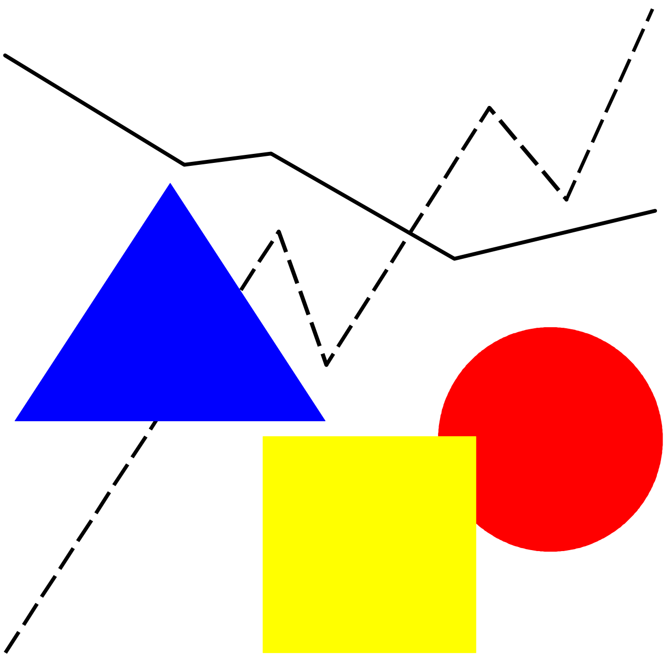
Statistisches Bundesamt

Peter von der Lippe

Chain Indices

A Study in Price Index Theory



Volume 16 of the Publication Series
Spectrum of Federal Statistics

METZLER
POESCHEL

Herausgeber: Statistisches Bundesamt, Wiesbaden

Der Titel ist auch als Buch erschienen.

Verlag:Metzler-Poeschel, Stuttgart

Verlagsauslieferung:
SFG – Servicecenter Fachverlage GmbH
Postfach 43 43
72774 Reutlingen
Telefon:0 70 71 / 93 53 50
Telefax:0 70 71 / 93 53 35
Internet:http://www.s-f-g.com
E-Mail:staba@s-f-g.com

Erschienen im März 2001

Preis: DM 38,80 / EUR 19,84

Bestellnummer: 1030516-01900

ISBN: 3-8246-0638-0

Recyclingpapier aus 100 % Altpapier.

Published by: Federal Statistical Office, Wiesbaden

The titel is also available in the form of a printed publication.

Publishers: Metzler-Poeschel, Stuttgart

Delivery:
SFG – Servicecenter Fachverlage GmbH
PO box 43 43
72774 Reutlingen / Germany
Tel: (+ 49 - 70 71) 93 53 50
Fax: (+ 49 - 70 71) 93 53 35
Internet:http://www.s-f-g.com
E-mail:staba@s-f-g.com

Published in March 2001

Price: DM 38.80 / EUR 19.84

Order Number: 1030516-01900

ISBN: 3-8246-0638-0

100 % recycled paper.

Contents

List of Figures

List of Tables

Foreword

The publication on chain indices as part of the series “Spectrum of Federal Statistics” again shows very clearly the importance and benefits of intensive cooperation between the academic community and official statistics.

The bridging between theory and practice as well as an intensive and frequent exchange of experience with the world of teaching and research are of special concern to the statistical offices.

The way this publication originated is a model illustration of the suitability of that approach. The conditions were favourable in two respects: First, the integration of the theory of chain indices into both the general index theory and the theory of price measurement and, second, the special topicality of that issue in the discussion of harmonising consumer price indices in Europe.

With this compendium, a comprehensive overview of the pros and cons of chain indices is given for the first time. Professor von der Lippe, a committed advocate of the pure price comparison, supports the arguments and procedures of German official price statistics. This contributes to defining positions and clarifying the situation, thus giving rise to the hope that the discussion about possibilities and limits of chain indices will now become more objective.

Wiesbaden, February 2001

Johann Hahlen

President of the Federal Statistical Office

Preface

This book is designed as a sort of pamphlet in which the author tries to convince the reader, as follows:

- the reasons given for the alleged advantages chain indices enjoy over direct indices are not conclusive,
- there are definitely many undeniable shortcomings of chain indices such as difficulties in aggregation and deflation, path dependence and inapplicability of axiomatic reasoning, and
- those who advocate chain indices, not infrequently, have misunderstood and neglected some of the ideas underlying the traditional direct Laspeyres index, ideas of which this book tries to remind the reader.

It is primarily the idea of making pure comparisons, or of comparing "like with like" which is ignored by chainers, but which is on the other hand a cornerstone of index theory, if not of statistics in general. Therefore I think it is time to reconsider carefully the decisions in favor of chain indices, because disadvantages of this type of index formulas abound, and there is no conceptual basis aside from some ingratiating rhetoric in terms of "flexibility", "adaptability" "up–to–dateness", "relevance" or so. It is first and foremost this kind of fashionable talk about indices, which decisively promotes the cause of chain indices and to call this into question is the primary concern of this book.

The history of this book began with some TES – courses (in English) on "Price Index Theory and Price Statistics". TES means "Training of European Statisticians" and is an organization (it became an institute meanwhile) of Eurostat (the Statistical Office of the European Union) with which I cooperated for some years fairly intensively. I was glad that I got so much support from the Statistisches Bundesamt, the German Statistical Office (GSO for short), by sending lecturers who presented the practical part of the topic, while I delivered the theory. I am particularly grateful to Mr. Johann Szenzenstein, Mr. Günther Elbel and Mr. Helmut Mayer from the GSO, as well as to varying lecturers from Eurostat who gave me this valuable support.

In these courses, I also discussed among other things repeatedly the most contentious chain index issue, and I prepared some comprehensive English texts on a number of theoretical aspects of index formulas. Some parts of this collection of papers, esp. parts of chapters 3, 4 (on axiomatics), 5 (on deflation), and in the main chapter 7, which was devoted to chaining forms the basis of this book on chain indices.

The TES courses gave cause to think over chain indices, and all arguments of chainers to be heard also in these lectures (I remember most vividly the discussions with Mr. Walström from Sweden, who was a vigorous fighter for chain indices while I tried hard to fight back). I heard most if not all arguments in favor of chaining, which were in general brought into play. So I had lots of opportunities to give some thought to them. As a result I published some articles in the German language on chain indices, but of course I was told that these articles were written in vain, as the number of those who understand this terribly complicated language is small, and unfortunately is unrelentingly on the wane. So I decided to improve and amend the texts mentioned above, in order to make them fit for an eventual publication which has become this present book.

In this phase I was assisted by some people who particularly deserve being mentioned here. My assistant Mr. Michael Westermann took care of the conversion of the original Word–texts into LaTeX, the formulas, corrections and layout and the final preparation for publication. I also wish to thank my assistant Dr. Andreas Kladroba and Mrs. Ursula Schapals for work at the initial stage of the text. I'm also very grateful to Ms. Nina Schlotter a student assistant of mine who spent a year or so in the United States and therefore was able to check the whole text with respect to language most conscientiously before we gave the text to a native speaker. This part was taken up by Mr. Paul Arthur a lecturer of English at the local Management Academy. I am grateful to him because he gave us the reassuring feeling that the book should be tolerably readable for Englishmen like himself, and for Americans as well. I do know that the work he did was not easy because the contents is difficult to follow for a non–statistican and we can easily find many texts which are more elegantly expressed in English.

I owe once more much to the GSO, which did not only support me with the TES–Courses, but also by incorporating this book into one of its publication series. I wish to thank the president of the GSO, Mr. Johann Hahlen, for taking this decision, and of course also Mrs. Sibylle von Oppeln–Bronikowski, who headed the group responsible for publications, and Mr. Wolfgang Buchwald, the president of the Price Statistics Department, who have both probably done a lot to prepare this decision.

Finally I wish to express my hope that this book will contribute to a discussion of chain indices, and to help the GSO in the pursuit of its index philosophy, and that it will not contain so many mistakes and errors that chainers may ignore this book. In any case there is only myself who should be blamed for errors.

Essen, December 2000

0 Introduction

0.1 The scope and structure of the book

a) The nature and relevance of chain indices

There are two ways of looking at the comparison of two periods, 0 and t, or "binary" temporal comparison with index numbers, the traditional approach (better called "direct" approach, see fig. 1.1.a), and the chain approach.

1. The common feature of all *traditional* (price) index formulas, like the well known formulas of Laspeyres and Paasche is to make a comparison between two periods (or more general, two "states"), 0 and t, where these periods are taken in isolation.

2. The chain principle on the other hand, consists explicitly in taking into account *all intermediate* periods $1, 2, \ldots, t-1$, in order to perform a comparison of the two end points 0 and t. Thus, not only 0 and t, but the whole time series of prices and quantities *in between* will be of interest in constructing a price index.

Therefore a chain index is essentially an index conceived as a measure of the *cumulated effect* of successive steps from 0 to 1, 1 to 2, $\ldots$,$t-1$ to t. Chaining is more a specific type of temporal aggregation , and description of a *time series* rather than a "pure" comparison. Hence for chain indices, as well as for the Divisia index a sequence is essential. They only apply to intertemporal comparisons, and in contrast to direct indices they are not applicable to cases in which no natural order or sequence exists. Thus the idea of a chain–index for example has no counterpart in interregional or international price comparisons, because countries cannot be sequenced in a "logical" or "natural" way (there is no $k+1$ nor $k-1$ country to be compared with country k).
Chain indices are recommended by the revised SNA of 1993. Some National Statistical Offices (NSOs) have already made use of them for a long time, with much enthusiasm, whereas a minority of other NSOs (esp. Germany) disapprove of and dislike strongly this index approach. The topic therefore is not only interesting from a theoretical point of view but is also politically relevant.
In Europe chain indices have already been used for many years, namely in France, the United Kingdom and Sweden, and since recently also in the Netherlands and in Luxembourg. There is a strong tendency in discussions on future practices in official statistics to make them mandatory for all European countries by Eurostat. Many countries have already moved from a traditional direct Laspeyres index to a chain index.
It is therefore most likely that the case is already decided in favor of "chainers", and the faction of "non–chainers" is definitely defeated. So a critique of chain indices as presented in this book is a bit like swimming against the tide. On the other hand such an exercise does not appear to be completely useless. It is never too late for an analysis of methods of statistics, in particular when shortcomings are obvious. Therefore

a position critical of chaining is not necessarily condemned to remain an outsider's position.

b) Aims of the presentation

The objective of this book is to present aspects of chaining with respect to problems both, theoretical *and* practical. We aim at a balanced assessment of chain indices, not concealing our reservations of course, but accounting for both

- index number *theory* considerations related to the justification (foundation) and formal aspects of index functions (formulas) using "axioms" for example, as well as
- problems concerning the practice of price statistics in official statistics has to be dealt with.

The theory of index numbers is dominated by complicated mathematical considerations, concerning index functions apparently in growing sophistication and complexity, such that index theory becomes less and less accessible to "ordinary" economists and statisticians. On the other hand practitioners have to deal with difficulties like how to organize a system of regular compilation of prices, the accuracy of data for an update of "weights", adjustments for quality changes and so on. These type of problems are habitually ignored (or at least treated as less relevant) in index theory, partly because some authors are not aware of them, and partly because it will be hard to express such aspects in an exact mathematical manner (whereby they are perhaps of little or no attraction to mathematicians).[1] Moreover the two groups, index theoreticians on the one hand, and official price statisticians on the other, appear to diverge rather than converge. For statisticians in statistical offices the variety of opinions, "axioms" and "approaches" in index theory is not easy to understand. Hence a comprehensible text might be useful and therefore the book also aims at building a bridge between mathematical theory and practice in statistical offices and ministries. This also implies a kind of presentation which throughout this book is rather moderately mathematical.

Of course the book also (or primarily) attempts to convince the reader of a certain view of chain indices. At this point a **personal remark** might be appropriate.

The present author didn't always take a dislike to chain indices. As many other students of this issue I was originally impressed by some of the arguments set forth for chain indices. I also only gradually became aware of the drawbacks of this approach (see chapters 3 and 5). There are many severe disadvantages of the chain approach, and due to some more or less obvious properties, the performance of chain indices is poor. This is especially true in the field of aggregation and deflation.

[1] No doubt, the choice of an index formula dominates index theory. But it may well be questioned if the formula itself really should occupy the center of interest. Perhaps some of the problems of getting the right data to enter the formula have an even greater effect on the result of index calculation than just the formula.

I tried to find out more of what might be called a "theory", underlying chain indices (see chapter 4) without much success, however. I realized that there is not much going on beyond the fixation with, not to say obsession, up–to–dateness of weights, which largely goes unquestioned. Most of the indices known to me are some kind of weighted average, but I never encountered an average in statistics where the concern about *weights* dwarfs all other considerations in such an amazing manner as in the case of chain index apologetics. In my view index theory should not be reduced to a search for the most up–to–date weights[2] but should first of all strive for a meaningful interpretation.

This is all the more true when indices are compiled in *official* statistics where care should be taken for all attempts to give an index function a clear meaning.

I not only realized that chain indices have little to do with axioms but also that "chain–logic" as the theoretical background of chaining is inconclusive or even contradictory. I gradually realized that chain indices could be viewed as describing the particular pattern of a *time series*, rather than comparing *two situations* as indices in general do, and should do.

Furthermore a closer look at arguments put forward to "prove" the alleged superiority, and the largely non existent "advantages" of chain indices revealed a kind of reasoning with much vagueness, contradiction, and ambiguity (see chapter 6). Surprisingly it turned out that indeed the "arguments usually advanced to substantiate this superiority are not ... very convincing" (SZULC (1983), p. 537).

In essence chaining contradicts comparability in a number of aspects, and I cannot understand why ensuring comparability is taken most seriously when individual goods and services are taken into consideration, whilst on the other hand it is treated most carelessly when commodities[3] are aggregated.

Finally, weight should also be given to practical aspects such as cost, confidence of users, the extra burden of data collection, and that the chain approach clearly facilitates manipulations of all (not only serious) kinds by emphasising "flexibility" at the expense of comparability and meaningful interpretation (see chapter 8). As results of price statistics are used, and interpreted in politics mathematical elegance should not be the only criterion.

Since it is often maintained that the so called "Divisia index" provides a theoretical underpinning to the chain index methodology, the Divisia approach will also be briefly discussed in this book (chapter 7). Both methods, chain indices and the Divisia index have in common that they arrive at a comparison of two points defining an interval in time, say 0 and t by subdividing the interval into a number of adjacent, non overlapping subintervals. They are both a kind of *temporal aggregation*, that is they perform a

[2] There are clearly many other aspects deserving our attention besides "representativity" or "relevance" of weights, and being responsible for what is really measured by an index. Moreover an index using more recent weights is by no means proven to be automatically "better" (see sec. 4.4 below).

[3] Throughout the following text the term "commodities" is frequently used instead of the much longer expression "goods and services".

binary comparison over an interval by aggregating the results of the comparisons of successive subintervals.

To understand the discussion much better, an introductory overview of index theory which unfortunately turned out to be longer than originally intended is given in chapter 2. It not only deals with the axiomatic approach but also with some problems of aggregation and deflation.

Thus the book contains eight chapters in total. Each chapter attempts to be more or less self contained, and is divided into a number of *sections* numbered at two levels. The first level refers to the chapter, and the second to the order of occurrence of the section within the chapter. *Equations* are indicated by three numbers "x.y.z", the first two referring to the section (x.y), and the third to the order of occurrence within that section. Tables, figures, and examples are indicated in the same manner. *Matrices* are indicated by bold face uppercase symbols, and *vectors* by bold face lowercase symbols. Some conclusions, summaries, and noteworthy findings are found in highlighted boxes.

