sum_{i=1}^{n} q_{it} \left(\frac{p_{i0} + p_{it}}{2} \right)}{\sum_{i=1}^{n} q_{i0} \left(\frac{p_{i0} + p_{it}}{2} \right)} \right] \right] \times 100$$

<u>F_____:</u>

$$Q_{0t}^{F} = \sqrt{Q_{0t}^{L} \times Q_{0t}^{P}}$$

Choosing Between Indices: 3 TESTS

(1) Time __________ Test:

P_{it}=price of the i^{th} good at time t.

For the single good: $$\frac{P_{it}}{P_{i0}} = \frac{1}{\left(P_{i0} \big/ P_{it} \right)}$$

I.e. Information about price change is the same, regardless of **reference date**.

Example:

Gum: Price in 2000 is \$1.15.
Price in 2005 is \$1.75.

$$\frac{1.75}{1.15} = \frac{1}{\left(1.15 / 1.75\right)} \Leftrightarrow (1.75)\left(1.15 / 1.75\right) = 1.15$$

$$1.52173913 = 1.52173913$$

● This motivates the idea of the **Time Reversal Test** for indices.

If it passes:

$$(P_{0t} * P_{t0}) = 1$$

where:

P_{0t} is base at 0 to time t and

P_{t0} is base at t to time 0.

<u>Example:</u> Laspeyres' Index:

$$P_{01}^{L} = \left[\left[\frac{\sum_{i=1}^{n} p_{i1}q_{i0}}{\sum_{i=1}^{n} p_{i0}q_{i0}}\right]\right] \quad \text{(assume the base value is 1 not 100)}$$

$$P_{10}^{L} = \left[\left[\frac{\sum_{i=1}^{n} p_{i0}q_{i1}}{\sum_{i=1}^{n} p_{i1}q_{i1}}\right]\right]$$

Test: $\left(P_{01}^{L} \times P_{10}^{L}\right) = 1$

Fails if: $\left(P_{01}^{L} \times P_{10}^{L}\right) \neq 1$

Laspeyres' generally fails the Time Reversal Test.

Using the data from the previous example:

$$P^{L}_{1990,2010} = \left[\left[\frac{\sum_{i=1}^{n} p_{i2010}q_{i1990}}{\sum_{i=1}^{n} p_{i1990}q_{ii1990}}\right]\right] = \left[\frac{(1.80)(1) + (2.70)(4)}{(0.75)(1) + (0.50)(4)}\right] = \frac{12.6}{2.75} = 4.5818$$

$$P^{L}_{2010,1990} = \left[\left[\frac{\sum_{i=1}^{n} p_{i1990}q_{i2010}}{\sum_{i=1}^{n} p_{i2010}q_{i2010}}\right]\right] = \left[\frac{(0.75)(2) + (0.50)(3)}{(1.80)(2) + (2.70)(3)}\right] = \frac{3}{11.7} = 0.2564$$

Test: $\left(P^{L}_{1990,2010} \times P^{L}_{2010,1990}\right) = (4.5818)(0.2564) = 1.1748 \neq 1$

Laspeyres' Price index fails the Time Reversal test.

Generally: Time Reversal Test >1 for Laspeyres' and < 1 for Paasche.

(2) **<u>Factor __________ Test:</u>**

For a single good, <u>price ratio</u> is $\left(\frac{p_1}{p_0}\right)$ and <u>quantity ratio</u> is $\left(\frac{q_1}{q_0}\right)$.

The change in **<u>e__________</u>** is the product: $\left(\frac{p_1 q_1}{p_0 q_0}\right) = \left(\frac{p_1}{p_0}\right)\left(\frac{q_1}{q_0}\right)$.

⇨The **<u>Factor Reversal</u>** test for an index checks:

$$\left[P_{01} * Q_{01}\right] = \left[\frac{\sum p_{i1} q_{i1}}{\sum p_{i0} q_{i0}}\right]$$

$Where\ P_{01}$ = price index

Q_{01} = quantity index

Suppose P_{01} shows a 20% rise and expenditure has risen by 50%, then Q_{01} should show a 25% rise:

$$\left[P_{01} \times Q_{01}\right] = \left[\frac{\sum p_{i1}q_{i1}}{\sum p_{i0}q_{i0}}\right]$$

$$(1.20) \times (1.25) = 1.50$$

Both Laspeyres' and Paasche' Indices ____ the Factor Reversal Test.

Example: Using the previous example:

$$P_{01}^{L} = \left[\frac{\sum_{i=1}^{n} p_{i1} q_{i0}}{\sum_{i=1}^{n} p_{i0} q_{i0}} \right] \qquad Q_{01}^{L} = \left[\frac{\sum_{i=1}^{n} q_{i1} p_{i0}}{\sum_{i=1}^{n} q_{i0} p_{i0}} \right]$$

$$\textit{The test:} \quad \left(\mathrm{P}_{01}^{\mathrm{L}} \times Q_{01}^{L} \right) = \left[\frac{\sum_{i=1}^{n} p_{i1} q_{i1}}{\sum_{i=1}^{n} p_{i0} q_{i0}} \right]$$

$$P^{L}_{1990,2010} = \left[\left[\frac{\sum_{i=1}^{n} p_{i2010}q_{i1990}}{\sum_{i=1}^{n} p_{i1990}q_{ii1990}}\right]\right] = \left[\frac{(1.80)(1) + (2.70)(4)}{(0.75)(1) + (0.50)(4)}\right] = \frac{12.6}{2.75} = 4.5818$$

$$Q^{L}_{1990,2010} = \left[\left[\frac{\sum_{i=1}^{n} q_{i2010}p_{i1990}}{\sum_{i=1}^{n} q_{i1990}p_{i1990}}\right]\right] = \left[\frac{(2)(0.75) + (3)(0.5)}{(1)(0.75) + (4)(0.5)}\right] = \frac{3}{2.75} = 1.0909$$

$$LHS \mapsto \left(P^{L}_{1990,2010} \times Q^{L}_{1990,2010}\right) = (4.5818)(1.0909) = 4.998$$

$$\sum p_{it}q_{it} = (1.8)(2) + (2.7)(3) = 11.7$$

$$\sum p_{i0}q_{i0} = (0.75)(1) + (0.5)(4) = 2.75$$

$$RHS \mapsto \frac{\sum p_{it}q_{it}}{\sum p_{i0}q_{i0}} = \frac{11.7}{2.75} = 4.2545$$

Since, $4.998 \neq 4.2545$, fails test.

<u>Marshall-Edgeworth</u>

⇨______ time-reversal test

⇨_____ factor-reversal test

<u>Fisher's Ideal</u>

⇨______ time reversal test

⇨______ factor reversal test

(3) Time Period Sequences:

- So far we have considered only 2 points in time.

▲ **How do we measure price changes from period 1 to period 2?**

Either: (1) Form P_{12} directly (using period 1 as the base year)

Or (2) take $\left[{P_{02}}/{P_{01}} \right] = P_{12}$ directly.

Should get the same result either way.

Note:

$$P_{02} = \frac{P_2}{P_0}$$

$$P_{01} = \frac{P_1}{P_0}$$

$$\frac{P_2 / P_0}{P_1 / P_0} \text{ rearranging:} \frac{P_2}{P_0} * \frac{P_0}{P_1} = \frac{P_2}{P_1} = P_{12}$$

Suppose: Prices rise by 20% from period 0 to 1.
Prices rise by 10% from period 1 to 2.

Overall rise from period 0 to period 2 is (1.2 × 1.1) =1.32 (32% rise)

Index Values: 100 120 132
(0) (1) (2)

This motivates the **C____________ Test:**

Index passes this test if: $\boxed{P_{02} = (P_{01} \times P_{12})}$

For example, consider the Laspeyres':

$$P_{01}^{L} = \left[\frac{\sum_{i=1}^{n}(p_{i1}q_{i0})}{\sum_{i=1}^{n}(p_{i0}q_{i0})}\right]$$

$$P_{12}^{L} = \left[\frac{\sum_{i=1}^{n}(p_{i2}q_{i1})}{\sum_{i=1}^{n}(p_{i1}q_{i1})}\right]$$

$$But: \quad \left(\mathrm{P}_{02}^{\mathrm{L}}\right) = \left[\frac{\sum_{i=1}^{n}(p_{i2}q_{i0})}{\sum_{i=1}^{n}(p_{i0}q_{i0})}\right] \neq \left[P_{01}^{L} \times P_{12}^{L}\right]$$

Laspreyres' index _____ the circularity test.

Is this important? YES!!–the CPI is essentially a Laspreyres'-type of index. ☞ Fails tests and tends to ____**-state** price rises.

▲All indices mentioned so far fail this test, except for the weighted aggregative index with fixed weights.

We have only considered Aggregative Indices up to this point. Now consider:

<u>**(II) Averages of __________ Indices**</u>

A **<u>price relative</u>** is: $\left[P_{it} / P_{i0} \right]$.

▲The assumption is that each commodity is used ______ as much.

But, an index should take into account the differing quantities of the items used.

⇨Weighted price indices permit the consideration of the <u>relative</u> importance of the commodities in the basket of goods. I.e. use quantities as weights.

A weighted average of price relatives index weights the various price relatives in a market basket by the total amount **s____** on that commodity.

Weighted Arithmetic Mean of Price Relatives

$$P_{0t}^{wm} = \left[\frac{\sum v_i \left({P_{it}}/{P_{i0}} \right)}{\sum v_i} \right] * 100$$

$$\boxed{P_{0t}^{wm} = \left[\frac{\sum p_{it} q_{i0}}{\sum p_{i0} q_{i0}} \right] = P_{0t}^{L}.}$$

Note: if $v_i = \left(p_{i0} q_{i0} \right)$, then $P_{0t}^{wm} = P_{0t}^{L}$

$(i.e. base\ year\ weights -- \exp enditure)$

v_i represents the weight.

Geometric Mean of Price Relatives

$$P_{0t}^{G} = \left[\prod \left(\frac{p_{it}}{p_{i0}} \right) \right]^{1/n} \Leftarrow \text{can also be weighted.}$$

Harmonic "Mean of Price Relatives"

$$P_{0t}^{H} = \left\{ \left[\frac{\sum v_i \left(P_{it} / P_{i0} \right)^{-1}}{\sum v_i} \right] \right\}^{-1}$$

Note: if $v_i = \left(p_{it} q_{it} \right)$, then $P_{0t}^{H} = P_{0t}^{P}$

$(i.e.\ current\ \ year\ weights -- \text{exp}enditure)$

Converting data in index number format: Measuring the level of real national output

When we are measuring the level of national income we often make use of **index numbers** to track what is happening to real GDP. In the table below we see the value of consumer spending and also real GDP expressed in £ billion. I have chosen 1995 as the base year for our index of spending and output. So the data for consumer spending and real GDP has an index value of 100.0 in 1995.

To calculate the index number for consumer spending in 1996 we use the following formula

Index (1996) = (consumer spending (1996) / base year consumer spending) x 100

	Consumer spending	Index of consumer spending	Real GDP	Index of real GDP
	£ billion	1995 = 100	£ billion	1995 = 100
1995 (Base)	512.6	100.0	857.5	100.0
1996	531.9	103.8	880.9	102.7
1997	551.1	107.5	908.7	106.0
1998	572.3	111.6	938.1	109.4
1999	598.8	116.8	966.6	112.7
2000	625.1	121.9	1005.5	117.3
2001	644.9	125.8	1027.9	119.9
2002	667.4	130.2	1048.5	122.3
2003	684.8	133.6	1074.9	125.3
2004	710.2	138.5	1108.9	129.3

Using Price indices: "Price D________"

It is common practice to distinguish between **nominal** and **real** values for a series.
I.e. current dollar terms versus constant dollar terms.