0.2 Introduction to price indices and price statistics

a) Definition of "index"

A price index P_{0t} is a *relative*[4] figure expressing a *comparative* statement (comparing a situation with a *base*–situation) designed to measure a change (or difference) of prices with respect to a complex phenomenon (an aggregate) of *several* commodities. It is a single figure showing how a *set* of n related variables has changed (over time), or differs (from place to place). It is a measure of aggregative comparison, a certain function[5] $P : \mathbb{R}^{kn} \rightarrow \mathbb{R}$ mapping real valued vectors with kn dimensions[6]

$$P_{0t} = P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t) \tag{0.2.1}$$

into a one dimensional positive real number for comparative purposes.

Note that this definition does *not* apply to chain indices nor to the so called "true cost of living index" (COLI) known in the "economic theory" of index numbers.

The function $P(\cdot)$ should satisfy certain functional equations (*axioms*) and have a meaningful interpretation. In case of a *binary temporal* comparison, to which index theory originally was exclusively devoted, it is common practice to use the following notation: P_{0t} (the price *index*) denotes the level of prices at time (period) t, the *current* period (or reporting period), compared to (relative to) time 0, the period known as

[4] The word index is also used in a much wider sense of an *absolute* figure. There are "indices" to represent a summary measure of some more or less complex magnitude (or "construct") like welfare, social status, the business cycle etc. in absolute terms.

[5] To express the plural of "index" there was originally no other English term in use than "index numbers". But it has become more and more common in the last few years to say "indices" or also "indexes" instead of index numbers. Some writers distinguish an "index" (in the sense of "index function" or "index formula") and "index number" which is a specific value of this function.

[6] With two price vectors and two quantity vectors for n commodities we have $4n$ dimensions ($k = 4$).

the reference *base* period (or simply “base”) because it is the basis of comparison in the sense that the level of prices at period 0 is 1 or 100%. Hence the result $P_{0t} = 1.25$ means an increase in prices between 0 and t amounting to 25% (per cent of the level of the base period)[7]. Note that the subscripts [or labels] 0 and t have a definite order, interchanging is not permissible since P_{t0} and P_{0t} definitely are different index numbers.
The word *base* is ambiguous because it may refer to the period to which

1. we compare the current state (*reference base*), or to which
2. the weights refer (*weight base*).

Especially in case of chain indices the distinction between reference base and weight base is of some importance. Some comments on the key words of the definition should be added.

- By *variables* we may understand prices and quantities of commodities bought or sold, wages, industrial production, all (or selected) inputs, or the output of the manufacturing industries etc. The only limitation is that the variables should be positive and measurable on a *metric* scale. Otherwise there will not be any *real* numbers to be combined to an “index”.

- The *set* (or aggregate) of prices can be consumer prices in the case of a consumer price index as opposed to prices of raw materials or wholesale prices. The set is defined only *qualitatively*, i.e. consumer prices refer to a collection of commodities which is said to be representative for the “cost of living”, whereas wholesale or retail prices belong to a set, defined somewhat differently. The type of variables and the definition of the “set” constitutes the *type* (or *name*) of the index number in question, e.g. whether we deal with a production index, an index of new orders, a price index or more specific a consumer–, retail–, agriculture–sales–price index, or agriculture–input–price index etc.

- The term *comparison* needs further specification. As to the number of “situations” to be compared we may distinguish between “binary” and “multi–situations” comparisons. According to the kind of “situations” to be compared we have

  - *intertemporal* comparisons (between periods[8] in time), the topic of primary concern in index theory,
  - *interspatial*, or “place to place” comparison, that is in practice international and interregional comparisons, and

[7] In practice it may be more common to state the result as $P_{0t} = 125$ but multiplying with 100 is only an operation to improve comprehension by users who are not so familiar with statistics.
[8] Intervals like months, years etc., or (in case of stocks) points in time (like dec. 31st 1999).

– *structural* comparisons for example the comparison of index numbers of different type (the well known "terms of trade", i.e. the ratio between export– and import prices) or among various sectoral indices (sub–indices) each of them being part of the same global (overall) index[9].

b) Objectives of price indices, concept of deflation

Price statistics deals with the collection of *individual* prices in *absolute* terms (i.e. expressed in $, £ or so), and with synthetic expressions of "price levels" referring to *aggregates* of goods and services in *relative* terms of indices. For some purposes absolute prices are interesting results as such, but in most parts they are used as raw material only to compile price indices. In intertemporal comparisons price indices have to serve the following purposes

a) to describe the "price level" of an aggregate, i.e. some sort of an average price for a set of commodities on a certain market, and

b) to "deflate" an amount of money representing a purchase or sale of commodities or simply an expenditure (e.g. payment of taxes), that is an estimation of the underlying quantity component or of "real" income (see sec 2.3a for the distinction between these two types of deflation).

In general the same formal requirements have to be met in both applications of index numbers (price level measurement and deflation), but there are also situations in which a conflict may arise, and there is no need to require a single index formula to be able to equally serve both purposes. An aggregate at *current* (period t) prices,

$$V_t = \sum_i p_{it} q_{it}, \tag{0.2.2}$$

is called a *value*, V_t (or "nominal" aggregate) and the same (with respect to the selection of commodities and their quantities) aggregate valued at *constant* prices of the base period 0,

$$Q_t = \sum_i p_{i0} q_{it} \tag{0.2.3}$$

is called a *volume* (or "real" aggregate). The objective of "deflating" an aggregate V_t, or a *value index* (value ratio)

$$V_{0t} = \frac{V_t}{V_0} \tag{0.2.4}$$

is to estimate the volume Q_t which a value represents, given the value V_t, or to derive a quantity index Q_{0t} from a value index V_{0t}.

[9] An example is the comparison of subdivisions of say the Consumer Price Index (CPI), which is usually structured into major categories like food and beverages, housing, clothing, personal care etc.

c) Two basic methodological principles of price statistics

It should be recalled that the relevant *object of observation* in price statistics is *not* merely the *price, but* the whole *act of buying* or selling in which the price is *one* element, but not the only one and perhaps not even the most decisive one. Other factors are the quantity and the quality of the commodity, the shop (outlet) in which the sale takes place, a bonus granted or services rendered in connection with the sale if applicable, arrangements made concerning delivery, availability of spare parts, insurance and others. All these "side conditions" are also elements that have to be constantly observed, and taken into account in price statistics in order to avoid comparisons that are clearly unfair and misleading. We may call them *price determining characteristics* (**PDC**). With this concept it is easy to define the two fundamental methodological principles of data collections on prices and index compilations (fig. 0.2.1):

Figure 0.2.1: Principles of price statistics and price index construction

The need to take price determining characteristics (**PDCs**) into account calls for

representativity	**pure price comparison**
to select contracts governed by the most typical **PDCs***	to have as similar **PDCs** prevailing in both situations to be compared as possible

* Price statistics and index calculations always implies a problem of selection (sampling) among the universe of prices.

> From the general objectives of price indices it follows that care should be taken to arrive at a selection of commodities, that is both representative and sufficiently comparable. Representativity in the context of consumer price indices, means being typical for the consumption pattern a certain household has at the moment or had in the recent past. Ideally the selection in question would be perfectly comparable, if it were kept unchanged in the course of time. In practice the idea of pure comparison would require that, as changes cannot be ignored, their influence should at least be accounted for properly.

It is easy to see, that both requirements are hard to reconcile if conditions change as they actually do. As changes take place more rapidly, the task becomes even more difficult and a compromise is needed.

> The *representative principle* consists in choosing those PDCs that are typical, often met in every day purchases, and that may be looked upon as being the "average" specification of the purchase in question.

The type of selection procedure (e.g. random sample or other), is left open in this context. The number of purchases every day is immense, it is completely impossible to observe them all in any systematic manner, and the contracts would also vastly differ if we would go into more detail describing and classifying them. Thus a statistic of *all* contracts and prices would not only be prohibitively costly, but also of very limited use. Since a selection has to be made it seems reasonable to concentrate on the average, or on "typical" contracts.

> The *principle of pure price comparison* on the other hand requires conditions under which any two price observations are made to be as exactly the same as possible except for time or location, i.e. except the aspect under consideration. The reason is: the rise or decline in a price should not be attributable to other factors than the different time or location of observation.

Comparisons are only called "pure" if different prices can *not* be "explained" by factors (like different quality, type of shop etc.), other than time or place of observation. Note that "pure" is defined negatively: *not* reflecting other factors, *not* affected by structural change etc. Sometimes a really "pure" international comparison can hardly be made, simply because it is impossible to find the same goods or at least "equivalent" goods in the other country. That is the reason why e.g. expenditures for housing, accommodation or other services are often excluded from international price comparisons.

Another problem is the emergence of a new product B, gradually replacing A. The principle of pure comparison requires the continuing observation of the price of A, while in order to comply with the principle of a representative price statistics it would be better to change over to B, at least once B has become the more typical product than A.

As a consequence of the principle of pure price comparison, *adjustments* for quality change have to be made on observed prices in case of a *quality change*, in order to distinguish between a rise in price due to quality improvement, and a "genuine" rise in price, the latter being what price statistics should exclusively attempt to measure.

In practice it is by no means easy to comply with *both* principles *simultaneously*, because if one of them is adhered to most strictly, the other will almost automatically fail. Applied absolutely the two principles will be contradictory. It should be noticed that in practice, in one way or another some *compromise* has to be found, which tries to comply with both conflicting principles in an acceptable manner. Of course no solution can fulfill each of them to the fullest satisfaction.

This also applies to index formulas. There is for example an index concept, like direct Laspeyres indices, in which a lot of attention is given to the idea of "pure" comparison at the expense of the representativity of the basket to which the index refers. On the other hand, to want an index which takes also the most recent changes of the consumption pattern into account would make a "chain index" preferable. But this implies an "impure" comparison in the sense that what is supposed to be a measure of a change in *prices* is to some extent also reflecting a change in quantities and structure. Thus the result has some ambiguity which to avoid is in essence the rationale of the principle of pure price comparison.

Pure price comparison is uncontested in the case of a *single commodity*. Much attention is given to all sorts of problems that may impair comparability, like quality change or nonavailability of similarly specified products in two different countries. There is a vast amount of discussion on how to handle these types of problems. However surprisingly, when applied to a collection, "bundle" or *basket* of *commodities*, the same principle seems to be forgotten or even turned into its opposite. To incorporate the most recent representative basket in index calculation appears to dominate all other considerations. It is generally accepted that each individual commodity should be comparable over time or space, but it is also widely accepted that this can well be neglected in case of a "composite" commodity, that is in the case of indices[10]. The problem is to reconcile two conflicting principles of index construction:

1. pure price comparison over a *long* interval in time (constant weights for many years)
2. representativity of weights for *any* subinterval.

A reasonable compromise is in our view not a chain index (focusing on 2 at the expense of 1) nor unchanged weights of a Laspeyres index (pursuing exclusively principle 1) but a Laspeyres index in which weights are reviewed and readjusted in intervals of say five years or so.

d) The variety of "approaches" in index theory

Mostly index theory has to do with the assessment of index formulas. It is perhaps a fruitless effort to count all the different index formulas that have been invented, and proposed as being the "best" or "ultimate" formula for the tasks of price level measurement, and deflation of aggregates. Furthermore there are different *approaches* to derive (or justify) index formulas, and the most important distinction is perhaps between

a) the "economic theory" approach, which aims at a microeconomic foundation (in terms of utility) of index formulas, and

[10] Obviously this remark refers to chain indices. They are vigorously recommended in the SNA 93 (System of National Accounts revision 1993). On the other hand the authors of SNA rejected so called "unit value" indices which "measure the change in the average value of units that are not homogeneous, and may therefore be affected by changes in the mix of items as well as by changes in their prices. Unit value indices cannot therefore be expected to provide good measures of average price changes over time" (SNA (1993), no.16.13). It did not occur to them that the same criticism would apply to chain indices.

b) the "formal" (or "axiomatic") approach, focusing on *criteria* (desirable mathematical properties, or "axioms" of index functions) useful to evaluate index formulas.

The chain index and the Divisia index approach do not fit well into this structure as they neither provide a new "theory" nor a new systematic evaluation in terms of "axioms". The dominant idea here is simply

- to partition an interval into a number of discrete subintervals (or to stipulate "time" as a continuous variable), such that a binary temporal comparison is made in many steps rather than in one step only, as implied in the definition of eq. 0.2.1, for which the axiomatic theory of index numbers is designed, and

- to make use of the most "recent" and "representative" weights.

Thus for an assessment of this approach the ideas of for the axiomatic theory are strictly speaking not applicable. There is no economic reasoning leading directly to the chain approach either. On the other hand it should be noted that "axioms" as tools to enable users to decide on which formula to apply also have some shortcomings. At least some of these axioms are *open to debate*, they are not selfevident, and the motivation of many axioms is solely a mathematical elegance[11]. Axioms are often in conflict with one another, and we cannot expect to be able to "prove" in such a situation which criteria should come first and which should come second and therefore could safely be sacrificed.

0.3 Simple comparisons (relatives) and aggregative comparisons (indices)

a) Simple comparison (a single commodity)

Consider a time series of prices for a certain commodity i ($i = 1, 2, \dots, n$), i.e. a sequence of prices recorded for the same commodity in successive equally spaced time intervals, $p_{i\tau}$ ($\tau = 0, 1, \dots, t$) or simply $p_0, p_1, \dots, p_t$. The *price relative*, a^i_{0t} or simply a_{0t} to compare prices of commodity i between two periods (e.g. months or years), 0 and t is defined as

$$a_{0t} = \frac{p_t}{p_0}. \tag{0.3.1}$$

[11] Some of the recommendations of the SNA seem to be best viewed in this way. There is nothing "wrong" in aiming at more mathematical elegance. The question is, however, whether this gain should be at the expense of other criteria, like understandability, comparability of results, ease to provide empirical data, "acceptance" by non–statisticians and the like.

Relatives are also called *simple* index numbers[12] or *elementary* index numbers in contrast to index numbers (sometimes also called *aggregative* index numbers). A price relative is a pure number, free of any unit of measurement. The word “relative” makes clear that the price is not measured in absolute terms, but relative to the price of the same commodity at the base period. In the same manner a quantity relative b_{0t} and a value relative c_{0t} is defined as follows:

$$b_{0t} = \frac{q_t}{q_0} \tag{0.3.2}$$

and

$$c_{0t} = v_{0t} = \frac{v_t}{v_0} = \frac{p_t q_t}{p_0 q_0} = a_{0t} b_{0t} \tag{0.3.2a}$$

where q denotes the quantity and v the value of the commodity i (note that the sum of values over all $i = 1, \dots, n$ commodities will be called $V = \sum v$). Price, quantity and value relatives have the six axioms as listed in tab. 0.3.1 in common. They are fulfilled trivially when applied to relatives but applied to index numbers, that is to $n \geq 2$ commodities

- at least some of them may pose problems, and will not be satisfied, and
- there are some axiomatic requirements to add, like mean value property or proportionality which make sense only if the number of commodities is $n \geq 2$.

The first three axioms (identity, dimensionality and commensurability) are minimum requirements to be met also by any index function in order to avoid ambiguity.

By axiom 1 (*identity*) a clear meaning is given to $a_{0t} = 1$ as opposed to $a_{0t} \neq 1$. Otherwise a relative would lack a kind of reference point. Identity provides a normalisation condition to decide over the direction of change.

By axiom 2 (*dimensionality*) the calculation is independent of the currency in which the prices are expressed. Dimensionality is also called “homogeneity of degree 0 in the prices”.

By axiom 3 (*commensurability*) independence is also given with respect to the arbitrary choice of units (e.g. quarts, gallons, pounds, ounces etc.) to which the price quotation refers (it is of course something different if the price of a quart or of a gallon is referred to).

[12] A price relative (like an index) is often also expressed as a percentage requiring the multiplication by 100. Following the definition of indices introduced above we also refrain from multiplying by 100 in defining relatives.

Table 0.3.1: Axioms satisfied by price and quantity relatives

no.	axiom	definition and interpretation	
1	identity	$a_{00} = a_{tt} = 1$	uniqueness of the reference point
2	dimen-sionality	$\frac{\lambda p_t}{\lambda p_0} = \frac{p_t}{p_0}$	a_{0t} is independent of the currency in which the prices are expressed[a]
3	commen-surability[b]	$\frac{\lambda p_t}{\lambda p_0} = \frac{p_t}{p_0}$	independence of the unit of quantity to which the price of commodity i refers[a]
4	time reversal test	$a_{t0} = \frac{1}{a_{0t}}$	consistency of relatives with different base periods
5	factor reversal test	$c_{0t} = a_{0t}b_{0t}$	the value change is decomposable in price change and quantity change
6	transitivity	$a_{0t} = a_{0s}a_{st}$	for all three periods 0, s and t

[a] in both periods, 0 and t.
[b] if only $n = 1$ commodity is involved the mathematical representation of 2 and 3 cannot be distinguished.

Most index formulas satisfy these minimum requirements as relatives always do but the remaining three axioms are often *not* met.[13] By *transitivity* (or chainability) is meant that the relative change as measured for a long interval from 0 to t should be the same as the corresponding change measured over a sequence of sub–intervals and "linked" (multiplied) together. This should be true for *every*[14] subdivision of the interval into two or more adjacent subintervals, like from 0 to s and from s to t. This principle sounds reasonable and is automatically fulfilled in case of relatives, but applied to indices it is responsible for much debate and much confusion. Transitivity may be regarded as a kind of consistency in temporal aggregation (over intervals in time)

An alternative to relatives with a *fixed base* value (p_0 in the definition of a price relative a_{0t} and q_0 in b_{0t}) is a *chain based* comparison, using "links" to be multiplied (chained)

$$l_t = \frac{p_t}{p_{t-1}} = a_{t-1,t}. \tag{0.3.3}$$

[13] See chapter 2 for a discussion of the time reversal test and the *factor reversal* test.

[14] It is important to underline this point with much emphasis. There are many misunderstandings with respect to "chain indices" that simply result from not giving much attention to the word "every", in this context. For different subdivisions of the same time interval for example a chain index will usually *not* yield the same result.

"Chaining"[15] means to represent a_{0t} as a product of *link relatives*, like links of a chain, because

$$a_{0t} = l_1 l_2 \ldots l_{t-1} l_t. \tag{0.3.4}$$

A link relative is a growth *factor*, closely related to the growth *rates* of a price defined by

$$r_t = \frac{p_t - p_{t-1}}{p_{t-1}} = \frac{\Delta p_t}{p_{t-1}} = \frac{p_t}{p_{t-1}} - 1 = l_t - 1$$

Another alternative to relatives would be to use so called log–changes

$$Da_{0t} = \ln\left(p_t/p_0\right), \tag{0.3.5}$$

which can also be viewed in combination with the *logarithmic mean* L of two positive numbers, p_t and p_0

$$L(p_t, p_0) = \frac{p_t - p_0}{\ln(p_t, p_0)} = L(p_0, p_t)$$

as a kind of growth rate since

$$Da_{0t} = \frac{p_t - p_0}{L(p_t, p_0)}. \tag{0.3.6}$$

b) Aggregative comparison (two or more commodities)

Indices are different from relatives mainly because of aggregation problems. The transition from $n = 1$ to $n \geq 2$ changes the situation fundamentally[16] in that

- both, "price" and "quantity" have to be redefined: applied to a "bundle" (or "basket") of commodities, and their meaning can no longer be the same as in case of a single commodity;
- aggregation (across commodities) will destroy many useful properties simple relatives will be satisfied by definition;

[15] Some authors call the operation of multiplying (single commodity) link relatives "linking" as opposed to "chaining" which is the same operation applied to indices ($n > 1$ commodities).

[16] Such problems give rise to a special statistical theory of price index numbers not being simply a part of the theory of means. We may even say, that index theory in general is that part of statistics that deals with the measurement of "a complex that is made of individual measurements for which no common physical unit exists" as Ragnar Frisch put it (quoted by ALLEN (1975), p. 5).

- a structural problem, unknown to the single commodity case emerges and leads to questions like: to what extent does a price index measure the increasing *level* of prices, and to what extent does it also reflect a change in the *structure* of prices or in the consumption pattern?[17]

In case of a single, "homogeneous" commodity i ($i = 1, 2, \ldots, n$) "quantity" is an amount of physical units q_i (as measured in tons, square inches, gallons, or hours in case of services etc.). This concept is also used in economic theory, e.g. in demand and supply curves. However when *different* types of goods and services are to be summated, "quantities" cannot be aggregated in physical units. In contrast to a sum of prices summing over quantities in absolute terms may not even be feasible. In general the sum $\sum q_{it}$ or $\sum q_{i0}$ is not defined due to different units of measurement (by contrast to $\sum p_{i0}$ or $\sum q_{it}p_{i0}$). The idea in representing the aggregative "quantity" by a "volume" $\sum q_{it}p_{i0}$ as a proxy (because the concept is otherwise not measurable) of some kind of "global quantity" is the following: If prices (and we should add: also qualities) were constant, the amount of money $\sum q_{it}p_{i0}$ represents (by contrast to $\sum q_{i0}p_{i0}$) is affected only by changes in quantities. Thus the volume–change[18] should equal the "real" quantity–change *under these circumstances*, that is *if* prices (and qualities) are held constant. Hence it is the (imputed) *constancy of prices that makes aggregation of quantities possible*. This should be recalled in case of the interpretation of "volumes" gained by using chain price indices as deflators (see chapter 5).

Aggregation over commodities is meaningfully possible only in relative terms. Another problem unknown in the single commodity case is structural change. Once $n \geq 2$ commodities are involved (and a non–zero variance of price relatives is given)[19] it will be necessary to distinguish between a change of the price *level* on the one hand, and a change of the price level *and* the *structure* of prices on the other hand (the same applies to quantities). A key problem of index theory appears to be the *elimination of structural change in aggregates* in order to provide a pure price comparison, or pure quantity comparison.

It is only in the presence of this structural change that most of the different index formulas will yield different results. Thus different index formulas can also be viewed as different ways *to deal with structural change by some kind of standardization*. The need for such a standardization and the significance of structural aspects in aggregation can easily be demonstrated in the case of an index of wages (see ex. 0.3.1)

[17] In the single commodity case there is no problem in separating and isolating these two aspects, and in defining a "pure" price–level–component in contrast to a mixed (price/quantity) component or a structural component.

[18] Volume change is measured by a quantity index. A "quantity–index" might also be called a "volume index" (a term which is highly ambiguous, however).

[19] In case of zero variance, almost all index formulas would yield the same result. A problem of selecting the appropriate index function only arises if there is some amount of dispersion, i.e. if there is some *structural* change involved.

Example 0.3.1

The task is to compare the average wage for two periods. Imagine an economy with only two industries A and B, and wages of $10 and $16 paid at base period:

Situation in base period			
industry	*wage*	*hours*	*payment*
A	*10*	*50*	*500*
B	*16*	*50*	*800*
*sum**	*13*	*100*	*1300*

* *or average*

It would be highly misleading to simply compare the average wage per hour presently paid with the average wage formerly paid at the base year. Assume two alternative situations (presented for demonstrative purposes) in t*:*

Situation in t			
industry	*wage*	*hours*	*payment*
A	*15*	*90*	*1350*
B	*24*	*10*	*240*
total	*15.9*	*100*	*1590*

alternative situation in t			
industry	*wage*	*hours*	*payment*
A	*15*	*10*	*150*
B	*24*	*90*	*2160*
total	*23.1*	*100*	*2310*

The average wage (price) per hour changed from 13 to 15.9 (hence grew by 22.3% only) or in the alternative situation from 13 to 23.1 (by 77.69%) although both industries experienced exactly the same rise in wages by 50%. Thus the average wage obviously does not reflect the rise in wages (at an equal rate of 50% in both branches) correctly. The reason is that the structure of employment has changed in favor of the low–earning industry or alternatively the high–earning industry. ◀

An *index* of wages (which is formally nothing else than a price index), should provide weights for the wages, derived from the *same* employment structure in both situations 0 and t in order to show the true ("pure") rise of the wagelevel[20], which is undoubtedly 50%. Any wage index not showing a rise of 50% does not appear reasonable.[21] Sums or simple averages, like values $V_t = \sum p_{it}q_{it}$ or the "unit value" $\tilde{p}_{it}$

$$\tilde{p}_{it} = \frac{\sum p_{it}q_{it}}{\sum q_{it}} = \sum p_{it}\frac{q_{it}}{\sum q_{it}} \tag{0.3.7}$$

provided the sum $\sum q_{it}$ is defined as an indicator of the "average price" are inappropriate to compare aggregates, because they are sensitive to a change in structure[22].

[20] In this case it is irrelevant whether the Laspeyres– or the Paasche type of wage index is used since the growth rates of wages are the same in both industries.

[21] This tacitly requires a measure which satisfies *proportionality*: if all prices (wages) rise by 50% the index should not have any other result than 50%.

[22] The value (sum) and the unit value (average) of say automobile–output can rise only because of a switch from low to high priced models, even though the total number of automobiles (quantity) as well as the volume (quantity valued at *constant* prices) produced remains unchanged.

> Structural change has to be eliminated by some method of weighting. Weighting with a constant base year structure (like in a Laspeyres type wage index) can be viewed as a kind of standardization (elimination of structural change).

By the same token (elimination of structural change[23]) it appears sound to take *volume* instead of value as indicator of quantity. Chain index methodology consists largely as a denial of this standardisation idea. Comparisons are made which reflect pure price movement *and* structural change as well, simply because updating of weights is deemed more important than comparability.

[23] In ex. 0.3.1 due to differential movement of quantities (hours worked), in what follows due to differential price movement.

1 Towards a clear notion of a chain index

As can easily be seen, much controversy about advantages or disadvantages of the chain index design is owed to ambiguities in the relevant concepts[24]. Different views concerning the chain approach often result from misunderstandings of the notion "chain index". Therefore it is of necessity to start with some terminological clarifications, and with a detailed definition of chain indices.

1.1 Direct and chain approach, chain and links

A distinction between "chain–" and "fixed based" indices is often made. From this we might conclude that a chain index has a different kind of "base" than the "ordinary" index formula of Laspeyres or Paasche. Another distinction widely used is between "fixed weighted" index and chain index. This again is misleading. It is for example consistent with the distinction between "fixed–" and "chain–weighted" to call $P^P_{0t} = \sum p_t q_t / \sum p_0 q_t$ a "fixed weighted" Paasche index[25] but the weights q_t vary. As t proceeds from 0 to 1, to 2 etc. weights are q_1, q_2 etc. There is no point in calling P^P_{0t} "fixed" weighted in contrast to "chain" weighted, when the weights of P^P_{0t} are updated no less frequently than the weights of a chain index.
More importantly, the distinctions introduced so far are degrading from the traditional approach to index numbers, and they aid and abet misunderstandings concerning the chain index approach. As everything changes and perhaps these days more rapidly than before the term "fix" or "fixed" raises some negative emotions. In addition this terminology might promote some (erroneous) ideas, like chain indices

- having solved the problem of choosing a suitable base period, and
- incorporating the most recent weights only[26].

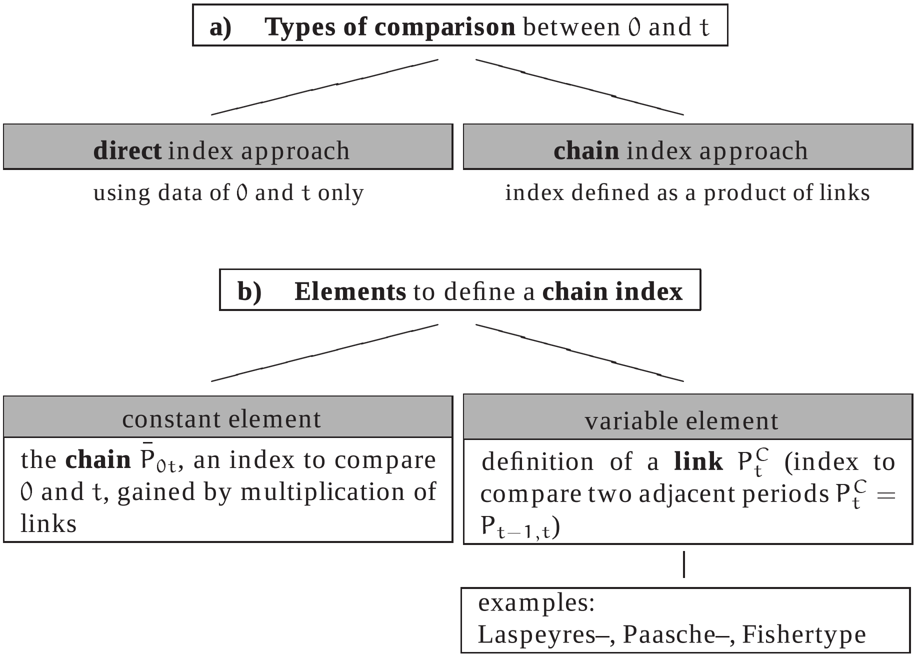
Therefore we will henceforth distinguish between a **direct** and a **chain** approach to compare the two periods, 0 and t. To make clear that the results will in general be different, we introduce the symbol $\bar{P}_{0t}$ to denote a chain index in contrast to P_{0t} which is the corresponding direct index.
What really matters is not the type of "base" or "weight", but that the chain principle consists explicitly in taking into account the *intermediate* periods $1, 2, \ldots, t-1$ and not only the two end points 0 and t.
A chain index essentially is an index conceived as a measure of the *cumulated effect* of successive steps from 0 to 1, 1 to 2, ..., $t-1$ to t.

---

[24] Another source of confusion is that axioms as evaluation–criteria, known in the context of direct indices, do not exist in the case of chain indices. This will be shown in sec. 3.2 .
[25] This will be called "direct" Paasche index later, as opposed to the Paasche chain index.
[26] The common prejudice of course is that such weights are the "best" weights and chain indices thereby measure a change of prices more "accurately", and more in line with the relevant consumption pattern than all other indices. See sec. 4.4 .

As indicated already chaining is a more specific type of temporal aggregation and description of a time series rather than a comparison of two states taken in isolation.