Recall: the real value is obtained by:

$$\text{Real Value} = \frac{\text{Nominal Value}}{\text{Index Number}} \times 100$$

Example: invest money for a year at 7% p.a. interest. Now suppose prices also rise by 2%.

Real Rate of Interest: $\left(\frac{1.07-1.02}{1.02}\right)=\left(\frac{1.07}{1.02}-1\right)\approx 5\%\ p.a.$

You can purchase approximately 5% more goods with each dollar than was possible a year ago.

Convert a ________ variable into a "real" variable by **deflating** – divide "nominal" value by price index to get "real" value.

Price Index used as a price d________.

Must choose the appropriate price index for the purpose of price deflator.

Real income is commonly referred to as the **purchasing power of the money income.**

In comparing incomes, wages, rents, GNP, and personal income per capita of different countries, the use of an appropriate deflator is common practice. Hence, the **real value** is more easily recognized.

$$\text{Real Exports} = \frac{\text{Nominal Exports}}{\text{Export Price Index}} \times 100$$

$$\text{Real Consumption} = \frac{\text{Nominal Consumption}}{\text{CPI}} \times 100$$

$$\text{Real Payroll} = \frac{\text{Nominal Value of Payroll}}{\text{Wage Rate Index}} \times 100$$

Obviously, increases/decreases in “nominal” series may be in the same/ opposite direction to those in corresponding “real” series.

Final Interest Rate Data

Year	CPI	Nominal Interest Rate	Inflation Rate	Real Interest Rate
1	100	--	--	--
2	110	15%	10%	5%
3	120	13%	9.1%	3.9%
4	115	8%	-4.2%	12.2%

5.3.3 Splicing techniques

Splicing in this context refers to the combining or joining of more than one method to form a complete time series. Several splicing techniques are available if it is not possible to use the same method or data source in all years. This section describes techniques that can be used to combine methods to minimise the potential inconsistencies in the time series. Each technique can be appropriate in certain situations, as determined by considerations such as data availability and the nature of the methodological modification. Selecting a technique requires an evaluation of the specific circumstances, and a determination of the best option for the particular case. It is *good practice* to perform the splicing using more than one technique before making a final decision and to document why a particular method was chosen. The principal approaches for inventory recalculations are summarised in Table 5.1.

5.3.3.1 OVERLAP

The overlap technique is often used when a new method is introduced but data are not available to apply the new method to the early years in the time series, for example when implementing a higher tier methodology. If the new method cannot be used for all years, it may be possible to develop a time series based on the relationship (or overlap) observed between the two methods during the years when both can be used. Essentially, the time series is constructed by assuming that there is a consistent relationship between the results of the previously used and new method. The emission or removal estimates for those years when the new method cannot be used directly

TABLE 5.1
SUMMARY OF SPLICING TECHNIQUES

Approach	Applicability	Comments
Overlap	Data necessary to apply both the previously used and the new method must be available for at least one year, preferably more.	• Most reliable when the overlap between two or more sets of annual estimates can be assessed. • If the trends observed using the previously used and new methods are inconsistent, this approach is not *good practice*.
Surrogate Data	Emission factors, activity data or other estimation parameters used in the new method are strongly correlated with other well-known and more readily available indicative data.	• Multiple indicative data sets (singly or in combination) should be tested in order to determine the most strongly correlated. • Should not be done for long periods.
Interpolation	Data needed for recalculation using the new method are available for intermittent years during the time series.	• Estimates can be linearly interpolated for the periods when the new method cannot be applied. • The method is not applicable in the case of large annual fluctuations.
Trend Extrapolation	Data for the new method are not collected annually and are not available at the beginning or the end of the time series.	• Most reliable if the trend over time is constant. • Should not be used if the trend is changing (in this case, the surrogate method may be more appropriate). • Should not be done for long periods.
Other Techniques	The standard alternatives are not valid when technical conditions are changing throughout the time series (e.g., due to the introduction of mitigation technology).	• Document customised approaches thoroughly. • Compare results with standard techniques.

Linking ("Splicing") Indices

⇨ Suppose we have a Laspeyres' price index, P_{0t}^{L} for the period t=1963,...,1973. The weights have become outdated over this time period.

⇨ You decide to form new weights in 1973 and construct a separate P_{0t}^{L} series for t=1973,1974,1975,...,1990.

Although these are different indices, you may need a continuous price time series, so combine the two series into one linked series:

Year	Series 1	Series 2		Linked
1963	100			
1964	120			
⋮	⋮			⋮
1973	160	100	(100) × 1.6 =	160
1974		110		
⋮		⋮		
1990		185	(185) × 1.6	

To make the 100 in year 1973 be equal to 160, simply multiply (100)(1.6)=160.

☐ **Relative prices are preserved over time.**

We could link with any other base year:

Year	Series 1		Series 2	Linked
1963	100	$\div$ 1.6		
1964	120	$\div$ 1.6		
	⋮	⋮		
1973	160	(160) $\div$ 1.6 =	100	
1974			110	
			⋮	
1990			185	

In each case, the linked series suggests prices are **1.96 times** higher in 1990 than in 1963.

$$\frac{185-62.5}{62.5}=1.96$$

$$\frac{296-100}{100}=1.96$$

Constructing the CPI

❑The CPI is based on "price _________" for each item in each location. Then they are aggregated to get regional indices and indices for groups. Eventually, you get to an "All groups" index for Canada.

❑The construction of the CPI involves the "***weighted average of price relatives***" approach.

The weights are (base period) ____________, so this is the Laspeyres' Index.

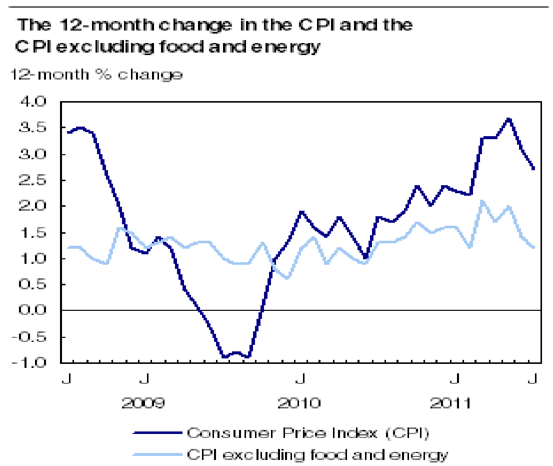
- Base year updated regularly -- used to be every 4 years.

❑Weights revised to reflect changing expenditure patterns (2009 is the latest.)

Note:

Laspeyres: weighted aggregative index, with base period __________ as weights.

CPI: weighted average of price relatives, with base period e____________ as weights.



Revisions and seasonal adjustment

The CPI allows one to compare, in percentage terms, prices in any given time period to prices in the official base period which, at present, is 2002=100. The official time base was changed from 1992=100 to 2002=100 starting with the CPI for May 2007. The change is strictly an arithmetic conversion, which alters the index levels, but leaves the percentage changes between any two periods intact, except for differences in rounding.

A seasonally adjusted series is one from which seasonal movements have been eliminated. Seasonal movements are defined as those which are caused by regular annual events such as variations in climate and regular institutional events such as vacations and statutory holidays. Seasonally adjusted series are calculated using X-12-ARIMA. Time series with no identifiable seasonal pattern remain unchanged from the official series.

The official (2002=100) unadjusted series for the All-items CPI, each of the eight major component indexes and four special aggregates (Core CPI (Bank of Canada definition), All-items CPI excluding eight of the most volatile components identified by the Bank of Canada, All-items CPI excluding food, All-items CPI excluding food and energy) are seasonally adjusted independently.

Each month, the previous month seasonally adjusted index is subject to revision. On an annual basis, the seasonally adjusted values for the last three years are revised with the January data release. Since these revisions can lead to changes in both the levels and movements of the indexes, users employing the CPI for indexation purposes are advised to use the unadjusted indexes.

ALL-ITEMS, CPI

1. Food
2. Shelter
3. Household operations and furnishings
4. Clothing and footwear
5. Transportation
6. Health and personal care
7. Recreation, education and reading
8. Alcoholic beverages and tobacco products

Data quality, concepts and methodology: Data quality, concepts and methodology

Previous

Definition

The Consumer Price Index (CPI) is an indicator of the changes in consumer prices experienced by the target population. The CPI measures price change by comparing, through time, the cost of a fixed basket of commodities. This basket is based on the expenditures of the target population in a certain reference period, currently 2009. Since the basket contains commodities of unchanging or equivalent quantity and quality, the index reflects only pure price movements.

Separate CPIs are published for Canada, the ten provinces, Whitehorse, Yellowknife and Iqaluit. Some CPI information is also available for an additional sixteen urban centres. Since the CPI is a measure of price change from one time period to another, it cannot be used to indicate differences in price levels between provinces or urban centres.

Population coverage

The population targeted by the CPI consists of families and individuals living in urban and rural private households. For practical reasons, residents of the Territories outside Whitehorse, Yellowknife and Iqaluit are not represented by the index. Prior to February 1995, the target population consisted of private households in Canadian urban centres with a population of 30,000 or more.

Time reference

The CPI compares, in percentage terms, prices in any given time period to prices in the official base period which, at present, is 2002=100. The official time base was changed from 1992=100 to 2002=100 starting with the CPI for May 2007. The change is strictly an arithmetic conversion which alters the index levels but leaves the percentage changes between any two periods intact, except for differences in rounding.

The Consumer Price Index

The CPI measures changes in consumer prices facing the Canadian consumer.

It is derived by comparing the cost of a _____basket of commodities purchased by consumers over time.

The "basket" is assumed to contain commodities of ***unchanging*** quantity and _______ and hence is believed to reflect only "pure price movements."

Between 1913 and1926 both the Department of Labour and the Dominion Bureau of Statistics (Statistics Canada) each produced separate CPI's. The present CPI series descended from these sources.

Some component indexes in the CPI begin with the date they were introduced into the CPI.

The CPI is widely used as an indicator of the **changes in the rate of inflation.** Hence, consumers can monitor changes in their personal income.

<u>The CPI is generally employed in four ways:</u>

1) Evaluate and escalate the p________g power over time of:
   -wages -rents
   -leases -child or spousal support allowance
   -private and public pension programs (Old Age Security and Can. Pension)
   -personal income tax deductions
   -government social payments

2) Used to _______ current dollar estimates to obtain constant dollar estimates.

3) Setting and evaluating economic policies
   - Bank of Canada's monetary policies
   - Assessment of public policy– food prices, etc.

4) Economic analysis and research by economists
   - explore effects of inflation
   - causes of inflation
   - regional disparities in price movements

CPI is based upon prices (including taxes) paid by consumers in private retail outlets, government stores, offices and other consumer service establishments.

CPI considers a constant "______" of goods over time – prices in cities greater than or equal to __,___.

"Price movements of the goods and services represented in the CPI are weighted according to the relative importance of commodities in the total expenditure of consumers."

CPI basket weights are collected from multiple types of _______ relating to a specific year. The weights are currently based on 2009 consumer expenditure data. The current time base of the index is 2002 = 100.

<u>Note</u>: The CPI does not measure changes in the “cost of living” – the latter is the change in income needed to keep a consumer as “well off” – depends on individual preferences.