Figure 1.1.1: Terminological distinctions referring to chain indices



Some authors regard this in itself as an advantage of the chain approach. If this really is convincing, will be discussed later. In the framework of the "direct approach" on the other hand, the relevant periods of a price index P_{0t} are 0 and t only. The result of the direct index, P_{0t} is not influenced by what happens in the intermediate points in time. In other words the whole period from 0 to t is regarded as one single step in the direct approach, in contrast to a sequence of steps in the chain approach.
To sum up the idea of fig. 1.1.1a: The distinction between "direct" and "chain" refers to whether or not the result of the index comparing two periods, 0 and t is influenced by how prices and quantities develop in the intermediate periods $1, 2, \ldots, t-1$, that is by the "links" connecting 0 with 1, 1 with 2, ...,$t-1$ with t. It is interesting to note some striking features of the chain approach at the outset of our discussion:

> There always are necessarily **three sources of variation** responsible for the result we get in the case of a chain index:
> 1. the difference in prices in t as compared with 0,
> 2. the change of weights (quantities) have undergone (in response to a change in prices) in comparing 0 and t first, and
> 3. prices and quantities in the intermediate points in time, that is in 1, 2, $\ldots, t-1$.

The **first source** of variation refers to what a price index (as opposed to a quantity index for example) primarily has got to measure. In essence the idea of "pure" is to make index calculation dependent on this source only.

The **second source** of variation accounts for the change of quantities ($q_0 \to q_t$) in response to changes in prices. In view of the idea of a "substitution bias" and the "economic theory", making use of weights related to *both* periods, 0 and t, an index should also reflect this (second) aspect. Hence on the face of it chain indices and some direct "superlative" indices[27] have the first and the second source of variation in common. The difference between them lies primarily in the third source of variation reflecting the impact of past prices and quantities, on present (or most recent) ones. But there is no explicit "economic theory" model of that type of substitution behavior implied in the chain approach, and by which chain indices differ from some direct "superlative" indices.

Interestingly there is a lot of discussion about why and in which way a chain index is better than a direct index, exclusively using base year weights (like in the case of direct Laspeyres indices), but there is little or no reason given for their superiority over those direct indices which also take current weights into account.

The **third source** of variation clearly is something specific of chain indices. A chain index is accounting for the *path* leading to the final result, thereby reflecting the specific shape of the time series of prices and quantities, its own "history". With the exception of the Divisia index no other index type reflects not only the difference between two situations, but also the specific path taken in order to get from 0 to t. Yet this seems to be an unintended by–product of chaining rather than a target of price level measurement or a requirement justified on theoretical grounds.

Moreover the operation of chaining (multiplying) is not motivated by some causal model explaining how present prices evolve from past prices but by aiming at consistent temporal aggregation (see sec. 4.1). Interestingly this consistency (that is "chainability" or "transitivity") is just the opposite of "path dependency", the latter being characteristic for both approaches, chain indices and the Divisia index.

To define the notion of "chain index" carefully and to avoid much of the confusion met in this field we should introduce the following *two* "elements" of the definition (see fig. 1.1.1b):

[27] As for example the (direct) formulas of Fisher, Törnquist or Walsh to be discussed later (see for example fig. 1.3.2.

1. A constant element is the "chain" $\bar{P}_{0t}$, which always is a product of "links" P_t^C, each of which being a direct index comparing t with the preceding period $t-1$, and
2. a variable element, the "link"[28] $P_t^C = P_{t-1,t}$ as there are numerous solutions we might think of; in this sense we might distinguish between Laspeyres– (P_t^{LC}), Paasche– (P_t^{PC}), Fisher– (P_t^{FC}) or other chain index numbers, according to the type of "links" that are multiplied to obtain the chain.

It is useful to introduce the symbol $\bar{P}_{0t}$ (or $\bar{P}_{0t}^C$) in order to make clear that $\bar{P}_{0t}$ can differ from the result of a direct comparison denoted by P_{0t} and to distinguish the two elements, that is the link, P_t^C, and the chain, $\bar{P}_{0t}$. There are two important reasons to insist on a distinction between *two* elements of the definition:

- A closer look at arguments put forward in favor of chain indices (see chapter 6) reveals that most of them result from calling "chain index" what in fact is "only" the link, and from comparing a link P_t^C with P_{0t} instead of correctly comparing the chain $\bar{P}_{0t}$ with P_{0t}.
- There is some discussion about which formula of the link to prefer, completely ignoring that the properties of the chain cannot in general be concluded from properties of the links.

Note that a link is an index comparing a period t with another (the precedent) period in the usual manner. *Only the link is an index* in the sense of satisfying or violating certain "axioms" of the type discussed in sec 2.2. But a chain is something different. It is *not* an index and can violate axioms, despite being "made" of links that satisfy them all. Hence we arrive at the following definition of a chain index in two steps:

First step:
All types of chain price[29] indices $\bar{P}_{0t}$ (as opposed to direct indices) have in common, that a binary comparison of two periods, 0 and t, being two or more periods apart from one another, is gained by the multiplication of "links":

$$\bar{P}_{0t} = P_1^C P_2^C \dots P_t^C = \prod_{\tau=1}^{\tau=t} P_\tau^C. \tag{1.1.1}$$

This may be called the constant, defining, constituent feature of the chain approach.

Second step:
The variable element on the other hand is the specific formula of the link. The Laspeyres link, as an example is defined as follows

$$P_t^{LC} = P_{t-1,t}^L = \frac{\sum p_t q_{t-1}}{\sum p_{t-1} q_{t-1}}, \tag{1.1.2}$$

[28] Since a link always compares the reference period t with the *preceding* period $t-1$ there is no need for two subscripts. It is sufficient to use only one subscript, t.
[29] The definition applies mutatis mutandis also to quantity indices.

so that a Laspeyres chain is by definition

$$\bar{P}^{LC}_{0t} = \bar{P}^{L}_{0t} = P^{LC}_{1} P^{LC}_{2} \dots P^{LC}_{t} \tag{1.1.2a}$$

Other chain indices are defined correspondingly by defining other "links", as for example the Paasche link to construct a Paasche chain

$$P^{PC}_{t} = P^{P}_{t-1,t} = \frac{\sum p_t q_t}{\sum p_{t-1} q_t} \tag{1.1.3}$$

or a Fisher link to construct a Fisher chain index which is the product of the following links

$$P^{FC}_{t} = P^{F}_{t-1,t} = \sqrt{P^{LC}_{t} P^{PC}_{t}}. \tag{1.1.4}$$

1.2 Product representation and chainability

Due to multiplication the chain $\bar{P}_{0t}$ is in general a function of the price and quantity vectors $\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_1, \mathbf{q}_1, \dots, \mathbf{p}_{t-1}, \mathbf{q}_{t-1}, \mathbf{p}_t, \mathbf{q}_t$, and not only of the first and last pair of vectors (eq. 0.2.1 does not apply). Hence it should be unlikely that applying the chain index approach and the traditional direct approach to the same data will yield the same results. But note that the existence of a product as such is not sufficient to characterize a chain index. To see this compare the Laspeyres *chain* index $\bar{P}^{LC}_{03}$ which is

$$\bar{P}^{LC}_{03} = \left(\sum \frac{p_1}{p_0} \frac{p_0 q_0}{\sum p_0 q_0}\right) \left(\sum \frac{p_2}{p_1} \frac{p_1 q_1}{\sum p_1 q_1}\right) \left(\sum \frac{p_3}{p_2} \frac{p_2 q_2}{\sum p_2 q_2}\right) \tag{1.2.1}$$

with the direct Laspeyres index P^{L}_{03} given by

$$P^{L}_{03} = \left(\sum \frac{p_1}{p_0} \frac{p_0 q_0}{\sum p_0 q_0}\right) \left(\sum \frac{p_2}{p_1} \frac{p_1 q_0}{\sum p_1 q_0}\right) \left(\sum \frac{p_3}{p_2} \frac{p_2 q_0}{\sum p_2 q_0}\right)$$

$$= \sum \frac{p_3}{p_0} \frac{p_0 q_0}{\sum p_0 q_0}. \tag{1.2.2}$$

Note that the factors in eq. 1.2.2 are *not* the "ordinary" indices, P^{L}_{01}, P^{L}_{12} and P^{L}_{23}, because P^{L}_{0t} is *not* transitive. The right hand side of eq. 1.2.2 is a sequence of *rebased* Laspeyres indices[30]:

$$P_{01} = \frac{P_{01}}{P_{00}} = \frac{\sum p_1 q_0}{\sum p_{01} q_0}, \quad P_{12(0)} = \frac{P_{02}}{P_{01}} = \frac{\sum p_2 q_0}{\sum p_1 q_0},$$

[30] It is surprising that Forsyth and Fowler (1981) concluded from an inspection of eq. 1.2.1 and 1.2.2, that direct indices were transitive, and chain indices not. For them there is a (false) alternative "transitivity versus representativity", direct indices being transitive and chain indices representative respectively. Their misunderstanding was that in eq. 1.2.2 we don't have indices P_{12} and P_{23} as factors (links) but rebased indices $P_{12(0)}$, $P_{23(0)}$. In fact both types of indices, the direct and the chain Laspeyres index are *not* transitive.

and

$$P_{23(0)} = \frac{P_{02}}{P_{02}} = \frac{\sum p_3 q_0}{\sum p_2 q_0}.$$

The difference between eq. 1.2.1 and 1.2.2 in the first place is that in P^L_{03} *only* the prices are updated and not the quantities (which appears less reasonable) whereas in $\bar{P}^{LC}_{03}$ *both*, prices and quantities are updated continuously (apparently economically more meaningful). With quantities remaining constant there would obviously be no difference between P^L_{03} and $\bar{P}^{LC}_{03}$. With quantities reacting to prices, however, it is not astonishing that $\bar{P}^{LC}_{03}$ will in general differ from P^L_{03} which is known as *drift* of the chain index (see sec. 3.3). The term "drift" does not mean that the incorrect chain index is drifting away from the correct direct index. We may as well think of the direct index drifting away from the (correct) chain index. Note that the Paasche chain index is given by

$$\bar{P}^{PC}_{03} = \left(\sum \frac{p_1}{p_0}\frac{p_0 q_3}{\sum p_0 q_3}\right)\left(\sum \frac{p_2}{p_1}\frac{p_1 q_3}{\sum p_1 q_3}\right)\left(\sum \frac{p_3}{p_2}\frac{p_2 q_3}{\sum p_2 q_3}\right). \qquad (1.2.3)$$

There are some points worth being stressed concerning the interpretation of eqs. 1.2.1 to 1.2.3:

Calculation, accuracy
Note that P^L_{03} can be calculated in two ways by multiplication *and* directly. In calculating $\bar{P}^{LC}_{03}$ there is no such choice. A chain index $\bar{P}_{0t}$ is not only *defined* as a product, multiplication is also the only way to compile $\bar{P}_{0t}$.
The implication of this is significant: $\bar{P}_{0t}$ depends on $\bar{P}_{0,t-1}$, $\bar{P}_{0,t-2}$, and so on, it is a *dependent*, or *indirect* comparison of t with 0. On the other hand P_{0t} is independent of $P_{0,t-1}$ etc. in the sense that the results of $P_{0,t-1}$, $P_{0,t-2}$ have no influence on P_{0t}. Each element of the sequence P_{0t} is an independent or "direct" comparison. This has some significance when chain indices and direct indices are compared in terms of accuracy (see sec. 3.7).

General and special model of an index
A comparison of eqs. 1.2.1 and 1.2.2 gives rise to widely different interpretations. According to non–chainers (or defenders of the direct Laspeyres index) eq. 1.2.2 shows that weights are not really constant but up–dated at least as far as prices are concerned. On the other hand the comparison of the two equations demonstrates for chainers, that a chain index is the more general approach

- if quantities are not changed in eq. 1.2.1 we simply get eq. 1.2.2 as a special case,
- hence chaining operates only when quantities are adjusted, and when this is not done we still have the direct Laspeyres index we are used to.

The emergence of this argument owes much to the painful experience made in the European harmonisation project where the chain issue turned out to be highly controversial and emotional (MAKARONIDIS (1999)). The idea of an all embracing *general*

concept from which both chain indices and direct indices might be derived as special cases was therefore welcomed as calming and reconciling.
The problem of this kind of reasoning is the choice of the product representation as a starting point. There seems to be no more of a general approach than chain indices (they *are* already the alleged general concept) and a direct index appears only as a badly justified (all quantities held constant) special case; but why constant? Once we adopt the chain index as the general concept why should we keep quantities constant?[31] In chapter 7 we will show that it is of limited use to compare two index functions on the basis of an equation which has a widely different significance for the main ideas of the index in question. The rationale of P^L_{0t} is based on a directly calculated expression $\sum p_t q_0 / \sum p_0 q_0$ a term in which essential elements are constant "weights" in the sense of *quantities*. It is only a by product of this approach that it can also be seen as chaining of links where weights in the sense of *expenditure shares* are only partially (with respect to prices) up–dated. If the rationale of P^L_{0t} were based on chaining it would be difficult to justify why updating is done only imperfectly.
On the other hand: if we compared $\bar{P}^{LC}_{0t}$ and P^L_{0t} on the basis of the explicit (or implicit) "ratio–of–expenditure" expression, it would be the chain index $\bar{P}^{LC}_{0t}$, and not P^L_{0t} which would appear poorly reasoned because its numerator turns out to be a highly artificial construct (see eq. 3.1.1).

Transitivity
Many writers conclude from a chain index, being defined as a product, that the chain index is transitive (in the sense of Fisher's circular test), or "chainable" *by definition* since

$$\bar{P}_{0t} = \bar{P}_{0k}\bar{P}_{kt} \tag{1.2.4}$$

holds. According to BANERJEE (1975) for example "the circular test is …satisfied by the chain index" (p. 55).
But the idea of the chain test is that the result ($\bar{P}_{0t}$) should be the same for *any* k, irrespective of how the interval $(0, t)$ is partitioned. But in general this is *not* true for chain indices.[32]

- It can easily be shown, that for example the product $P_{01}P_{12}\dots P_{56}$ and $P_{02}P_{24}P_{46}$ does not yield the same result $\bar{P}_{06}$ (see sec. 4.1).
- Moreover, it can be shown that the idea of chainability as a requirement is necessarily in conflict with other important aspects of index construction, like additivity and adaptability of weights.

[31] In a sense the constant $P_{0t} = 1$ could also be viewed as another special case of chaining in that not only quantities but also prices are held constant.
[32] For more details concerning both points see chapter 4.

To sum up:

> The name "chain index" is misleading: multiplication ("chaining") may give rise to the impression chainability was met. But this is not true. Different types of subdividing an interval and chaining (aggregating) over these subintervals will yield different results and these results will also differ from a direct comparison.

Chaining is a device of *temporal aggregation*, by integrating subintervals.[33] A correct understanding of transitivity (chainability, circularity) requires that *each* chaining over the *same* interval in time should yield the same result, that is indirect comparisons of *all* types should be consistent with direct comparison and also consistent among themselves.

Indirect comparison superior to direct comparison?

According to many authors a chain index is superior to a direct index simply because it takes into account all data of a time series (there is some appeal in this argument by leaving open, what precisely is meant by "taking into account") in contrast to the traditional direct approach, amounting "to establish a trend by a kind of end–point method" (CRAIG (1969), S. 147). But:

> Multiplication is not a unique, defining feature of chain indices. Nor are there any desirable properties of chain indices to be concluded from multiplication alone.

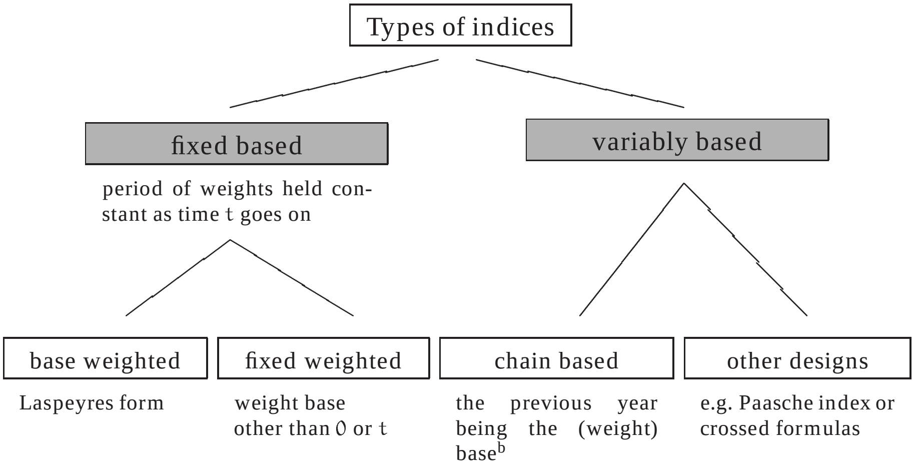
Such erroneous conclusions are often implied in justifications of the chain approach. They are often found in texts of "chainers". Chaining seems to create some additional information (the exact nature of which is not specified, however) or to make better use of data (by letting more data enter the formula). But in general in statistics more phenomena affecting a result will not necessarily give a more "informative" result. In fact more often the very opposite is true. Not infrequently in statistics an effort is made to reduce the effect of irrelevant and incomparable data and to construct a measure in which such influences are eliminated, rather than to make calculations sensitive to a greater number of influences. Also most of the assessment of index formulas is done by examining rather *simple* stylised processes as for example in the case of "axioms". To study figures resulting from a complicated mix of influences is, however, in general of rather limited value.

[33] In sec 4.1 it will also be shown that multiplication of links is justified by (implicitly) assuming that prices will change in proportion when change is measured using *different* weights, such that the weight structure will not matter. On the other hand the change of weights on an ongoing basis (most often viewed as a major advantage of chain indices) is justified on just the opposite grounds: using different (more recent) weights will make a difference, and will result in a more "relevant" measurement of the change in prices.

1.3 Weights in the chain approach

By many authors a distinction is made between fixed based indices and variably based indices respectively (see fig. 1.3.1). This may be useful for some purposes and may have some merits but it does not make clear how to classify the chain approach with respect to its type of weights. Note that "chain based" as defined in fig. 1.3.1 only applies to the link, not to the chain. Moreover $\bar{P}^{LC}_{0t}$, or $\bar{P}^{FC}_{0t}$ and other types of chain indices don't have one but many weighting structures.[34]

Figure 1.3.1: Traditional classification of weighting schemes [a]



[a] Other possibilities exist. The weight base can be periodic other than 0 or t in the interval $(0, t)$ or it can be periodic (e.g. each starting point of a business cycle).
[b] Note that this is true only for the **links**.

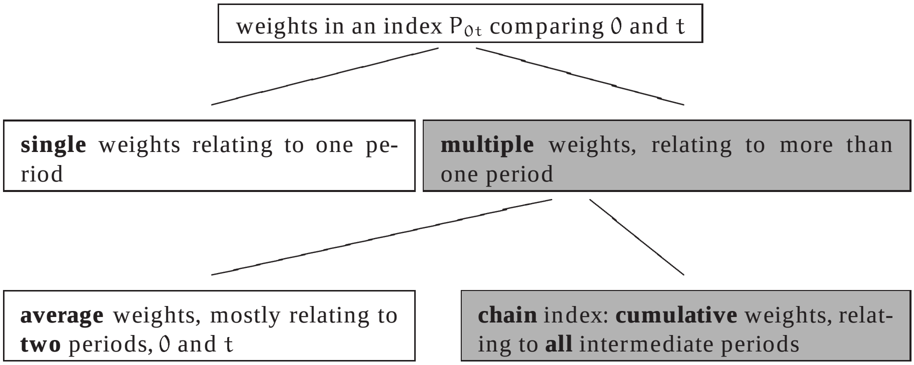
The typology of fig. 1.3.1 is also implicitly using the term "base" with two different meanings. It is useful (as done already in sec. 0.2.a) to distinguish between two types of "base", the *reference base (RB)*, that is the period to which the comparison is related (the period set 100), and the *weight base (WB)*, that is the period to which the weights belong, respectively.

It is often maintained that to change the RB is a purely "formal" maneuver where no problem whatsoever arises. But this is only true for (simple) price relatives, where – interestingly – there is no problem to choose a WB and to aggregate over commodities. But it is common practice to give the following interpretation to say "1990 = 100": the results obtained for 1991, 1992, ... are expressed "in terms of" the level of 1990. In

[34] The chain index appears to be a kind of variably based index only if we look at the individual links instead of the chain. The link, not the chain has a single weighting scheme only. As to the chain, $\bar{P}_{0t}$ will be better characterized by saying that it has multiple or more precisely *cumulative* weights (see fig. 1.3.2).

view of this interpretation there should be some relation between RB and WB and to change the RB should no longer be viewed as a purely “formal” maneuver.

Figure 1.3.2: Alternative principles of defining weights *



* This scheme compares **direct** indices with **chains**, not direct indices with **links** (as fig. 1.3.1 does). A direct index will in general be single–weighted, like P^L or P^P, or average weighted, like P^{ME}, P^T or P^W, whereas a chain index *always* has cumulative weights.

Note that each index has one, and only one RB, but may have several WBs, as for example weights being an average of several weights.[35] It should also be remembered

- that weights related to the base period 0, and weights related to t are *not* logically equivalent, in the sense that for example P^L_{0t} (WB: 0) and P^P_{0t} (WB: t) are equally “logical”, since 0 is kept constant for at least a certain time, thus 0 represents *one* single period, whilst t is variable ($t = 1, 2, \ldots$), denoting *several* periods, and
- that there is a strange inclination to prefer weights related “symmetrically” to both periods 0 and t, as if this “impartiality” were an advantage as such.

A closer look at the nature of 0 and t shows that the idea of a logical “equivalence” of the Laspeyres and Paasche formula[36], and of some alleged advantage of a “symmetric” construction of weights is not tenable.

[35] Examples for this quoted in fig. 1.3.2 are the Edgeworth–Marshall (ME), the Törnquist (T) or the Walsh (W) index.

[36] One of the messages MUDGETT (1951) repeatedly set out, is that the Laspeyres and Paasche formula are equally good or bad. They are logically on “the same footing”, “one rests on just as solid a logical foundation as the other” (p.21), “neither can be considered preferable to the other for logical reasons” (p. 28). Mudgett vigorously advocated chain indices, predominantly because he rejected fixed weights: “It is my view, however, that fixed weights are the greatest weakness of modern indexes, and one purpose of the following pages is to show that a basic analysis of the problem of index number measurements leads infallably to this conclusion” (p. vi). Mudgett certainly has done a lot to support the obsession with most recent weights and the idea of logical superiority of the chain approach to the direct approach.

Both periods, RB and WB coincide in the case of the direct Laspeyres index. Equality of RB and WB is just the main characteristic of the Laspeyres approach: P^L_{0t} (WB = RB) is said to be a *base weighted* index. By contrast a *fixed weighted* index – to be denoted by $P_{0t(s)}$ – also has a constant WB (which is 0), but a WB being neither 0 (the RB) nor t but s.
Take for example a direct Laspeyres index P^L_{0t}, rebased to the reference base s ($\neq 0, t$) by the operation

$$P^L_{0t}/P^L_{0s} = \frac{\sum p_t q_0}{\sum p_s q_0} = P^L_{st(0)} = P_{st(0)}. \tag{1.3.1}$$

The last symbol also allows a more general notation for all single weighted index formulas. Thus we have $P^L_{0t} = P_{0t(0)}$ and $P^P_{0t} = P_{0t(t)}$. With this notation we also get

$$P^L_{03} = P_{01(0)}P_{12(0)}P_{23(0)}, \tag{1.3.1a}$$

with factors having a moving RB but constant WB (in brackets) in contrast to

$$\bar{P}^{LC}_{03} = P_{01(0)}P_{12(1)}P_{23(2)}, \tag{1.3.2}$$

where both, RB and WB are moving.[37] In the same manner, in $\bar{P}^{PC}_{03} = P_{01(1)}P_{12(2)}P_{23(3)}$ links are moving with respect to both, RB and WB.
Equation 1.3.2 also makes clear that $\bar{P}_{0t}$ has three weighting schemes, not only one. We often find comparisons, that might be called "unfair". An example is to give the false impression as if $\bar{P}_{0t}$ had only one single WB like P^C_t, or P_{0t}.[38]

[37] Eqs. 1.3.1a and 1.3.2 are identical with eqs. 1.2.1 and 1.2.2.
[38] There is no single WB in $\bar{P}_{0t}$ it is also far from clear why $\bar{P}_{0t}$ should have an economic–index–theory interpretation (or "true cost of living", COLI interpretation). There is no economic theory index referring to a *sequence* of price and quantity vectors.

2 Elements of (direct) index theory

This chapter deals with some fundamental issues of index theory, starting with the two perhaps most frequently used index formulas, the (direct) index formula of Laspeyres and of Paasche respectively, and also discussed are some recent controversies about the Laspeyres "fixed basket" approach. A succinct introduction into the axiomatic index theory will also be given, and the final part of this chapter will be devoted to aggregation and deflation problems.

2.1 Index formulas of Laspeyres and Paasche

a) Dual interpretation: mean of relatives and ratio of expenditures

The early theory of index numbers suggests the following conclusion: In order to satisfy the commensurability axiom[39] a price index should be a mean of ratios (relatives), instead of a ratio of mean prices, like $\bar{p}_t/\bar{p}_0$ (Dutot's index) or $\tilde{p}_t/\tilde{p}_0$ (a formula of Drobisch, using unit values as defined in eq. 0.3.7). The reason is that price relatives are invariant under a change of physical units (as well as the currency unit)[40] to which price quotations refer. Moreover a mean of relatives necessarily meets the *mean value condition*, that is its value cannot be greater (lower) than the greatest (lowest) individual price relative.

The formula of Laspeyres is a weighted arithmetic mean[41] of price relatives

$$P_{0t}^{L} = \sum_i \frac{p_{it}}{p_{i0}} g_i, \quad \text{where weights } g_i = \frac{p_{i0}q_{i0}}{\sum p_{i0}q_{i0}} \tag{2.1.1}$$

are base period expenditure shares. Analogously the Paasche price index is a weighted harmonic mean of price relatives, weights being expenditure shares in the current period t (eq. (2.1.2) in tab. 2.1.1).

Both index formulas, Laspeyres and Paasche have an intuitive and straightforward dual interpretation, not only as

- *weighted means of price relatives,* but also as
- a *ratio of expenditures* (for example in case of consumer prices).

[39] When commensurability is not satisfied a change in the unit of measurement of *some* commodities (e.g. prices referring to gallons [4 quarts] instead of quarts) does not leave the index unchanged.

[40] This is called "price dimensionality". Note that a change of the currency unit only affects prices, while a change of the unit of "quantity" to which price quotations refer will affect both, prices and quantities. This does not apply, however, to the concept of price level in international (by contrast to intertemporal price comparison: the function P, the "parity" has a dimension which is the relation of two currency units, say hfl / $.

[41] Irving Fisher systematically explored the effects of various choices concerning the type of mean (arithmetic, geometric, harmonic) and of weighting schemes.

According to the “changing cost of a basket” – interpretation of a price index should compare expenditures incurred in order to buy certain well defined goods and services. Thus it should show how changing (in practice mostly rising) *prices* affect the cost of buying a *comparable* “basket”. The Laspeyres approach consists in keeping the “basket” (or “budget”) constant. See tab. 2.1.1 for the relevant formulas.

Table 2.1.1: Dual interpretation of price indices

index	mean of price relatives	ratio of expenditures
Laspeyres	(2.1.1) $P_{0t}^{L} = \sum \frac{p_t}{p_0} \left(\frac{p_0 q_0}{\sum p_0 q_0} \right)$	(2.1.3) $P_{0t}^{L} = \frac{\sum p_t q_0}{\sum p_0 q_0}$
Paasche	(2.1.2) $\left(P_{0t}^{P}\right)^{-1} = \sum \frac{p_0}{p_t} \frac{p_t q_t}{\sum p_t q_t}$	(2.1.4) $P_{0t}^{P} = \frac{\sum p_t q_t}{\sum p_0 q_t}$

Note, a changing cost interpretation requires the same budget (which is the vector of quantities $\mathbf{q}$) in both situations 0 and t respectively

$$P_{0t} = \frac{\sum p_{it} q_i}{\sum p_{i0} q_i} = \frac{\sum p_t q}{\sum p_0 q} = \frac{\mathbf{p}_t' \mathbf{q}}{\mathbf{p}_0' \mathbf{q}}, \tag{2.1.5}$$

or in terms of a mean of price relatives:

$$P_{0t} = \sum \frac{p_t}{p_0} \cdot \left(\frac{p_0 q}{\sum p_0 q} \right) = \sum a_{0t}^{i} w_i, \tag{2.1.5a}$$

where $\sum w_i = 1$.
Hence to gain a “rising costs of a *comparative* budget” interpretation, quantities consumed of each commodity should be *equal* in *both*[42] periods; 0 and t.
In contrast to a price index, designed to measure the *price* component of the cost of living, the value index (or value relative, –ratio)

$$V_{0t} = \frac{\sum p_{it} q_{it}}{\sum p_{i0} q_{i0}} = \frac{\sum p_t q_t}{\sum p_0 q_0} = \frac{\mathbf{p}_t' \mathbf{q}_t}{\mathbf{p}_0' \mathbf{q}_0} \tag{2.1.6}$$

is affected by *both*, price changes and changes in quantities consumed. Furthermore V_{0t} cannot be regarded as a weighted average of price relatives, since in $V_{0t} = \sum \frac{p_{it}}{p_{i0}} w_i^*$ the “weights” $w_i^* = \sum p_{i0} q_{it} / \sum p_{i0} q_{i0}$ do not add up to 1.

42 This is done in order to eliminate the structural component [change of quantities] in the expenditures compared.

In both formulas, P^L_{0t} and P^P_{0t} the *same* quantities (either base or current quantities) are used in the numerator and in the denominator of the ratio–of–aggregates formula in order to avoid the ambiguity of a cost of living (or more general: value) index. Hence both price indices, P^L_{0t} and P^P_{0t} can be interpreted in terms of "changing costs".
It is worthy to note, however, that

> There is *no* symmetry between the Laspeyres– and Paasche formulas as commonly supposed. Whenever a series of successive periods $t = 0, 1, 2, \dots,$ thus a time series of index numbers is considered there is a remarkable difference between the two formulas, in particular with respect to:
>
> - data requirements,
> - an interpretation in terms of representativity and pure price comparison, and
> - the underlying concept of measuring a price movement.