The CPI is not a ***cost-of-living index***, though people frequently call it this. In theory, the objective behind a cost-of-living index is to measure price changes experienced by consumers in maintaining a constant standard of living. The idea is that consumers would normally switch between products as the price relationship of goods changes. If, for example, consumers get the same satisfaction from drinking tea as they do from drinking coffee, then it is possible to substitute tea for coffee if the price of tea falls relative to the price of coffee. The cheaper of the interchangeable products may be chosen.

We could compute a cost-of-living index for an individual if we had complete information about that person's taste and consuming habits. To do this for a large number of people, let alone the total population of Canada, is impossible. For this reason, regularly published price indexes are based on the ***fixed basket*** concept rather than the cost-of-living concept.

Highlights of the Canadian CPI

(i) Relates to prices in urban centres.

(ii) Based on prices of over 300 items.

(iii) 8- “Groups” of goods:

Food
Shelter
Household Operations and Furnishings
Clothing and Footwear
Transportation
Health and Personal Care
Recreation, Education and Reading
Tobacco and Alcoholic beverages

- There are separate indices plus an "All Groups" category.

(iv) Index computed monthly –regional indices also available.

(v) Weight given to each individual price change is based on expenditure as % of total household expenditure. The primary sources of expenditure data on consumer goods and services are from the Survey of Household Spending and the Food Expenditure Survey.

(vi) Prices of each group/service sampled at several outlets: 110,000 price quotations each month. The All-items index at the Canada level is based on an annual sample of over 500,000 prices quotes.

(vii) Some prices quoted twice/month; some once/quarter; some twice/year; some annual. EG. Food: 2 / month; Haircuts: 1 / 3 months; Car insurance: bi-annual

(viii) The extent and quality of the sample is constrained by budgets and the information available on which to base the sampling prices.

Cost of Living

It is reasonable to measure the change in the **cost of living** of an individual between two periods as ***the change in their money income which will be necessary for the individual to maintain his or her original standard of living*** – no more or no less.

It must be pointed out that if all prices change in the same proportion, that proportion will measure the change in the cost of living and there will be no problem of measurement.

☆However, all prices do not change in the same ___________.

Suppose in period 0 an individual spends his _______ by purchasing q_0 of various commodities at price p_0. Assuming he saves nothing, his total income equals his total expenditure: $\sum p_0 q_0$.

➢ In period 1, prices change to p_1. How much _____ does he need in period 1 to make him as well off as he was in period 0?

If the quantities of goods which he would need to buy to leave him exactly as well off as he was in period 0 are $\overline{q}_1$, then with an income of $\sum p_1 \overline{q}_1$ he would be exactly as well off.

It follows that the change in the cost of maintaining his original (period 0) standard of living will be given by

$$\boxed{C_{01}^{0} = \frac{\sum p_1 \bar{q}_1}{\sum p_0 q_0}}.$$

The superscript (0) indicates that the change in his cost of living is being measured in terms of his period 0 standard of living.

Unfortunately we do not know the $\bar{q}_1$'s because they are not the quantities he actually buys in period 1, but only what he would need to buy to be as well off as before.

➢ **<u>Consequently, we do not know the aggregate $\sum p_1 \bar{q}_1$.</u>**

However, we do know the aggregate $\sum p_1 q_0$. This is the amount of ______ he would require in period 1 to enable him to purchase the <u>quantities</u> he purchased in period 0.

It can be shown that $\sum p_1 q_0$ will be greater than $\sum p_1 \bar{q}_1$, provided his tastes have remained unchanged.

If he did have the income $\sum p_1 q_0$ in period 1, he would not buy the same quantities as he bought in period 0, (the q_0 quantities) because he would take advantage of the changes in relative prices and would alter his allocation of ______, buying relatively more of those goods whose prices had fallen relatively more or risen relatively less.

The fact that he would do this in preference to buying the q_0 quantities indicates that an income of $\sum p_1 q_0$ would make him **<u>better off</u>** than the original ______ of $\sum p_0 q_0$ and, hence better off than an ______ of $\sum p_1 \overline{q}_1$ which is its equivalent.

➢We have accordingly, $\sum p_1 q_0 > \sum p_1 \bar{q}_1$.

This assumption that his tastes must have remained unchanged is vital because if they have changed, we cannot conclude that the quantities he would buy in period 1 with an _______of $\sum p_1 q_0$ are preferred to those he actually did buy in period 0.

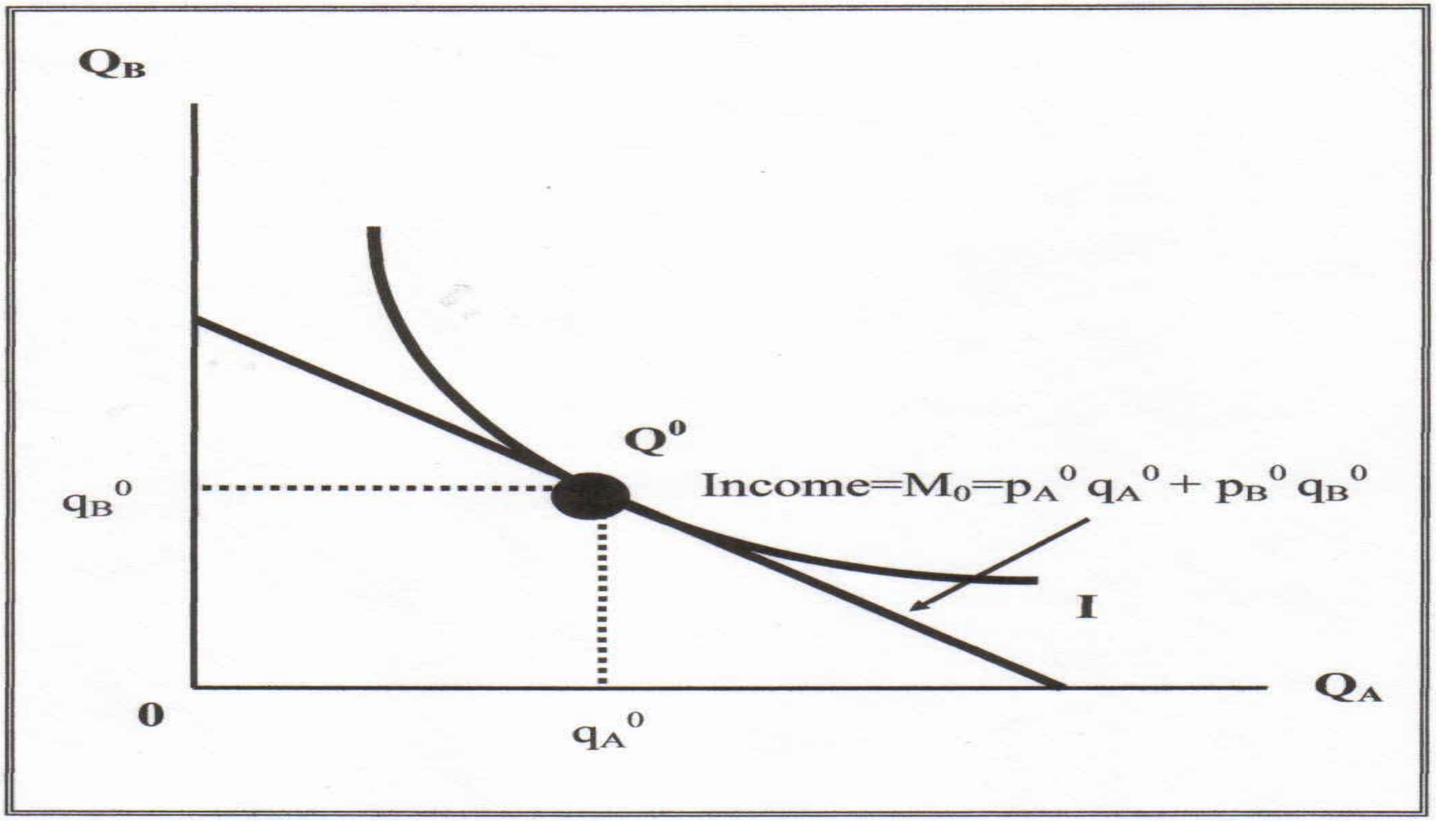
This can be illustrated, by means of **indifference curves**, in the case of an individual spending his _______ on two goods.

We call the two goods A and B. Let the price of A in period 0 be p_A^0 and of B be p_B^0. If the individual's money ______ in period 0 is M_0, we shall have:

$$\boxed{M_0 = p_A^0 q_A^0 + p_B^0 q_B^0}$$

where q_A^0 and q_B^0 are the quantities of A and B respectively which he can purchase under these circumstances.

2 Good Example: Time Period 0



Any combination of good A and good B can be purchased along or below the budget line.
Assume fixed preferences and one individual.

$$\text{Slope} = -\frac{P_A^0}{P_B^0}$$

On the diagram, q_A^0 is measured on the X-axis and q_B^0 on the Y-axis. The individual can take up any position on the line – his budget line, M_0. It has a slope given by the price of A relative to that of B.

➔ Slope: $\left(-\frac{p_A^0}{p_B^0}\right)$

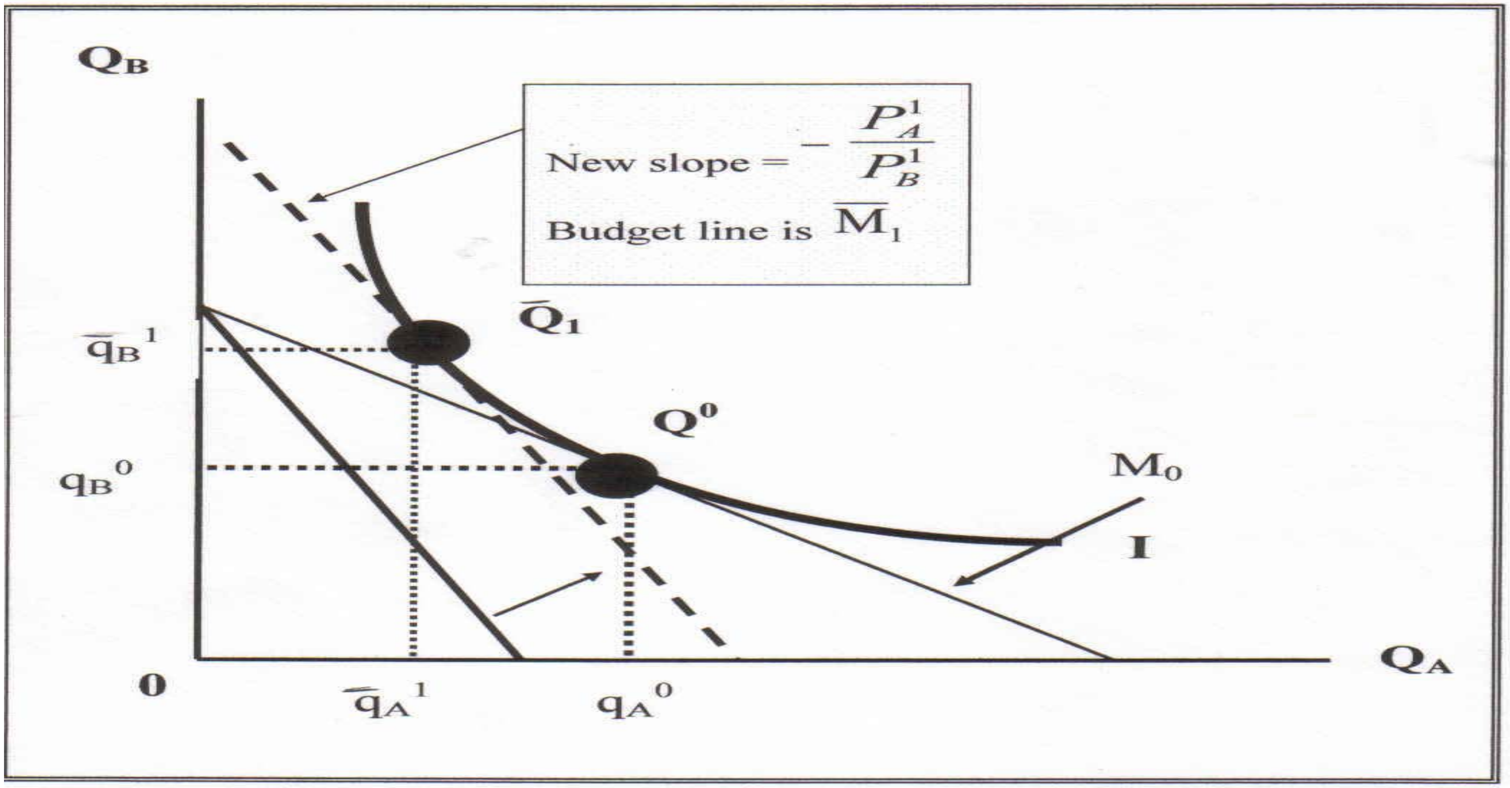
The individual will move to the point Q_0, where quantities of A and B, q_A^0 and q_B^0, meet on the budget line, at which point he reaches his highest indifference curve. The individual will purchase q_A^0 and q_B^0 in period 0.

Suppose, in period 1 prices begin to change. (Price of A increases.)

How much ______ will the individual now need to be as well off as he was before?

The individual will need sufficient to enable him to take up a position **just** on the indifference curve. We now draw the line $\overline{M}_1$ with a slope given by $\left(-\frac{p_A^1}{p_B^1}\right)$ such that it just touches the indifference curve.

2 Good Example: Time Period 1
Price of Good A increases.



New Budget line: $\overline{M}_1 = P_A^1 q_A^1 + P_B^1 q_B^1$

The individual will need income $\overline{M}_1$ to consume at the same level of satisfaction as period 0.

$C_{01}^0 = \overline{M}_1 / M_0 =$ change in the cost of living.

With an

With an ______ corresponding to this line, he would take up the position $\overline{Q}_1$, purchasing $\overline{q}_A^1$ of A and $\overline{q}_B^1$ of B.

This would require ______: $\overline{M}_1 = p_A^1\overline{q}_A^1 + p_B^1\overline{q}_B^1$, as against the old income of $M_0 = p_A^0 q_A^0 + p_B^0 q_B^0$.

■ But the individual is indifferent between the positions Q_0 and $\overline{Q}_1$, so that the ratio $\overline{M}_1 / M_0$ measures the change in his old income necessary to make him as well off in period 1 as he was in period 0.

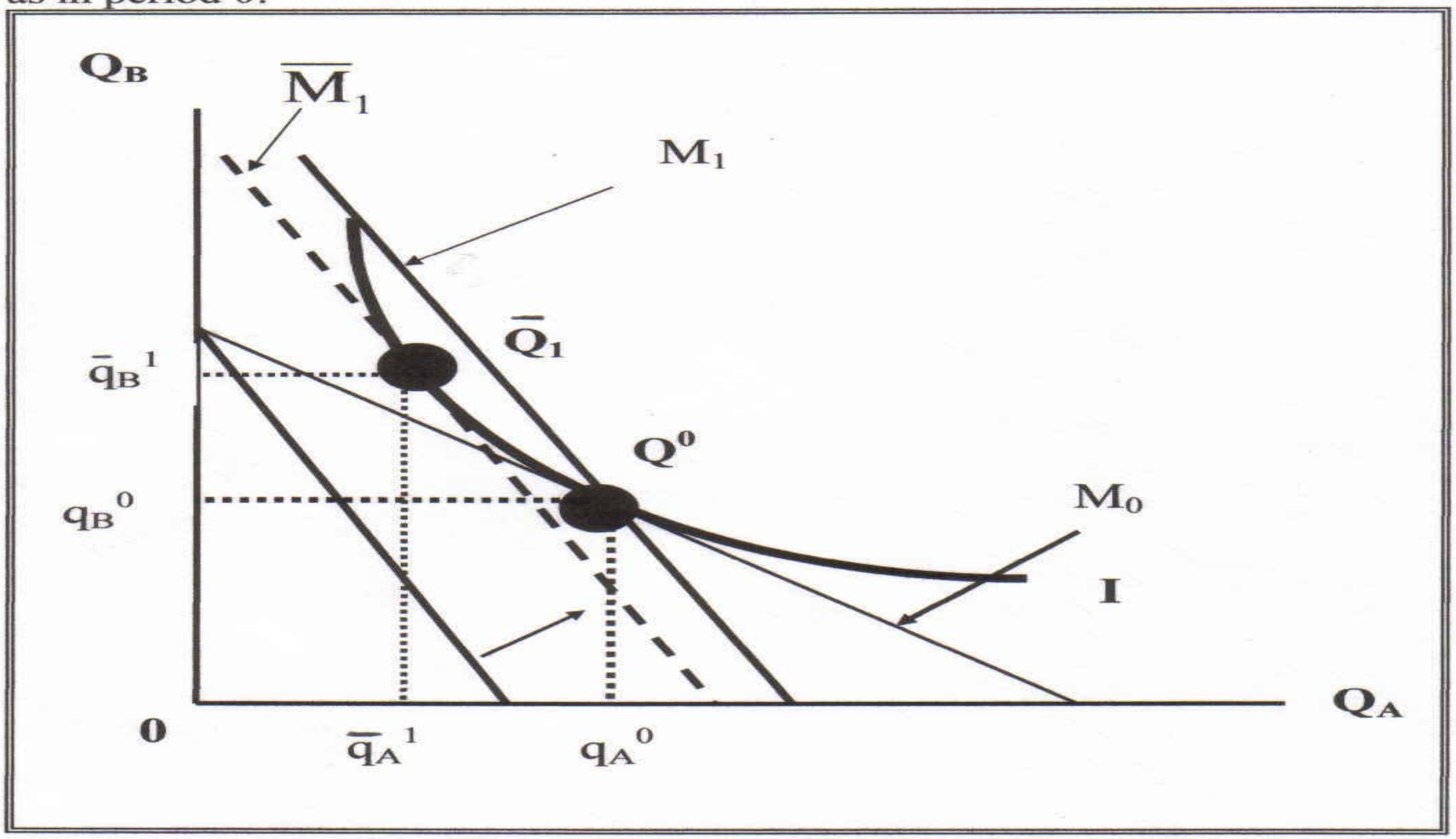
Suppose that the individual is given ______ in period 1 sufficient to enable him to purchase the same quantities as he did in period 0.

This ______ must be sufficient to make the budget line for the new prices pass through Q_0 and must be equal to $M_1 = p_A^1 q_A^0 + p_B^1 q_B^0$.

This is shown as the broken line in the diagram, and it must lie parallel to but to the right of $\overline{M}_1$, so that:

$$M_1 = p_A^1 q_A^0 + p_B^1 q_B^0 > \overline{M}_1 = p_A^1 \overline{q}_A^1 + p_B^1 \overline{q}_B^1 .$$

If given enough income to purchase the same quantities in period 1 as in period 0:



$$M_1 > \overline{M}_1$$

M_1 → Laspreyres' index uses base period quantities as the weights. (Hence the tendency for this index to overestimate a rise in price level.)

6

Price Indices and the "Cost of Living"

☐ ***Can a price index, such as the CPI, measure a change in the "cost of living"?***

⇨Individuals are affected ___________ by price changes.

Definition: A change in the **cost of living** is the change in ($) income needed to maintain ***original*** 'standard of living.'

➡This depends on individual **preferences.**

⇨Let $\overline{q}_{i1}$ be the quantity of good i that I need in period "1" to maintain the standard of living I enjoyed in period "0."

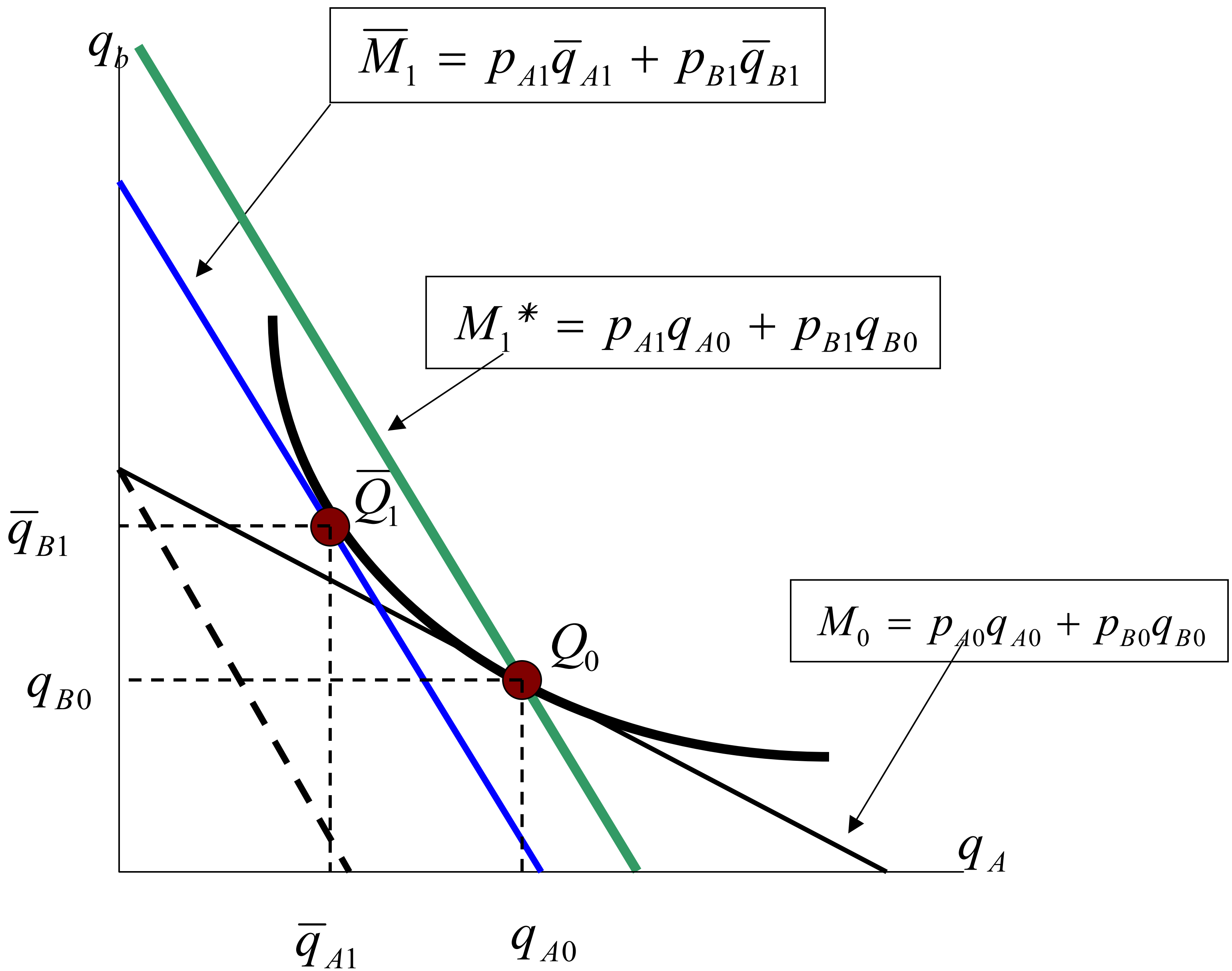
Cost of Living Index is:

$$C_{01}^{0} = \left[\frac{\sum p_{i1}\bar{q}_{i1}}{\sum p_{i0}q_{i0}} \right]$$

Compare C_{01}^{0} with price indices we have used.