The idea which lies behind this is, as already mentioned: period 0 denotes a *single, constant* (for the "life time" of an index) period whilst period t is a *variable* period referring to *many* different periods, $t = 1, 2, \dots$. There is only *one* constant base period 0, but a *multitude* of current periods, t as time goes on.
Therefore P^L_{0t} and P^P_{0t} are *not* symmetrically built formulas[43]. This can easily be verified by looking at the numerator, and the denominator of the two price index formulas in their aggregative form in the case of a *sequence* of price indices.
In the sequence

$$P^L_{01},\ P^L_{02},\ P^L_{03}, \cdots : \qquad \frac{\sum p_1 q_0}{\sum p_0 q_0}, \frac{\sum p_2 q_0}{\sum p_0 q_0}, \frac{\sum p_3 q_0}{\sum p_0 q_0}, \dots \tag{2.1.7}$$

indices differ from one another *only* with respect to prices in the numerator, while quantities remain the same throughout. In case of successive Paasche–indices, by contrast, numbers we get

$$P^P_{01},\ P^P_{02},\ P^P_{03}, \cdots : \qquad \frac{\sum p_1 q_1}{\sum p_0 q_1}, \frac{\sum p_2 q_2}{\sum p_0 q_2}, \frac{\sum p_3 q_3}{\sum p_0 q_3}, \dots \tag{2.1.8}$$

a series in which each element does not only differ with respect to prices, but to prices *and* quantities. There are some additional points in which P^L_{0t} and P^P_{0t} are markedly different, following a different rationale. But it is better to focus on some of Irving Fisher's ideas first.

[43] It can often be read in index literature that both formulas are conceptually equivalent to one another and based logically on the same footing. This is one of the points repeatedly emphasized for example in the book of MUDGETT (1951), who was one of the most enthusiastic proponents of chain indices. It should be noted that the "economic theory of index numbers" has also done a lot to create the impression of two equally reasoned formulas. Yet here again here it is common practice to consider 0 and t as two equivalent situations.

b) Price and quantity indices, the ill–conceived reversal idea

The idea of measuring volume changes by quantity indices arose much later than the price level measurement by price indices. In general three ways to receive a quantity index (denoted by Q_{0t}) should be distinguished:

- a direct method by *interchanging prices and quantities* (more exactly: price and quantity vectors) in the function $f(\cdot)$ such that we have the price index $P_{0t} = f(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t)$ and the direct quantity index $Q_{0t} = f(\mathbf{q}_0, \mathbf{p}_0, \mathbf{q}_t, \mathbf{p}_t)$,
- indirectly by dividing the value ratio (index) V_{0t} by the price index in question, such that $Q^*_{0t} = V_{0t}/P_{0t}$ where Q^* is the "co–factor" or "factor antithesis" quantity index,
- a third less common method in which both indices are derived *independently* or from a system of interrelated equations.[44]

The idea of the "*factor reversal test*" is to require identity of the direct and the cofactor quantity index $Q_{0t} = Q^*_{0t}$. This is the strict version of the (weaker) "*product test*" in which the product of P_{0t} and Q_{0t} ought to equal the value ratio (value index). Both indices, P_{0t} and Q_{0t} should satisfy certain axioms, but they do not need to possess the same mathematical form (the formula for Q_{0t} [or P_{0t}] resulting from interchanging vectors in P_{0t} [or Q_{0t} respectively]). Thus we get the formulas expressed in matrix (vector) notation as listed in tab. 2.1.2.
The basis for these "antithetic" relations, as Fisher used to call this is given by

$$V_{0t} = P^L_{0t} Q^P_{0t} = P^P_{0t} Q^L_{0t}, \tag{2.1.9}$$

a well known equation, also the starting point of many considerations related to deflation. Equation 2.1.9 also shows that the pair Laspeyres / Paasche is satisfying the product test, but not the much more demanding factor reversal test, because in general $P^L_{0t} Q^L_{0t} \neq V_{0t}$, as is also true in the case of the Paasche formulas $P^P_{0t} Q^P_{0t} \neq V_{0t}$.
It is justified to call Q^L_{0t} and Q^P_{0t} "quantity" indices[45] because both indices have a weighted average of *quantity* relatives interpretation coherent with the respective price indices defined as a weighted mean of *price* relatives. We get for example Q^L_{0t} as weighted arithmetic average of quantity relatives $\frac{q_{it}}{q_{i0}}$ with weights $w_{i0} = p_{i0} q_{i0} / (\sum p_{i0} q_{i0})$, in the same way as P^L_{0t} is an average of price relatives $\frac{p_{it}}{p_{i0}}$.

---

44 An example for this method is Stuvel's pair of indices.

45 Some authors prefer the term "volume" index instead of "quantity" index. The reason is that a price index should eliminate changes in quality thus a (co–factor) quantity index would not only measure changes in quantity but also in quality. But on the grounds set out above we chose not to follow this practice.

Table 2.1.2: Quantity indices of the Laspeyres and Paasche type

quantity index	Laspeyres index	Paasche index
direct quantity index	$Q^L_{0t} = \frac{\mathbf{q}'_t\mathbf{p}_0}{\mathbf{q}'_0\mathbf{p}_0}$	$Q^P_{0t} = \frac{\mathbf{q}'_t\mathbf{p}_t}{\mathbf{q}'_0\mathbf{p}_t}$
co–factor quantity index*	$Q^P_{0t} = \frac{\mathbf{q}'_t\mathbf{p}_t}{\mathbf{q}'_0\mathbf{p}_t}$	$Q^L_{0t} = \frac{\mathbf{q}'_t\mathbf{p}_0}{\mathbf{q}'_0\mathbf{p}_0}$

* (indirect) quantity index gained as a co–factor (factor antithesis) of price.

It was already known to Irving Fisher that the elements of the pair Laspeyres / Paasche are not only the "factor antithesis" but also "time antithesis" of each other, because

$$P^L_{0t} = \frac{1}{P^P_{0t}} \tag{2.1.10}$$

and

$$P^P_{0t} = \frac{1}{P^L_{0t}} \tag{2.1.10a}$$

holds[46] as does (due to eq. 2.1.9)

$$Q^L_{0t} = \frac{V_{0t}}{P^P_{0t}} \qquad \text{and} \qquad Q^P_{0t} = \frac{V_{0t}}{P^L_{0t}}.$$

It is well known that neither P^L nor P^P (or Q^L and Q^P) satisfy time reversibility and factor reversibility whereas the index $P^F_{0t} = \sqrt{P^L_{0t}P^P_{0t}}$, known as Fisher's "ideal" index is able to pass both tests, for example time reversibility, since

$$P^F_{0t}P^F_{t0} = \sqrt{P^L_{0t}P^P_{0t}}\sqrt{P^L_{t0}P^P_{t0}} = \sqrt{P^L_{0t}P^P_{0t}}\sqrt{\frac{1}{P^P_{0t}}\frac{1}{P^L_{0t}}} = 1.$$

In Fisher's view his "ideal index" was superior to both formulas, the Laspeyres and the Paasche formula as his index meets both tests, the time reversal as well as the factor reversal test.
There is, however, conspicuously not much to be found in index theory literature, to show why satisfying these tests should be an "advantage":

46 It can easily be verified that the same is true for the respective quantity indices.

- Deflation requires V_{0t} to be decomposable into a price and a quantity component, but it is far from clear why the product of P^P_{0t} and Q^L_{0t} (that is passing the *product* test) is not sufficient and passing the (much more restrictive) *factor reversal* test is needed[47].
- It is even more doubtful why time reversibility or the "symmetric treatment" of both periods, 0 and t, in choosing the appropriate weights should have any significance in the context of intertemporal comparisons[48].

Recently DIEWERT (1999) recommended the formula of Walsh P^W_{0t} as the best choice in what he called the "theory of pure price and quantity indexes" ("pricing out a constant 'representative' basket of commodities") because of three criteria

1. Diewert requires time reversibility and
2. the weights should be a *symmetric mean* which would allow for two solutions

$$P^W_{0t} = \frac{\sum p_t\sqrt{q_0 q_t}}{\sum p_0\sqrt{q_0 q_t}} \qquad \text{and} \qquad P^{ME}_{0t} = \frac{\sum p_t\,(q_0\,/2 + q_t\,/2)}{\sum p_0\,(q_0\,/2 + q_t\,/2)},$$

   and finally

3. the function of weights should be homogeneous of degree 0 in the components q_0 and q_t, which is another specification of the idea of symmetry[49], and which rules out the formula P^{ME}_{0t} of Marshall and Edgeworth.

In contrast to this position, it appears to us that the direct Laspeyres index fits much better to the idea of a "pure" price index than Walsh's index and that the three criteria chosen to single out this index are poorly reasoned. The criteria are if need be, justifiable on formal and esthetic grounds, at the expense of allowing for a meaningful interpretation. Emphasis on criteria of this kind is perhaps a legacy of Irving Fisher who contributed many most praiseworthy ideas to index theory, but also a philosophy in which

1. fairness, symmetry and the like
2. taking averages ("crossing") always played an important part.

---

[47] The only reason able to motivate this search for "ideal" indices (satisfying not only the product test but the factor reversibility criterion), is the desire to have only *one* index which can serve both purposes, deflation and price level measurement. But there are a number of reasons why we will need two different indices for the two purposes anyway.

[48] Interestingly criteria like those mentioned above, as well as transitivity gain importance in just that field of index theory to which chain indices do not apply: interspatial comparisons.

[49] It is conspicuous that this is again justified with reference to international comparisons, as Diewert pointed out: "the quantity vector of the large country may totally overwhelm the influence of the quantity vector of the small country", acknowledging, however, that this "is not likely to be a severe problem in the time series context".

Thinking in terms of "fairness" can for example be seen by how Fisher himself introduced what he called the "great reversal tests": "Index numbers to be fair ought to work both ways – both ways as regards... the two times to be compared or as regard the two sets of associated elements for which index numbers may be calculated – that is, prices and quantities" (FISHER (1922), p. 62). Thinking in terms of symmetry and fairness might have an intuitive appeal, but it is not sufficiently thought over, and it can even do some harm when it supersedes other more reasoned criteria. The second point is much more harmful. Firstly (less detrimental), it leads us to purely formally motivated index functions, like for example the formula

$$\frac{\sum p_t q_t}{\sum p_0 q_0}\left[\frac{\sum p_0 q_0 \frac{q_0}{q_t} \sum p_t q_0 \left(\frac{q_0}{q_t}+\frac{p_t}{p_0}\right) \sum p_t q_t \frac{p_t}{p_0} \left(\sum p_0 q_t\right)^2}{\sum p_0 q_0 \sum p_t q_0 \sum p_t q_t \frac{q_t}{q_0} \sum p_0 q_t \left(\frac{q_t}{q_0}+\frac{p_0}{p_t}\right)}\right]^{1/4}, \qquad (2.1.11)$$

which he arrived at after "double" crossing[50]. This formula received a better ranking in Fisher's quality scale than the indices of Laspeyres and Paasche, despite the arcane "message" it conveys. Unduly focusing on formal aspects is in itself not the problem. The second (more detrimental) point in thinking in terms of averages and "compromises" is that it thereby becomes *difficult*, if not impossible *to advocate for an extreme solution*. It is for example well known that under certain conditions the Laspeyres and Paasche index will be the upper and lower bound of an interval in which many reasonable formulas will fall.

Hence we are ready to accept P^L_{0t} as biased upwards (or P^P_{0t} as biased downwards) and to think of an index (for example a chain index) that will in general fall short of P^L_{0t} (or exceed P^P_{0t}) as being superior to P^L_{0t} (or toP^P_{0t} respectively). It is in our view intolerable to make such conclusions simply for that reason. Why should an extreme solution not be the correct one on *conceptual* grounds?

In the following we try to clarify the conceptual basis of P^L_{0t} and P^P_{0t} a bit further, and to present additional arguments supporting our view that, ideas which (Fisher was well convinced of), such as successive crossing would carry us nearer and nearer to the truth, or periods 0 and t should be treated symmetrically[51], are erroneous.

c) Interpretation of the Laspeyres and Paasche formula in terms of "pure" comparison

As to the ease of interpretation it certainly is an advantage that both index formulas, P^L_{0t} and P^P_{0t} offer *two* ways of looking at them[52]:

---

[50] Crossing means taking an average. Fisher also made a distinction between crossing (more general) and rectifying (geometric mean).

[51] This is not only the basis of time reversibility but also of the false idea that both, the Laspeyres (base weight) and Paasche (current weight) approach rest "on just as solid a logical foundation as the other", like MUDGETT (1951), p. 21 said, but so did many other authors.

[52] None of the two abovementioned interpretations apply to certain other formulas, as for example the Fisher index or all sorts of chain indices. Fisher's index is preferred because of passing reversal tests,

- in the *aggregative form* (eq. 2.1.1 and 2.1.1) they can be regarded as ratios of expenditures[53] (actually incurred or imputed), or as a comparison of "standardized" expenditure in the sense of referring to the same quantities as weights in both periods to be compared;

- they can be viewed as an average of price relatives (eq. 2.1.1 and 2.1.1), or in other words as a *typical* (average) price relative as well.

In both cases the "Laspeyres principle" consists of taking *constant* (base period) weights, i.e. quantities to be multiplied by prices, or expenditure shares to be multiplied by price relatives. The "Paasche principle" on the other hand consists of using *variable*, that is *current* period quantities or current period expenditure shares respectively.

Generally speaking weights should be capable of measuring the relative "importance" of items. In the case of relative expenditures (mean of price relatives form), this is performed in a way which appeals strongly to common sense: a commodity i is more important than another commodity j if the amount of money spent by an average consumer in buying i is greater than the amount spent on j. This argument applies to both periods, to the base period 0 and to the current period t as well, or: to the historical budget as well as to the actual (or most recent) budget. The problem is now which basket should be preferred.

For those who follow the chain index idea or some widespread prejudices to support chain indices there is no doubt: the actual budget appears – under *all* circumstances[54] – to be preferable to the historical budget. But there are two reasons to argue in favor of a constant, historical budget (or basket):

- the underlying concept of measuring a price movement or defining "volumes",

- data requirements.

The use of constant weights, as applied in the Laspeyres philosophy is justified on the following grounds.

1. Pure comparison

From year to year variations in the Laspeyres price index (or quantity index) can be attributed exclusively to variations in the prices (or in quantities respectively). A *constantly* weighted mean represents a *pure change* in prices (in P^L_{0t}), or in quantities (in case of Q^L_{0t}) by eliminating structural change (making standardized comparisons). Thus Laspeyres' principle is in line with the principle of "pure" price (quantity) comparison, whereas Paasche indices appear superior concerning the *representativity* criterion. But representativity has a price: successive Paasche indices $P^P_{01}, P^P_{02}, P^P_{03}, \ldots$ as well as $Q^P_{01}, Q^P_{02}, Q^P_{03}, \ldots$ are always affected by both variables, changing prices

which are irrelevant in our view, however.

[53] When the price index formula applies to consumer prices as supposed throughout this chapter.

[54] This is a point to be considered in more detail in sec. 4.4.

and changing quantities. The price component is incompletely isolated as t varies in case of P_{0t}^{P} (and likewise the quantity component in Q_{0t}^{P}). *Both* Paasche indices measure a *mix* of influences. On the other hand as the elements of the sequence $P_{01}^{L}, P_{02}^{L}, P_{03}^{L}, \ldots$ differ only due to different prices (eq. 2.1.7) so do successive volumes $\sum p_0 q_1, \sum p_0 q_2, \sum p_0 q_3, \ldots$, that is the numerators in the series of Laspeyres quantity indices $Q_{01}^{L}, Q_{02}^{L}, Q_{03}^{L}, \ldots$, as a result of deflation *only* differ with respect to *quantities*.