Assume:

(1) One individual
(2) _____ preferences


$\overline{M}_1 = p_{A1}\overline{q}_{A1} + p_{B1}\overline{q}_{B1}$
$M_1{}^* = p_{A1}q_{A0} + p_{B1}q_{B0}$
$M_0 = p_{A0}q_{A0} + p_{B0}q_{B0}$
q_b
q_A
$\overline{q}_{B1}$
q_{B0}
$\overline{q}_{A1}$
q_{A0}
$\overline{Q}_1$
Q_0

The price of A increases:

Relative prices change: “A” is now more __________.

Budget line shifts ⇨slope increases

Consumer is indifferent between: $\left\{ \begin{matrix} \left(\mathrm{q}_{A0}, \mathrm{q}_{B0}\right) \\ \left(\overline{\mathrm{q}}_{A1}, \overline{\mathrm{q}}_{B1}\right) \end{matrix} \right\}$

So, $C_{01}^{0} = \left[\dfrac{\sum p_{i1}\overline{q}_{i1}}{\sum p_{i0}q_{i0}} \right] = \dfrac{\overline{M}_1}{M_0}$.

To purchase the original bundle, consumer needs income of :

$(p_{A1}q_{A0} + p_{B1}q_{B0})$ (the green line). M_1 *

Note: $\sum p_{i1}q_{i0} > \sum p_{i1}\overline{q}_{i1}$

$(M_1{*}) \qquad (\overline{M}_1)$

If we divide both sides by :

$\sum p_{i0}q_{i0} \quad \rightarrow$ base period expendiure :

$$\left[\frac{\sum p_{i1}q_{i0}}{\sum p_{i0}q_{i0}}\right] > \left[\frac{\sum p_{i1}\bar{q}_{i1}}{\sum p_{i0}q_{i0}}\right]$$

or

$$\mathrm{P}_{01}^{\mathrm{L}} \quad > \quad \mathrm{C}_{01}^{0}$$

Laspeyres' Price Index **o________** rises in true cost-of-living, and **understates** falls in the true cost-of-living.

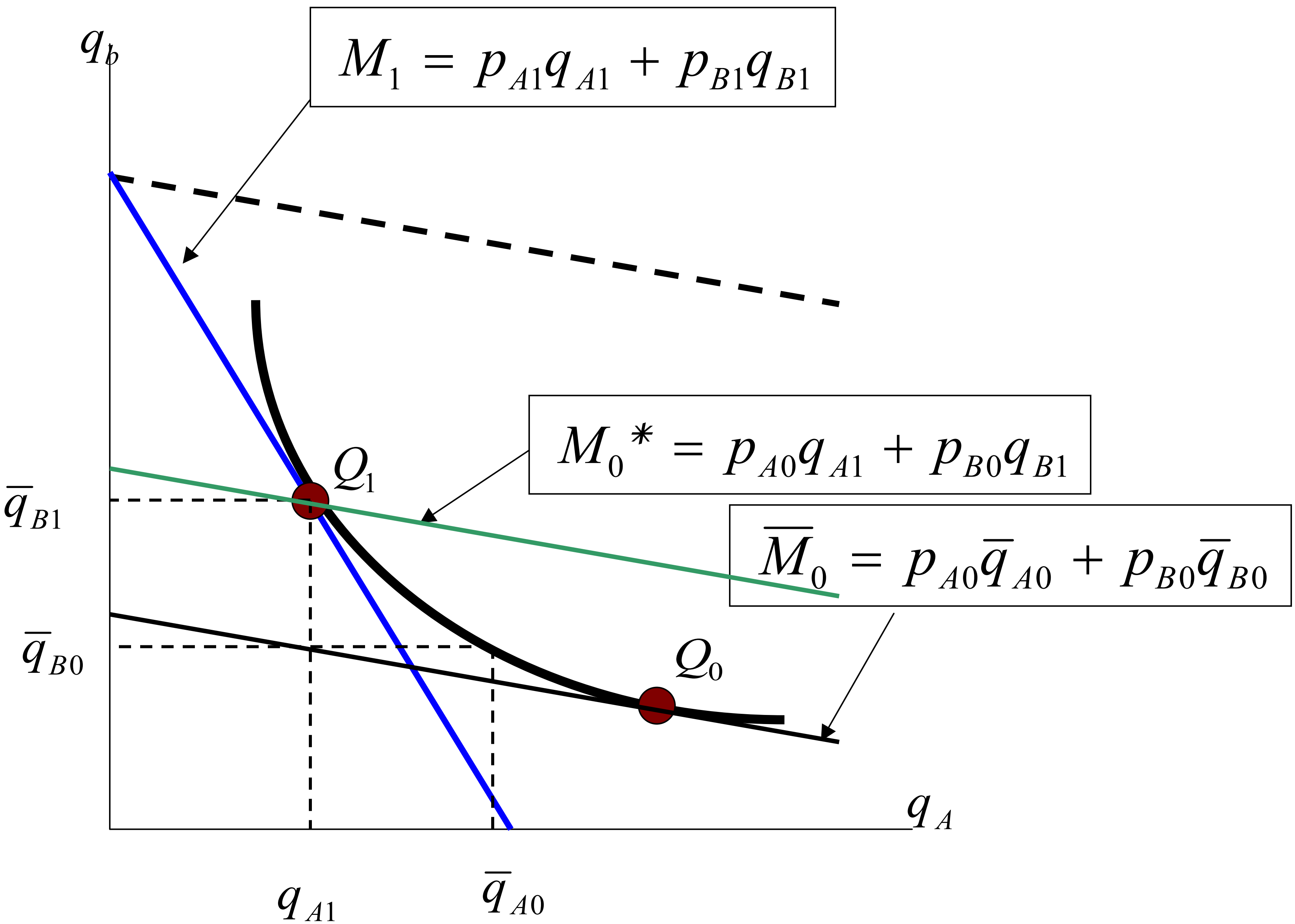
Can also define an alternative indicator of the change in cost of living:

"the individual's income in period "1" as a ratio of what would have been needed in period "0", at period "0" prices to be as well off":

$$C_{01}^{1} = \left[\frac{\sum p_{i1} q_{i1}}{\sum p_{i0} \bar{q}_{i0}} \right] = \frac{M_1}{\bar{M}_0}$$

where $\bar{q}_{i0}$ = quantity of good *i* in period "0" needed to maintain "welfare" experienced in period "1".

- *the superscript on C.O.L. index refers to the utility base period.*


$M_1 = p_{A1}q_{A1} + p_{B1}q_{B1}$
q_b
$M_0^* = p_{A0}q_{A1} + p_{B0}q_{B1}$
Q_1
$\bar{q}_{B1}$
$\bar{M}_0 = p_{A0}\bar{q}_{A0} + p_{B0}\bar{q}_{B0}$
$\bar{q}_{B0}$
Q_0
q_A
q_{A1}
$\bar{q}_{A0}$

Start at the blue line: $M_1 = p_{A1}q_{A1} + p_{B1}q_{B1}$

Suppose price is lower is the previous time period:

$$M_1 = (p_{A1}q_{A1} + p_{B1}q_{B1})$$
$$\overline{M}_0 = (p_{A0}\overline{q}_{A0} + p_{B0}\overline{q}_{B0})$$

$$Now: (p_{A0}q_{A1} + p_{B0}q_{B1}) > (p_{A0}\overline{q}_{A0} + p_{B0}\overline{q}_{B0})$$

green line black line

M_0^* $\overline{M}_0$

or

$$\sum p_{i0}q_{i1} \quad > \quad \sum p_{i0}\overline{q}_{i0}$$

If we divide by $\sum p_{i1}q_{i1}$:

$$\left[\frac{\sum \mathrm{p}_{i0}\mathrm{q}_{i1}}{\sum \mathrm{p}_{i1}\mathrm{q}_{i1}}\right] > \left[\frac{\sum \mathrm{p}_{i0}\overline{\mathrm{q}}_{i0}}{\sum \mathrm{p}_{i1}\mathrm{q}_{i1}}\right] \leftarrow \text{invert both sides \&}$$

reverse the inequality sign

$$\text{Paasche price index} \rightarrow \left[\frac{\sum \mathrm{p}_{i1}\mathrm{q}_{i1}}{\sum \mathrm{p}_{i0}\mathrm{q}_{i1}}\right] < \left[\frac{\sum \mathrm{p}_{i1}\mathrm{q}_{i1}}{\sum \mathrm{p}_{i0}\overline{\mathrm{q}}_{i0}}\right] \leftarrow C_{01}^{1}$$

understates rises & overstates falls in true C.O.L. index.

$$\boxed{P_{01}^{P} < C_{01}^{1}}$$

A Paasche price index ___________ rises and overstates falls in the true cost of living index!

Note:

- 2 different C.O.L. indices.

- Each index involves "______________" quantities.

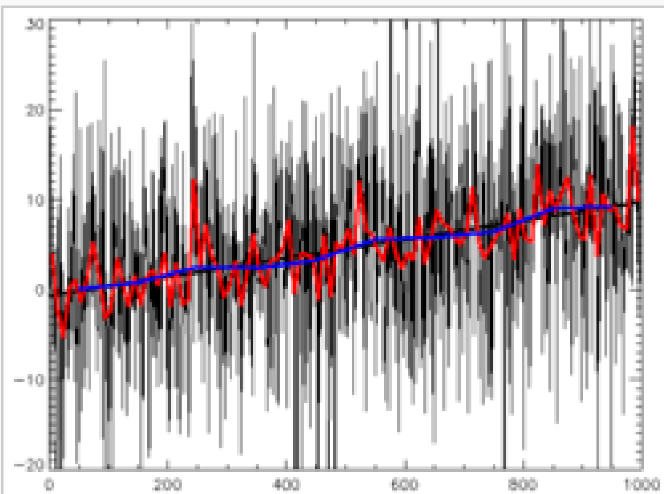
- Cannot tell extent of under/over-statement.

- Each individual has own **preferences**.

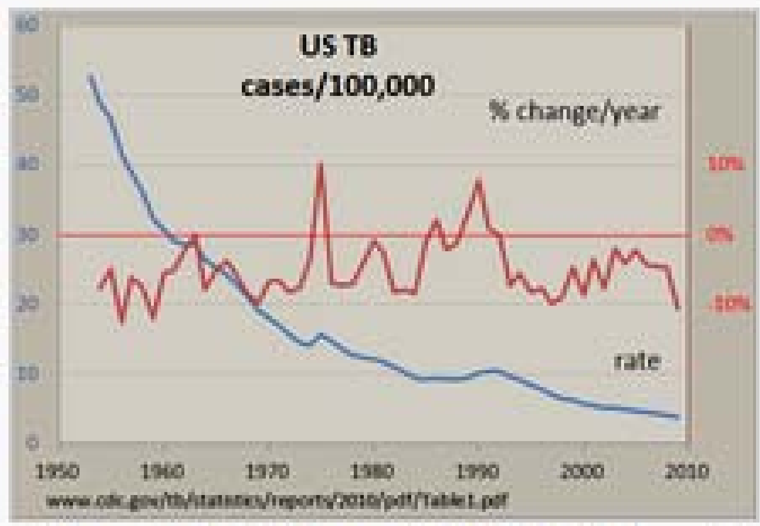
- A price index takes **no** account of preferences – cannot measure changes in true C.O.L. (Or of welfare").

Topic 3 – Part II
Analysis of Economic Time Series Data

Recording observations of a variable that is a function of time, results in a set of numbers called a **time series**.



Time series: random data plus trend, with best-fit line and different smoothings



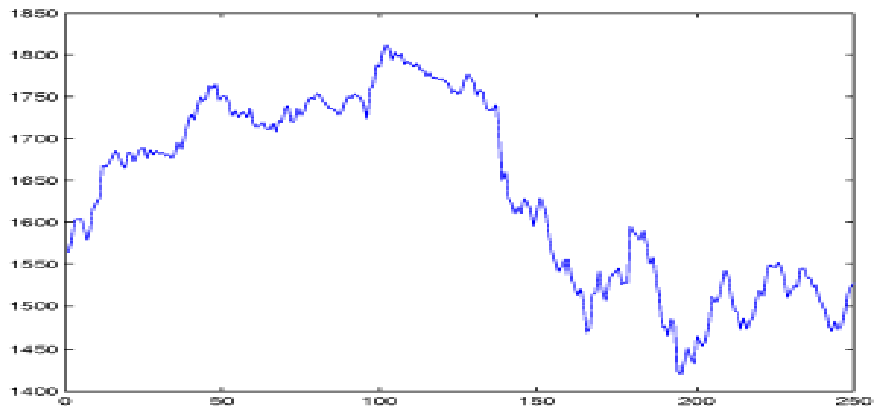
Tuberculosis incidence US 1953-2009

<u>Time Series Data</u>: a sequence of data values, gather over time – usually at some fixed interval.

Since data have a natural order, we can analyse a time series of data in terms of **<u>4</u>** basic components:

(i) T____: Long term "direction" of the series, over many years.

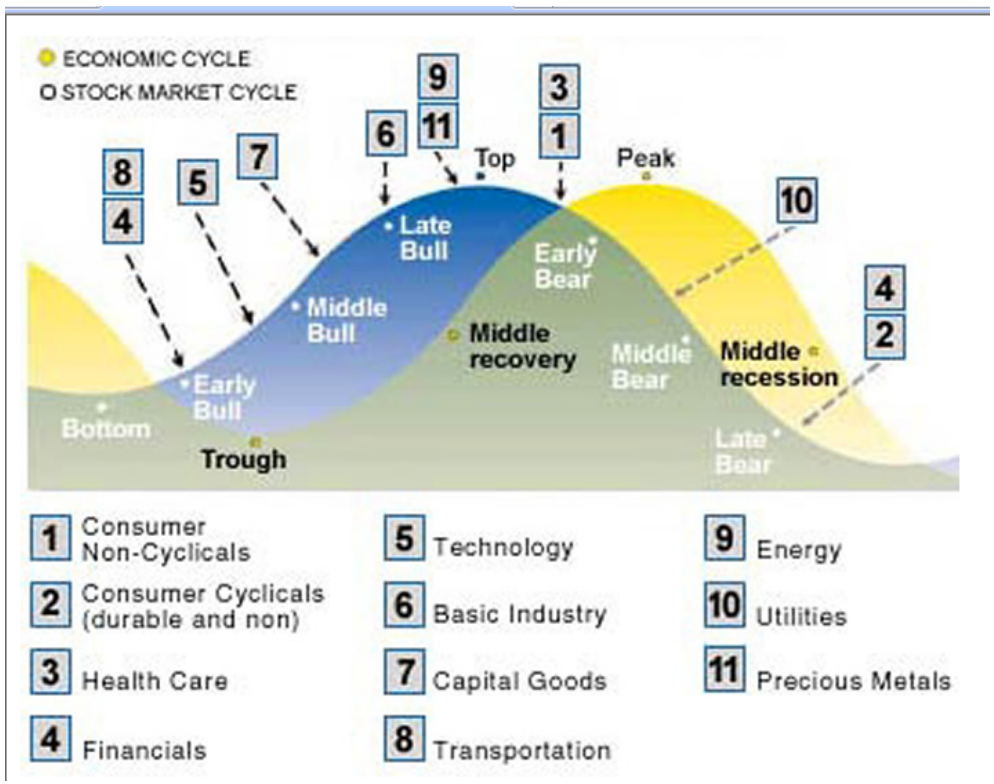
⇨Direction and shape of the series is important to us.



(ii) C_____: Represents a pattern repeated over time periods of different length, usually longer than a _____.

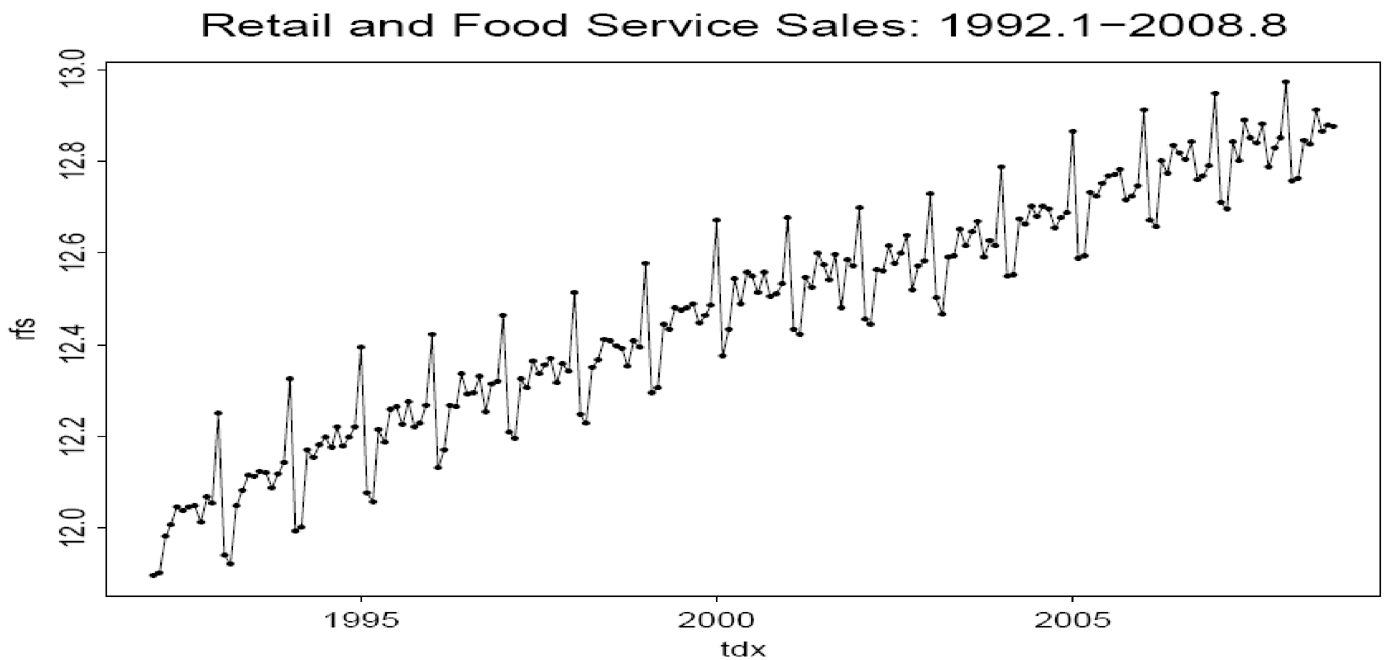
⇨"Wave-like" pattern – (3-5 duration)

⇨eg. Business cycles.



(iii) S________: represents fluctuations that repeat themselves within a fixed period of one ____.

⇨Same pattern each year.



(iv) I__________: fluctuation that is unpredictable, or takes place by chance or randomly.

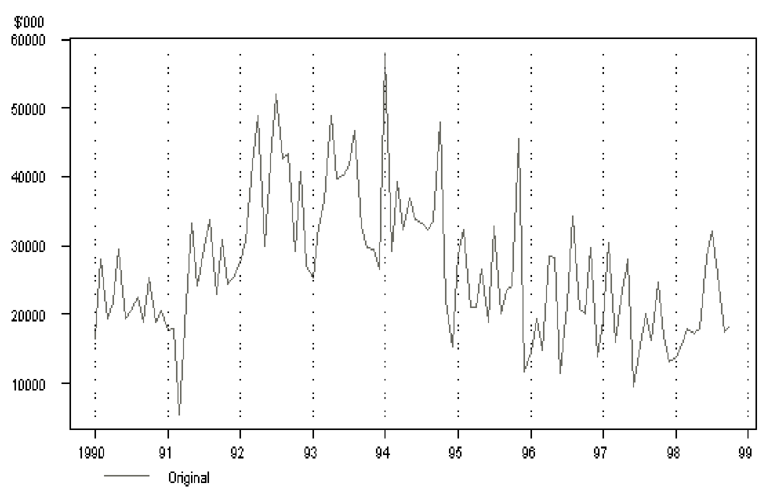
⇨Random "noise" in the data.

⇨Non-systematic.

WHAT IS AN IRREGULAR?

The irregular component (sometimes also known as the residual) is what remains after the seasonal and trend components of a time series have been estimated and removed. It results from short term fluctuations in the series which are neither systematic nor predictable. In a highly irregular series, these fluctuations can dominate movements, which will mask the trend and seasonality. The following graph is of a highly irregular time series:

Figure 2: Monthly Value of Building Approvals, Australian Capital Territory (ACT)



The Time Series Model:

It is convenient to "model" a time series using these components.

Time series models fall into 2 major categories:

(A) A________: Y=T+C+S+I ⇦ assume components are independent of one another.

(B) M________________: Y=T*C*S*I ⇦ 4 components are related to each other, yet result from different basic causes.

**(Multiplicative model is more important and connection between the two.)

Focus first on the **decomposition of Y into components**. Why?

(I) Components are useful in own right.
(II) Aids in forecasting of Y.

- For the purpose of estimating each of the components of a time series, models normally treat S, C, and I as **deviations from the _____**.

- Trend is usually estimated in units of Y, and the other components are measured in _____ form, with values greater than 100 indicating a deviation **above** the trend value and values below100 indicating a movement **below** the trend.

I.e. Represent cycle and seasonal as fraction of trend:

> 100% ⇨ above trend; < 100%⇨ below trend

The **Steps of Decomposition** of a time series into its components for a **multiplicative model**:

(1) Isolate Seasonal component by the ratio-to-moving average method.

⇨Determine seasonal index with a base value of 100.

(2) Remove Seasonality from Y, leaving (T×C×I).

⇨ Divide the values of Y by the seasonal index S, and multiply by 100 to obtain:

$$100\left(\frac{Y}{S}\right) = T \times C \times I$$

(3) Determine trend component, and remove, leaving (C×I)

⇨divide T×C×I by the trend value $\hat{Y}$ to obtain (C×I).

(4) "Smooth" remainder to eliminate irregular component
⇨leaves cycle.