2. Consistency

It is sometimes asserted that there is a kind of inconsistency in using P_{0t}^{L} for price level measurement but P_{0t}^{P} for deflation purposes. The idea of the factor reversal test is precisely that the same type of index should be used for both tasks. But there is no reason to require this when a *pure* comparison is intended. Only on the face of it eq. 2.1.9 is combining two equations, in essence in both situations

$$V_{0t} = P_{0t}^{L} Q_{0t}^{P} \qquad \text{(price level measurement), and} \tag{2.1.9a}$$

$$V_{0t} = P_{0t}^{P} Q_{0t}^{L} \qquad \text{(deflation),} \tag{2.1.9b}$$

the emphasis is put on the Laspeyres term. The Paasche index is only a byproduct or tool, its underlying concept of "price level" or "volume" is an *indirect* concept only.

The Laspeyres–approach is to infer rising prices directly from rising costs of a fixed budget while Paasche's approach to conceptualize price movement is more indirect: prices move by the amount by which actual costs diverge from those costs (expenditures), that were to be paid if prices remained constant (an analogous distinction can be made with respect to how "quantity movement" is measured).

The indirect nature of the Paasche approach becomes apparent when we examine the denominators of the price indices $P_{01}^{P}, P_{02}^{P}, P_{03}^{P}, \ldots$, where $\sum p_0 q_1, \sum p_0 q_2, \sum p_0 q_3, \ldots$ is a sequence tracing a *quantity* movement (by the same token the denominators of the *quantity* indices describe a *price* movement with reference to the constant quantities q_0).

The underlying rationale of the duality Laspeyres / Paasche, therefore is a certain division of labor between a price index, and a quantity index, in which the Laspeyres concept represents a "direct" measurement or "pure" component *isolated* in only the observable value movement $\sum p_1 q_1, \sum p_2 q_2, \sum p_3 q_3, \ldots$.

FORSYTH and FOWLER (1981), p. 234, criticize P_{0t}^{L} as this price index "is unacceptable because it utterly fails to represent the changes taking place over the time span" (by "changes" the authors meant changes in quantity). The question arises: if a price index ought to reflect quantity movement, then what is the task of a quantity index?

It is interesting to note in passing that one of the two authors (both most enthusiastic about chain indices) just quoted also presented a definition of "pure" price comparison, we would decidedly disagree with: "Any price index between time t_0 and t_n will give an estimate of pure price change if the (quantity) weights are the same in both the numerator and the denominator of the index" (FOWLER (1974), p. 81). According to this definition the expression $\bar{P}_{0t}^{LC} = \frac{\sum p_1 q_0 \sum p_2 q_1}{\sum p_0 q_0 \sum p_1 q_1}$ is also "a 'pure' index and not an average value index" (ibid., p. 83)[55] although not only the difference between p_0 and p_2 affects the result but also the difference between q_0 and q_1.[56]

To sum up:

> The Laspeyres formulas, P^L_{0t} in cases of price level measurement and Q^L_{0t} in cases of deflation respectively, perform a *pure* comparison, a result influenced only by the variable in question, that is prices only in the case of P^L_{0t}, and quantities only in the case of Q^L_{0t}. As a price index is the aim in the first case and a quantity index in the second there is no inconsistency and no drawback in violating the factor reversal test either.

The Laspeyres approach (price index as well as quantity index) is in keeping with the principle of pure price (volume) comparison. This does not apply to the direct Paasche formula, nor to the chain indices. The merits of the direct Paasche's index P^P_{0t} lie more in utilizing them for purposes of deflation (on the basis of eq. 2.1.9b), that is to estimate volumes in a consistent manner (see sec. 2.3).

3. Practicalities

Constant percentage weights $(p_{i0}q_{i0} / \sum p_{i0}q_{i0})$ allow to work with the same constant market basket for all successive years in which the same base period is maintained. Thus Laspeyres' formula is less demanding regarding data–collection (price– and quantity–reports) than Paasche's formula. The former can do with a constant weighting scheme for some periods, the latter, however, requires an update of the "market basket" (quantity structure) in every period. According to ASTIN (1999) Eurostat had to choose between P^L_{0t} and P^P_{0t} as the formula to be applied in compiling the European Harmonized Index of Consumer Prices (HICP), though both formulas rest on widely different concepts of (price) inflation.

P^L_{0t} was rejected because of its weights by which "it becomes increasingly unrealistic with the passage of time" (the usual verdict of those who prefer a chain index), and P^P_{0t} because of a "purely practical" reason (updating of weights).

Eurostat then decided to recommend a chain index requiring an update of weights, which turns out to be only a bit less frequently than in the case of P^P_{0t}. This shows that in making a choice between a direct and a chain approach and (given a direct one) between P^L_{0t} and P^P_{0t} *conceptual* considerations possibly were not of a high priority. To consider P^P_{0t} instead of P^L_{0t} indicates that it was perhaps not adequately recognized, that P^P_{0t} is more suitable for the purpose of deflation, while on the other hand P^L_{0t} enjoys the advantage of *pure* price comparison over P^P_{0t}, (let alone its straightforward interpretation) in the case of price level measurement.

[55] The term "average value" here most probably denotes "unit value".

[56] In our view the chain price index does not describe a "pure" price movement (i.e. not disturbed by quantity change).

Table 2.1.3: The interpretation of the Laspeyres and Paasche price index and quantity index formula

	Price–index formula (ratio of aggregates form)	
	Laspeyres	Paasche[a]
numerator	imputed expenditures, i.e. expenditures as they *were*, if quantities were kept constant	*empirical*, i.e. observed actual expenditures referring to actual quantities (*not* to constant quantities)
denominator	*empirical*, i.e. actual base period expenditures (being constant); empirically observed and constant	*imputed* expenditures as they *were*, if prices were constant; measures of volume (as substitute for quantity)
time series	time series interpretation possible because only the numerator varies	both, numerator and denominator vary, index numbers not comparable
price movement	*directly* measured: rising (descending) prices inferred from rising (decreasing) costs of a fixed budget[b]	indirectly[c] measured: rising prices because actual costs are higher than they were when prices remained constant
	Quantity–index formula (ratio of aggregates form)	
concept of quantity movement	*direct*: quantities (volume) increased to the extent to which expenditure valued at constant prices has increased (i.e. rising volume)	indirect[c]: quantities are rising, if a value at current prices rises more than a value at constant prices (for a fixed budget i.e. more than volume)

[a] It is assumed that the Paasche price (quantity) index is used to measure price (quantity) *level* movement, but the real merits of the Paasche formula can be seen in the case of deflation (see sec. 2.3).

[b] a quantitatively fixed budget.

[c] indirect approach means: by comparing value and volume (both referring to actual consumption quantities).

d) Theorem of Ladislaus von Bortkiewicz

A valuable equation in comparing the Laspeyres and the Paasche price index was derived by Ladislaus von Bortkiewicz. The theorem can be generalized for comparing any two linear indices, and a direct index with its corresponding chain index (see sec. 3.4a). Denoting price relatives and quantity relatives by a^i_{0t} and b^i_{0t} respectively we obtain $c^i_{0t} = a^i_{0t} \cdot b^i_{0t}$ for the value ratio (relative) of an individual commodity, and by using base period expenditure shares as weights (Laspeyres weights) it is easy to verify that the means of relatives take the following forms

$$P_{0t}^{L} = \sum_i a_{0t}^{i} w_i \quad \text{, where } w_i = \frac{p_{i0} q_{i0}}{\sum p_{i0} q_{i0}} \tag{2.1.12}$$

$$Q_{0t}^{L} = \sum_i b_{0t}^{i} w_i \quad \text{and} \tag{2.1.13}$$

$$V_{0t} = \sum_i c_{0t}^{i} w_i = \sum_i a_{0t}^{i} b_{0t}^{i} w_i. \tag{2.1.14}$$

The covariance C between (weighted) price and quantity relatives is given by

$$C = \sum_i \left(a_{0t}^{i} - P_{0t}^{L}\right)\left(b_{0t}^{i} - Q_{0t}^{L}\right) w_i$$
$$= V_{0t} - P_{0t}^{L} Q_{0t}^{L} = Q_{0t}^{L}\left(P_{0t}^{P} - P_{0t}^{L}\right) = P_{0t}^{L}\left(Q_{0t}^{P} - Q_{0t}^{L}\right), \tag{2.1.15}$$

or

$$C = V_{0t} - P_{0t}^{L} Q_{0t}^{L} \tag{2.1.15a}$$

which is known as the theorem of Bortkiewicz. Since the covariance C is defined as

$$C = r_{ab} s_a s_b = r_{ab} P^L Q^L V_a V_b,$$

where r is the correlation coefficient and V_a, V_b denote the coefficients of variation. The theorem of Bortkiewicz also reads as follows

$$\frac{V_{0t}}{P_{0t}^{L} Q_{0t}^{L}} = 1 + r_{ab} V_a V_b. \tag{2.1.16}$$

The interpretation usually given to this finding is as follows:

- If price and quantity relatives *correlate negatively* ($C < 0$, prices, and quantities move in an opposite direction), which is supposed to be the normal situation, the difference $P^P - P^L$ is negative, that is $P_{0t}^{L} > P_{0t}^{P}$
- on the other hand we get $P_{0t}^{L} < P_{0t}^{P}$ whenever the covariance C is *positive*, thus the value of the Paasche price (and quantity) index is greater than the Laspeyres price (and quantity) index if prices and quantities tend to move in the same direction between 0 and t.

Hence the *direction* of the divergence between P^L and P^P depends on the covariance, however the *extent* of the divergence also depends on the variances, s_a^2 and s_b^2 of the price relatives a and quantity relatives b.
An interesting implication of the formula of Bortkiewicz is the following statement: that $P_{0t}^{L} = P_{0t}^{P}$ is to be expected if:

1. in case of the proportionality test: *all* prices rise at the *same* rate λ ($p_{it} = \lambda p_{i0}$) or all quantities rise at the same rate λ (or remain constant [$\lambda = 1$]), that means: the variance of price ratios and / or of quantity ratios is zero,

2. when the covariance C between a_{0t} and b_{0t} vanishes.

From eq. 2.1.15 we may also conclude:

$$\frac{P^L_{0t}}{P^P_{0t}} = \frac{Q^L_{0t}}{Q^P_{0t}} = 1 - \frac{C}{V_{0t}} = \frac{P^L_{0t} Q^L_{0t}}{V_{0t}} \tag{2.1.17}$$

and by eq. 2.1.16

$$\frac{P^P_{0t}}{P^L_{0t}} = \frac{Q^P_{0t}}{Q^L_{0t}} = 1 + r_{ab} V_a V_b, \tag{2.1.18}$$

which means that the amount by which the Laspeyres formula (usually) exceeds the Paasche formula is the same in the case of the price index, and the quantity index. The so called “Laspeyres–effect”, that is the situation in which $P^L > P^P$ (and consequently also $Q^L > Q^P$) can easily be made plausible now. It occurs when the price of commodity i rises (that is $a^i_{0t} > 1$) the quantity tends to be reduced ($q_{it} < q_{i0}$ hence $b^i_{0t} < 1$) and vice versa. Since $P^L_{0t} = \sum \frac{p_t}{p_0} \frac{p_0 q_0}{\sum p_0 q_0}$ whilst $P^P_{0t} = \sum \frac{p_t}{p_0} \frac{p_0 q_t}{\sum p_0 q_t}$, a price relative which is greater than unity (that is $a^i_{0t} > 1$) will have a lower weight (since $p_0 q_t < p_0 q_0$) in the Paasche formula compared to the Laspeyres formula. On the other hand, if $a^i_{0t} < 1$ (falling price of commodity i) a higher weight will be assigned to it in P^P compared with P^L.
To sum up, whenever prices and quantities change in opposite directions P^Lgives a higher weight (lower weight) to rising (decreasing) prices than P^P does.

2.2 The axiomatic approach in index theory

In the framework of this approach index functions are explored with reference to a set of functional equations[57] called “axioms”. Axioms[58] are useful criteria to assess index formulas. We addressed some of them already. A situation in which axiomatic considerations are not applicable, which is largely true for chain indices (see sec. 3.2) is clearly unpleasant because not much can be said *in general* about the “behavior” of an index formula, and we are left with numerical examples and simulations.

[57] A functional equation is an equation describing a relation which holds for a function φ, for example $\varphi(y, x) = [\varphi(x, y)]^{-1}$.

[58] Some authors make a distinction between “axioms” and “tests” in order to single out those requirements which are so fundamental that they are deemed necessary to define “index”, as opposed to other functions not worth being called “index”. In other words axioms are those properties of a function which are elements of the definition of an index. In what follows no such distinction will be made and a property of an index may be called “axiom” although it might deserve only the label “test” in view of a certain definition.

a) Usefulness of "axioms" and axiomatic systems

In principle we are at liberty to think of any property of a function to serve as an "axiom", or to invent some "new" postulates, and call them "axioms". But it is of course desirable that,

- some motivation or justification for the axiom proposed can be given, that is it should be demonstrated that an index function (formula), satisfying the axiom is preferable to another which violates it, or in other words, satisfying such a requirement makes an index meaningful whereas violating makes it meaningless;
- an exact mathematical formulation of the idea can be found (to give an example: the property of "understandability" of a formula is no doubt highly desirable but attempts to find an exact mathematical expression of "understandability" were not yet successful).

It would be unsatisfactory to dispose of one axiom only, because the class of functions which would be admissible in the light of this axiom would in general be unduly wide[59]. Once two or more axioms are introduced it is first of all necessary to make sure that no axiom contradicts another axiom (or a group of axioms), such that in principle at least one function exists which is able to fulfill *all* axioms required simultaneously, and secondly that no axiom is simply an implication of other axioms. The first characteristic of an axiomatic system is called "consistency" and the second refers to "independence".
There will always be debate as to which axiom is more important to be satisfied and which less, since desirable properties of a formula are often only ensured at the expense of violating other axioms also deemed desirable. Hence some justification of an axiom should be given[60], and a *system* of axioms which is usually a collection of certain *minimum* standards an index should obey, is most useful.
A distinction is sometimes made between a constructive and deductive use of axioms

- If axioms or systems of axioms are used to evaluate some given (known) index formulas in order to arrive at a better understanding of them, a "constructive" use is made.
- The approach is "deductive" if an attempt is made to compare axiomatic systems, and to "characterize" a formula completely by certain axioms (in other words: to deduce the *unique* formula[61] related to the system of axioms).

[59] It is of course also possible to express axioms that can be fulfilled by only one function. But this is not the usual manner to devise axioms. An example would be to require a deflator to satisfy structural consistency which reduces the set of deflators to one element only, the direct Paasche price index (see sec. 2.3d).

[60] We may for example argue for a formula satisfying the factor reversal test. On the other hand this is very restrictive. Therefore doubt may arise in what justifies such a *restrictive* (and in our view unnecessary) criterion, and why the weak version of the test (that is the product test) should not suffice.

[61] In this case the particular index function in question is said to be "characterized" by the axioms in question.

Consistency of a system of k axioms is proved by showing that there exists at least one formula which satisfies all k axioms. The system of "tests" devised by Irving Fisher, however was inconsistent.[62] To prove independence it has to be demonstrated that any $k-1$ of the k axioms can be satisfied by an index which is unable however to fulfill the remaining (k^{th}) axiom.[63]
Axioms are powerful tools to describe the properties of index formulas, and to understand better the "behavior" of a formula by referring to general a priori properties found reasonable or indispensable in view of what an index should measure. A system of axioms, like the system of Eichhorn and Voeller (see fig. 2.2.1)[64] helps to decide which functions are "reasonable", and therefore to be admitted (allowed) and which are "meaningless" and to be rejected.