❑ Once all of the components of a time series are identified, forecasts of the value of the time series can be made by first estimating the value of the trend component at that point in time in the future, and then modifying this trend value by an adjustment that takes into account S and C components.

<u>Seasonality</u>

We are interested in the seasonal component and in __________ it from Y. (i.e. Seasonally adjusting the data).

Basic procedure involves **<u>"smoothing"</u>** the data using moving averages:

$$Y_t = \{\ 8 \quad 2 \quad 6 \quad 1 \quad 10 \quad 4 \quad 12 \quad 2 \quad 8\ \}$$

2 Period MA: $\frac{8+2}{2}=5 \quad \frac{2+6}{2}=4 \quad \frac{6+1}{2}=3.5 \quad \frac{1+10}{2}=5.5 \quad 7 \quad 8 \quad 7 \quad 5$

⇨More terms in average ➡ smoother
⇨Choose terms to span exactly one year.
⇨Quarterly data ➡ 4 terms; ⇨Monthly data ➡ 12 terms
Need to take account of "timing" of data.

Example: How To determine A Seasonally Adjusted Series (Multiplicative Model)

Year	Quarter	Y_{it}	4-Qtr. M.A.	Centred M.A. (Tit*Cit)	Ratio: $S_i \times I_t = \frac{Y_{it}}{(T_{it} \times C_{it})}$
1994	Q1	40			
	Q2	60	45.0		
	Q3	20	47.5	46.25	
	Q4	60	50.0	48.75	
1995	Q1	50	52.5	51.25	
	Q2	70	57.5	55.00	
	Q3	30	55.0	56.25	
	Q4	80	55.0	55.00	
1996	Q1	40	50.0	52.50	
	Q2	70	52.5	51.25	
	Q3	10		----	----
	Q4	90		----	----

Obtain Seasonal Indices: (Average)

$$\boxed{S_1 = (0.976 \times 0.762)^{1/2} \times 100 = 86.2\%}$$

$$\boxed{S_2 = (1.273 \times 1.366)^{1/2} \times 100 = 131.9\%}$$

$$\boxed{S_3 = (0.432 \times 0.533)^{1/2} \times 100 = 48.0\%}$$

$$\boxed{S_4 = (1.231 \times 1.455)^{1/2} \times 100 = 133.8\%}$$

<u>Obtain the geometric mean of S_i's</u>:

$$\overline{S}_i^{GM} = (86.2 \times 131.9 \times 48 \times 133.8)^{1/4} = 92.44 \neq 100\%$$

Since the geometric mean does not equal **<u>100%</u>**, we must divide <u>each</u> **<u>S_i</u>** by 92.44:

$$S_1^* = \frac{86.2}{92.44} \times 100 = 93.2$$

$$S_2^* = \frac{131.9}{92.44} \times 100 = 142.7$$

$$S_3^* = \frac{48}{92.44} \times 100 = 51.9$$

$$S_4^* = \frac{133.8}{92.44} \times 100 = 144.7$$

$$\bar{S}_i^{GM*} = (93.2 \times 142.7 \times 51.9 \times 144.7)^{1/4} \cong 100\%$$

$$(GM = 99.97)$$

Apply same **<u>S_i's</u>** to each year. To seasonally adjust the data:

$$Y_{it}^S = \left[\left(\frac{Y_{it}}{S_i^*}\right) \times 100\right];$$

$Y_{it} = T_{it} * C_{it} * \mathbf{S_{it}} * I_{it}$

$$Y_{it}^S = \frac{(Y_{it})}{(S_i^*)} * 100$$

Year	Quarter	Y_{it}	S_i*	Seasonally Adjusted Series
1994	Q1	40	93.2	42.918
	Q2	60	142.7	42.046
	Q3	20	51.9	38.536
	Q4	60	144.7	41.465
1995	Q1	50	93.2	53.648
	Q2	70	142.7	49.054
	Q3	30	51.9	57.803
	Q4	80	144.7	55.287
1996	Q1	40	93.2	42.918
	Q2	70	142.7	49.054
	Q3	10	51.7	19.268
	Q4	90	144.7	62.198

Differences Between Additive and Multiplicative Models

Multiplicative Model	Additive Model
1) <u>Seasonal factors should average to 1 or 100%:</u> $GM = \left(\Pi s_i\right)^{1/i} = 1$ Example: $GM = (s_1 \times s_2 \times s_3 \times s_4)^{1/4}$	1) <u>Seasonal factors should average to 0:</u> $\sum \frac{s_i}{i} = 0$ Example: $\bar{x} = \frac{s_1 + s_2 + s_3 + s_4}{4} = 0$
<u>2)When adjusting for seasonality</u>: ⇨Divide by index: $\mathrm{y}^{\mathrm{s}} = \frac{\mathrm{y}}{\mathrm{s}_{\mathrm{i}}^{*}}$	<u>2)When adjusting for seasonality</u>: ⇨Always subtract: $\mathrm{y}^{\mathrm{s}} = y - s_i^{*}$

<u>Seasonality and Additive Model:</u>

There are only **<u>two</u>** differences:

(1) Always ________ instead of divide: $Y_{it}^{S} = (Y_{it} - S_{i}^{*})$

(2) Seasonal factors should average to ____ (not 100).

$$\frac{S_1 + S_2 + S_3 + S_4}{4} = 0$$

Multiplicative or Additive?

Multiplicative: assume fixed proportion of Y_{it} is seasonal.

Example: retail trade turnover (December Versus September)
(generally reasonable for trended data.)

Additive: assumes fixed amount of Y_{it} is seasonal.
(May be appropriate if no trend; real terms.)

Refinements to Ratio-to-Moving-Average" Method

(1) **Trading ____ Adjustments**

- Different months have different number of trading days.
- May differ year after year.

(2) **O________:**

May need to "trim" data to avoid distorting seasonal factors.

⇨ Example: introduction of G.S.T.

(3) **"Evolving" __________ Pattern**

- S_i^***'s** usually allowed to change cyclically over years.

(4) **____ Point Problems**

Calculation of S_i^***'s** may be sensitive to Y_{it} values at end (s) of series.

The Trend –T

An important step in analyzing a time series is an estimation of the trend component, T.

One possibility is to use a "moving average" to smooth out ($T_t \times C_t$) to leave trend.

⇨Not favoured because the length of M.A. (and choice of "weights") arbitrary. Can be distortive.

☺The preferred approach is to identify a time series' trend, and then "model" the trend components:

⇨{Increasing / Decreasing}
⇨{Linear / Non-linear}

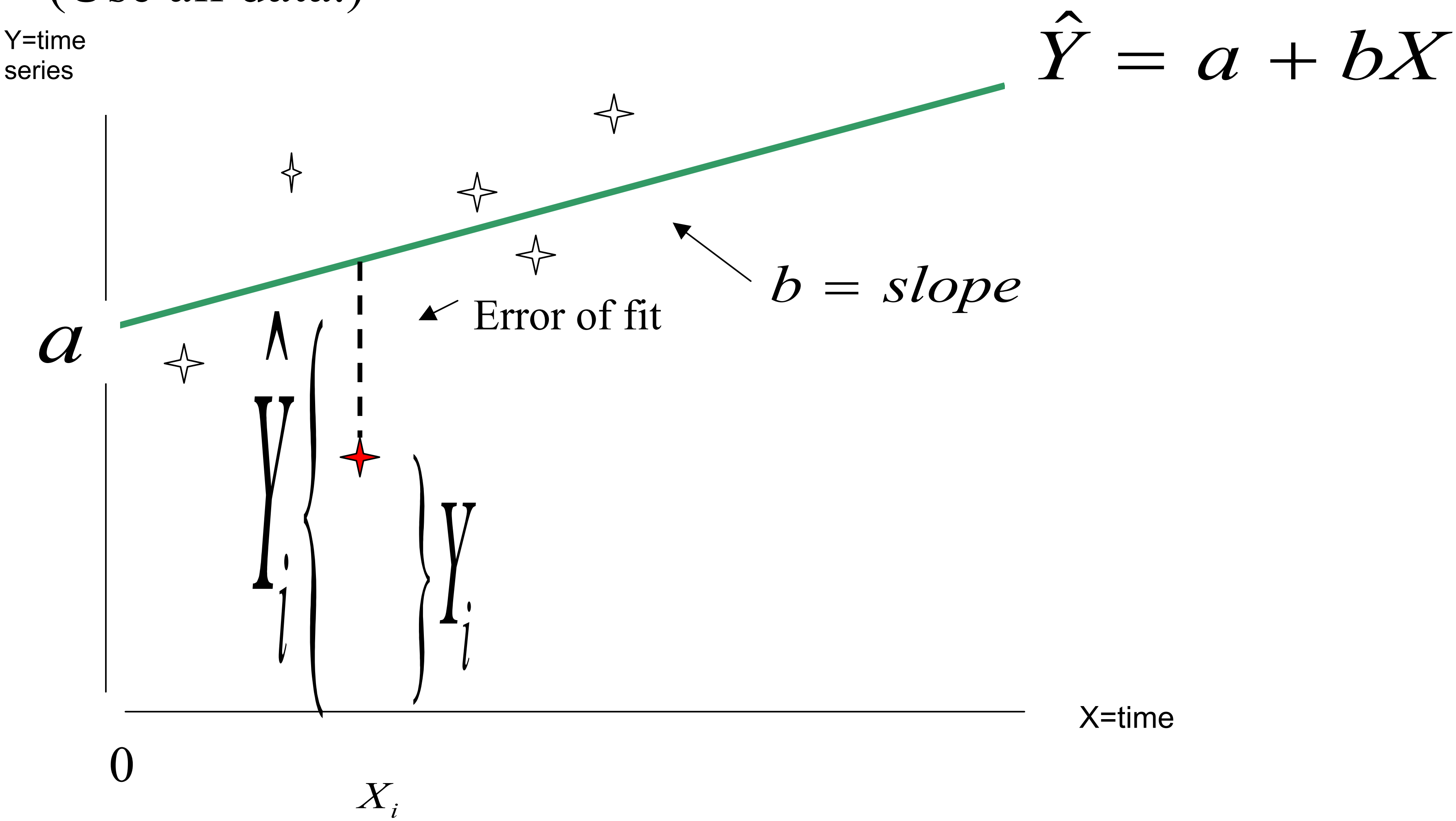
One principle we can use to fit a trend line to de-seasonalized data is the **<u>L_____Squares</u>** principle.

Let: Y_i=i^{th} observation on seasonally adjusted series.
X_i=i^{th} observation on "time."

Example: X = 1, 2, 3, 4, ...
X= 100, 110, 120, 130,...
X=-3, -2, -1, 0, 1, 2, 3, ...

Each observation for data gives us an (X_i, Y_i) point (plot Y_i against X_i).
Try to fit a line through ____ so it is "close" to as many data points as possible.

(Use all data.)


Y=time series
$\hat{Y} = a + bX$
$b = slope$
Error of fit
a
$\hat{Y}_i$
Y_i
0
X_i
X=time

Relationship is $Y_i = \alpha + \beta X_i$

(Not exact due to other components.)

Fitted line is $\boxed{\hat{Y}_i = a + bX_i}$

Error of fit= $\left(Y_i - \hat{Y}_i\right)$ (+/-) ← *Sums to Zero!!*

Place the line so as to <u>minimize</u>: $\boxed{\sum_{i=1}^{T} \left(Y_i - \hat{Y}_i\right)^2}$

(Choose a and b).

That is: $$\underset{a,b}{Min}\left[\sum_{i=1}^{T}\left(Y_i - a - bX_i\right)^2\right]$$

Solution:

$$a = \overline{\mathrm{Y}} - \mathrm{b}\overline{\mathrm{X}}$$

$$b = \frac{\left[\sum_{i=1}^{T}(X_i - \overline{X})(Y_i - \overline{Y})\right]}{\sum_{i=1}^{T}(X_i - \overline{X})^2} = \left[\frac{\left(\sum_{i=1}^{T} X_i Y_i\right) - T\overline{X}\overline{Y}}{\left(\sum_{i=1}^{T} X_i^2\right) - T\overline{X}^2}\right]$$

where T=# of observations.

Example: T=7

Y_i	10	12	12	13	16	17	16
X_i							

$$\overline{X} = 0$$

$$\overline{Y} = 13.7143$$

$$\sum X_i^2 = 28$$

$$\sum X_i Y_i = 32$$

$$b = \frac{32-0}{28-0} = 1.1429$$

$$a = 13.7143 - (1.1429)(0) = 13.7143$$

$$\hat{Y}_i = 13.7143 + 1.1429X_i$$

X_i	Y_i	$\hat{Y}_i$	$\left(Y_i - \hat{Y}_i\right)$
	10	10.2856	-0.2856
	12	11.4285	0.5715
	12	12.5714	-0.5714
	13	13.7143	-0.7143
	16	14.8572	1.1428
	17	16.0001	0.9999
	16	17.1430	-1.1430

<u>Some Properties of Least Squares</u>

(1) The predicted line passes through the _____ of X and Y, $(\overline{X}, \overline{Y})$:

$$\hat{Y}_i = a + bX_i = a + b\overline{X}$$

but,

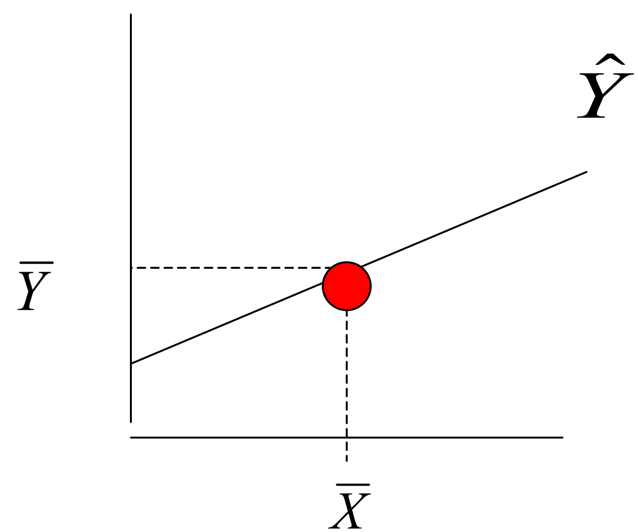
$$a = \overline{Y} - b\overline{X}$$

so,

$$\hat{Y}_i = \left(\overline{Y} - b\overline{X}\right) + b\overline{X} = \overline{Y}$$

when

$$X_i = \overline{X}.$$



(2) Scaling each X_i by a constant leaves $\hat{Y}_i$ __________.

Begin with:

$$\hat{Y}_i = a + bX_i$$

Form:

$$X_i^* = CX_i.$$

Then:

$$b^* = \frac{\left[\sum (X_i^* - \overline{X}^*)(Y_i - \overline{Y})\right]}{\left[\sum (X_i^* - \overline{X}^*)^2\right]},$$

where:

$$\overline{X}^* = C\overline{X};$$

So:

$$b^* = \frac{\left[\sum (CX_i - C\overline{X})(Y_i - \overline{Y})\right]}{\left[\sum (CX_i - C\overline{X})^2\right]} = \frac{1}{C}b$$

Similarly:

$$
\begin{aligned}
a^* &= \overline{Y} - b^* \overline{X}^* \\
&= \overline{Y} - \left(\frac{1}{C}\right) b\left(C(\overline{X})\right) \\
&= \overline{Y} - b\overline{X} = a
\end{aligned}
$$

So:

$$
\begin{aligned}
\hat{Y}_i^* &= a^* + b^* X_i^* \\
&= a + \left(\frac{1}{C}\right) b\left(C(X_i)\right) = a + bX_i \\
&= \hat{Y}_i \Leftarrow \text{ same fitted line.}
\end{aligned}
$$

(3) Adding a constant to each Xi leaves $\hat{Y}_i$ __________.

Begin with:

$$\hat{Y}_i = a + bX_i$$

Then form:

$$X_i^* = (X_i + k);\ \text{for some 'k'}.$$

$$b^* = \frac{\left[\sum (X_i^* - \overline{X}^*)(Y_i - \overline{Y})\right]}{\left[\sum (X_i^* - \overline{X}^*)^2\right]}$$

where:

$$\overline{X}^* = (\overline{X} + k)$$

So,

$$b^* = \frac{\left[\sum (X_i + k - \overline{X} - k)(Y_i - \overline{Y})\right]}{\left[\sum (X_i + k - \overline{X} - k)^2\right]}$$

$$b^* = \frac{\left[\sum (X_i - \overline{X})(Y_i - \overline{Y})\right]}{\left[\sum (X_i - \overline{X})^2\right]} = b$$

$$a^* = \overline{Y} - b^*\overline{X}^*$$
$$= \overline{Y} - b\overline{X}^*$$
$$= \overline{Y} - b(\overline{X} + k)$$
$$= (\overline{Y} - b\overline{X}) - bk$$
$$= (a - bk)$$

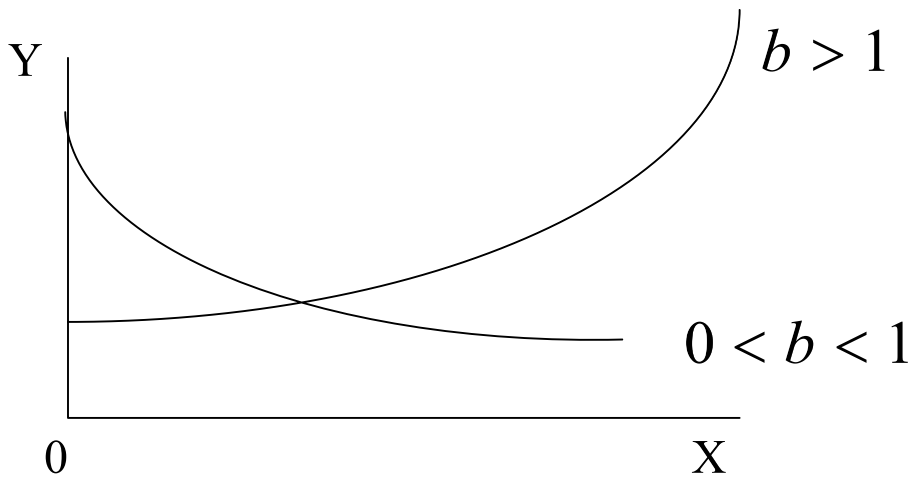
We are free to choose the X_i values just as we wish – the fitted line is unaffected by this choice.

So:

$$\hat{Y}_i^* = a^* + b^*X_i^*$$
$$= (a - bk) + b(X_i + k)$$
$$= a - bk + bX_i + bk$$
$$= (a + bX_i) = \hat{Y}_i \Leftarrow \text{ same fitted line.}$$

Often, a ______ trend is clearly not appropriate. There are several non-linear possibilities.

<u>Consider an Exponential Trend</u>:



$$\hat{Y}_i = ab^{Xi}$$

To make this trend linear in parameters, take the natural log:

$$\ln \hat{Y}_i = \ln a + X_i \ln b$$

or

$$\hat{Y}_i^* = a^* + b^* X_i \Leftarrow \text{ linear in parameters}$$

Obtain a* and b* by least squares, using Y_i^* $(= \ln Y_i)$ and X_i data.

Then, the fitted trend line is:

$$\text{Exp}(\hat{Y}_i^*) = \text{Exp}\left[\ln(\hat{Y}_i)\right] = \hat{Y}_i$$

Example: T=7

X_i	Yi	$Y_i^* \quad (=\ln Y_i)$
-3	1	0
-2	3	1.0958
-1	4	1.3863
0	6	1.7918
1	9	2.1972
2	8	2.0794
3	12	2.4849

$$\overline{X} = 0$$

$$\overline{Y}^* = 1.5769$$

$$\sum X_i^2 = 28; \quad \sum X_i Y_i^* = 10.2272$$

$$b^* = \frac{\left[\sum X_i Y_i^* - T\overline{X}\overline{Y}\right]}{\sum X_i^{\,2} - T\overline{X}^2} = \frac{10.2272}{28} = 0.36526$$

$$a^* = \overline{Y}^* - b^*\overline{X} = 1.5769$$

So,

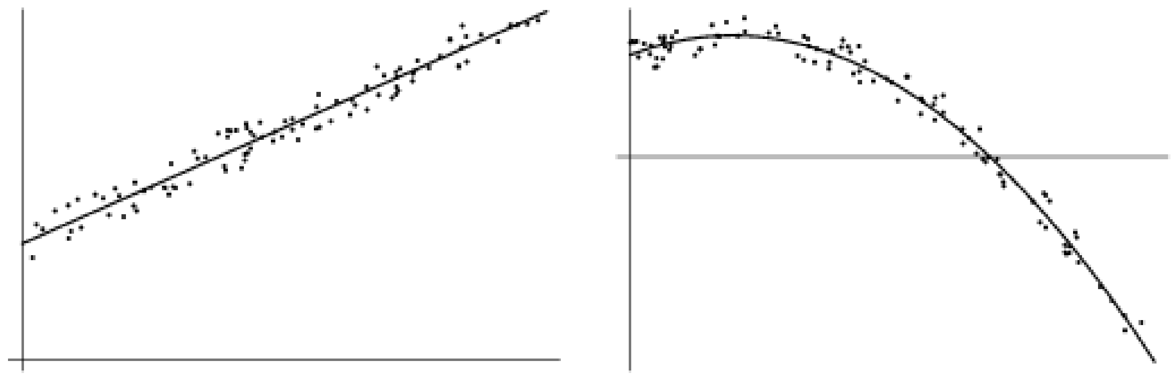
$$\hat{Y}_i^* = 1.5769 + 0.36526X_i$$

or

$$\hat{Y}_i = \text{Exp}\overbrace{(1.5769 + 0.36526X_i)}^{\hat{Y}_i^*}$$

$$\hat{Y}_i = \underset{a}{(4.8399)}\underset{b}{(1.4409)}^{Xi}$$

Xi	Yi	$\hat{Y}_i$	$(Y_i - \hat{Y}_i)$
-3	1	1.6178	-0.6178
-2	3	2.3311	0.6689
-1	4	3.3589	0.6411
0	6	***4.8399***=a	1.1601
1	9	6.9738	2.0262
2	8	10.0489	-2.0486
3	12	14.4790	-2.4790



For nonlinear least squares fitting to a number of unknown parameters, linear least squares fitting may be applied iteratively to a linearized form of the function until convergence is achieved. However, it is often also possible to linearize a nonlinear function at the outset and still use linear methods for determining fit parameters without resorting to iterative procedures. This approach does commonly violate the implicit assumption that the distribution of errors is normal, but often still gives acceptable results using normal equations, a pseudoinverse, etc. Depending on the type of fit and initial parameters chosen, the nonlinear fit may have good or poor convergence properties. If uncertainties (in the most general case, error ellipses) are given for the points, points can be weighted differently in order to give the high-quality points more weight.