Figure 2.2.1: Systems of axioms by Eichhorn & Voeller

five axioms system (EV–5)	four axioms system (EV–4)
1: strict monotonicity	1. strict monotonicity
2: (price) dimensionality	2: (price) dimensionality
3: commensurability	3: commensurability
4. identity	4. strict proportionality
5. linear homogeneity	
by implication	
4 + 5 → proportionality*	4 → identity
2 + 5 → homogeneity of degree -1	
2 + 3 → quantity dimensionality	
1 + proportionality → (strict) mean value property	

* Therefore EV–4 is included in EV–5 but the converse is not true.

It may be questioned whether linear homogeneity should be an element of a system of axioms or should the less demanding property of proportionality be regarded as sufficient, because if linear homogeneity is treated as indispensable, some well known

62 In contrast to consistency, inconsistency is difficult to prove. The fact that no formula can be found satisfying all k axioms of a system can only be a hint that inconsistency might exist. It is of course *not* a proof.
63 Thus in a three axioms system at least one example of an index function has to be found that violates axiom 1, but satisfies the remaining axioms, 2 and 3, or that violates axiom 2, but satisfies axiom 1 and 3 and so on.
64 Not all axioms mentioned here are already explained, but comments on their meaning will be given later.

index formulas would be ruled out[65]. Therefore the less ambitious system of four axioms was conceived. OLT (1996) made an attempt to relax the requirements of an axiomatic system even further.

In contrast to an *axiomatic theory*, by which a formula is evaluated without using data, and in which only two categories (to pass or to fail a test) are distinguished a so called "*quantum theory*" deals with the amount of deviation (or "bias") from fulfilling an axiom. The aim is to give an estimate for (or to determine upper and lower bounds for) the result an index function will yield when applied to data. The problem of such reasoning is that hypothetic "data" are perhaps erroneously assumed to be "representative" or "typical". Hence conclusions drawn from calculations with such "data" might happen to not be tenable under more general conditions concerning the data. In the absence of axiomatic reasoning it is difficult to establish *generally* valid judgements of index formulas.

b) A tentative hierarchy of axioms

Once a system of axioms has been established, each axiom is regarded as equally justified and necessary as every other. Thus there is no hierarchy among axioms. But to decide on which axiomatic system or which formula should be preferred to another, it is in general desirable to classify axioms according to their "importance" for the specific purpose the index function should serve, or to gain at least some idea on more or less "importance" of certain properties.

In what follows a tentative classification of axioms is presented (fig. 2.2.2), which is of course a personal view of the author, and not an issue to be settled by mathematical proof.[66] We distinguish between three types of axioms, that are

a) axioms expressing fundamental prerequisites of *measurement* in general, paramount to aspects important for specific applications (some properties any reliable measuring rod must have); these are essentially *invariance* axioms

   1. **commensurability** and
   2. (price) **dimensionality** to ensure independence of units of measurement (or "invariance" under a change in these units), and
   3. **identity** in order to avoid ambiguity of comparison, that is to make $P = 1$ a situation (reference point) from which a deviation is well defined like $P > 1$ and $P < 1$ reflecting a rise or decline respectively;

[65] To be more distinct, the functions of Stuvel, and the so called log–change index numbers.

[66] In principle many other classifications of axioms might be conceived and different views on the relevance of certain axioms may be adopted. Any attempt to classify axioms is inevitably affected by opinions on what makes axioms more or less justified, more or less dispensable and the like. Widely divergent views on this issue are possible without calling in question the general usefulness of the axiomatic approach.

b) axioms necessary (or at least desirable) with respect to the *purposes* for which an index is compiled, that is axioms inspired by the *use* of index numbers in *economic analysis*: (or interpretations intended): the axioms of this group refer to

- a correct reflection of *movement* from the unambiguously defined situation of unity (or 100%), a group of axioms to which belong: **monotonicity, linear homogeneity** and (by implication) **mean value property** (it will be shown that these criteria strongly appeal to an intuitive understanding of what an index intends to measure);
- *aggregative properties*: being a measure related to aggregates, an index should be consistently decomposable (broken down by commodities [sub–indices], or by a kind of variable [price vs. quantity index]) aggregative consistency, and volumes resulting from deflation should be capable of meaningful summation structural consistency;

c) a group of axioms introduced in *analogy* to relatives, based on the belief that index numbers should behave as if they were individual relatives, or axioms motivated by some *symmetry* considerations; this is the least reasoned group of axioms.

Note that the axioms 1 through 5 listed in fig. 2.2.2 are just the axioms of the system EV–5. Since some of the axioms have already been discussed, and will be discussed in more detail in what follows, some brief comments on the three groups of axioms will be sufficient here.

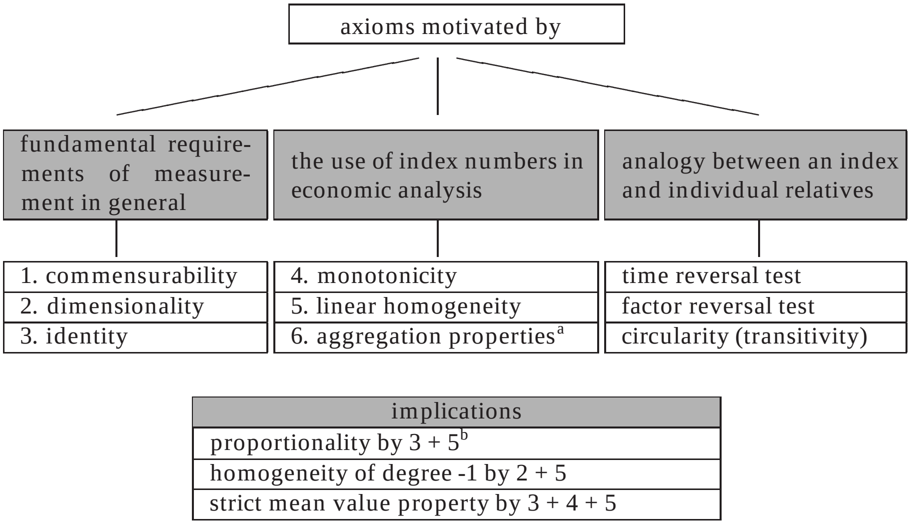
Ad a) and b): Prerequisites of measurement and correct reflection of movement

Obviously it is highly desirable that a price index P correctly indicates *direction* and *amount* of change at least in the cases in which an extremely *simple* kind of change is assumed (see fig. 2.2.3) as for example

- no change takes place,
- an isolated change, where all other prices remain constant, or
- prices change uniformly (at the same rate).

More difficult is a non–uniform change, but in this case we at least (or only) can reasonably postulate the mean value property. Another simple situation is a rise (or decline) which is consistently λ–fold higher as compared to another change (linear homogeneity is the axiom which addresses this situation).

Figure 2.2.2: Tentative classification of axioms and their uses



[a] various consistency requirements (see sec. 2.3) and also consistent decomposition of the value ratio into a price index and a quantity index (product test = weak version of factor reversibility)
[b] furthermore: when proportionality then also identity (the converse is not true)

Hence it appears to be a *minimum* standard only (but therefore precisely an indispensable standard), for an index to be used in analysing the movement of a price level to have

- a clear idea of a reference point ($P = 100\%$) and the amount of change to be reflected by an index under most simple hypothetical conditions, and
- the movement we want to measure separated from other (irrelevant) kinds of changes, like changes in units of measurement for example (note that this is an idea fundamental also to the notation of “pure price comparison”).

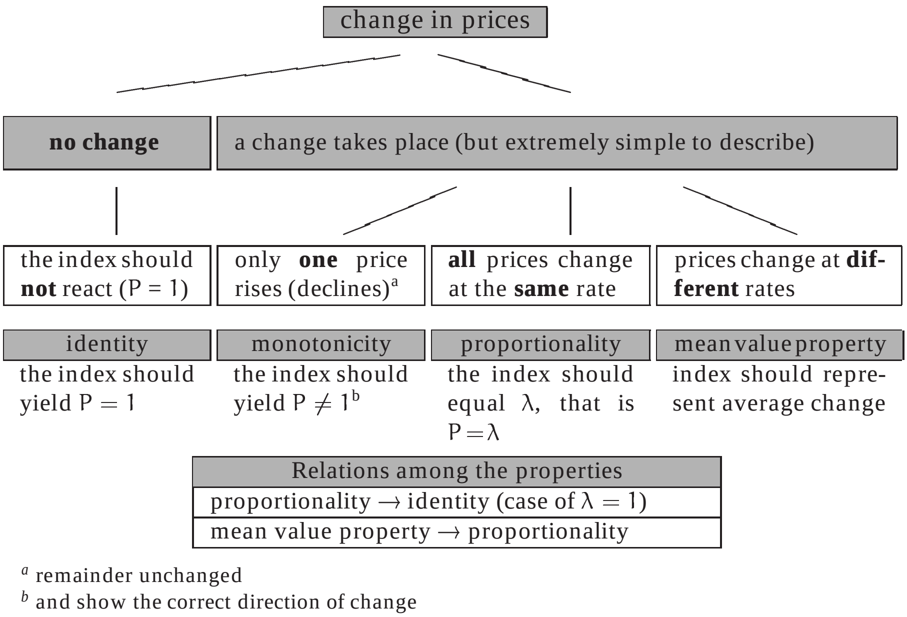
The idea is, when an index function is unable to differentiate correctly among different “degrees” of change in prices in such “idealized” simple situations, it will all the more lack the ability to measure price movement in more complicated, and more realistic situations.

Ad b): Usefulness in economic analysis
Whenever an “explanation” of a price movement, for example in the case of consumer prices is intended to show to which extent certain prices contributed to the overall price movements certain aggregation properties *of the index function* are highly desirable. This should be kept distinct from aggregation properties in the framework of

deflation. Volumes derived from using a "deflator"–price index should also be capable of meaningful and consistent decomposition or aggregation. Thus such properties, also known as consistency *of volumes* are also important to ensure that values as well as volumes (deflated values) will satisfy the same definitional equations.Note that we distinguish[67] aggregation properties of

- the index function itself, and
- of volumes (resulting from deflation).

Figure 2.2.3: Relations among some axiomatic properties



Ad c): Analogy to relatives

It is not legitimate to postulate properties an index should have simply by analogy to individual price relatives, because indices and relatives differ by a number of aspects (the impact of structural change for example).

Axioms like time reversibility and circularity (transitivity) might also be viewed as aspects of invariance in a certain sense: time reversal refers to irrelevance of the choice

[67] see sec. 2.3 for more detail.

of the base, and circularity (fundamental for chain indices, see also below and sec. 4.1) can be regarded as a type of consistency in aggregation over periods in time in the sense that a "direct" calculation P_{0t} is consistent with any "indirect" calculation referring to the same interval (0 to t), like $P_{0s}P_{st}$. But there are good reasons *not* to insist on such criteria in the *intertemporal* case. When no symmetry exists between 0 and t there is no point in interchanging 0 and t. Nor is there any reason why a relation between P_{0s}, P_{st} and P_{0t} should hold (or hold by definition) when there is no need to compare indices that are directly calculated with those indirectly calculated.
Some axioms need further explanation.

c) Some axioms explained in detail

This section comments on the meaning of some axioms, except for axioms related to aggregation to which sec. 2.3 is devoted. It is difficult to find a structure according to which axioms should be explained, because axioms are interrelated in a complicated manner. In general it is not possible to identify an axiom as more "general" or "restrictive" than another one, unless we explicitly introduce the terms "strict" and "weak".

1. Commensurability and dimensionality
Usually there is no problem with the two properties commensurability and dimensionality. Therefore a short comment will do. Commensurability can be expressed as follows

$$P\left(\mathbf{Lp}_0, \mathbf{L}^{-1}\mathbf{q}_0, \mathbf{Lp}_t, \mathbf{L}^{-1}\mathbf{q}_t\right) = P\left(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t\right), \tag{2.2.1}$$

where $\mathbf{L}$ is a $n \times n$ diagonal matrix with elements $\lambda_1, \ldots, \lambda_n$, such that

$$\mathbf{L} = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix} \quad \text{and} \quad \mathbf{L}^{-1} = \begin{bmatrix} 1/\lambda_1 & 0 & \ldots & 0 \\ 0 & 1/\lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1/\lambda_n \end{bmatrix}.$$

By virtue of commensurability (throughout quite a powerful property), the index function can be expressed in price relatives. Consider a matrix $\mathbf{L}$ as follows

$$\mathbf{L} = \begin{bmatrix} 1/p_{10} & 0 & \ldots & 0 \\ 0 & 1/p_{20} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1/p_{n0} \end{bmatrix}$$

with main diagonal elements $1/p_{i0}$. Then we obtain

$\mathbf{Lp}_0 = \mathbf{1}$, where $\mathbf{1}' = [1\ 1 \ldots 1]$ and

$\mathbf{Lp}_t = \mathbf{a}$, the vector of price relatives $\mathbf{a}' = [p_{1t}/p_{10}\ p_{2t}/p_{20} \ldots p_{nt}/p_{n0}]$.

Furthermore

$\mathbf{L}^{-1}\mathbf{q}_0 = \mathbf{v}_0$, the vector of base *values* $\mathbf{v}_0' = [p_{10}q_{10}\ p_{20}q_{20} \dots p_{n0}q_{n0}]$

and

$\mathbf{L}\mathbf{q}_t = \mathbf{v}_t$, the vector of *volumes* $\mathbf{v}_t' = [p_{10}q_{1t}\ p_{20}q_{2t} \dots p_{n0}q_{nt}]$.

Thus when commensurability holds, the index function can be expressed as follows:

$$P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t) = P(\mathbf{a}, \mathbf{v}_0, \mathbf{v}_t), \quad (2.2.2)$$

that is as a *function* of only $3n$ (instead of $4n$) variables, the *price relatives*, a_i the reference period volumes v_{it} (valued at base period prices) and the base period values v_{i0}.

By (price) **dimensionality** or *homogeneity of degree 0 in prices* the index function is not affected by a change in the currency unit to which prices refer (independence of the currency in which the prices are expressed):

$$P(\lambda\mathbf{p}_0, \mathbf{q}_0, \lambda\mathbf{p}_t, \mathbf{q}_t) = P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t). \quad (2.2.3)$$

Price dimensionality is a trivial property of (single commodity) price relatives as pointed out in sec. 0.3. In combination with commensurability quantity dimensionality (also called “weak commensurability”)

$$P\left(\lambda\mathbf{p}_0, \lambda^{-1}\mathbf{q}_0, \lambda\mathbf{p}_t, \lambda^{-1}\mathbf{q}_t\right) = P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t), \quad (2.2.4)$$

or more commonly expressed as follows

$$P(\mathbf{p}_0, \lambda\mathbf{q}_0, \mathbf{p}_t, \lambda\mathbf{q}_t) = P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t), \quad (2.2.4a)$$

is satisfied. If price dimensionality and commensurability is met then quantity dimensionality is also, but the converse is not true[68]. Quantity dimensionality is the special case of all $\lambda_1 \dots \lambda_n$ being equal ($\lambda_i = \lambda\ \forall\ i = 1, 2, \dots, n$), such that each commodity is subject to the *same* change in the unit of measurement and $\mathbf{L}$ reduces to the scalar matrix $\mathbf{L} = \lambda\mathbf{I}$.

2. Correct reflection of direction and amount of change

Axioms like identity, proportionality, monotonicity, additivity and mean value property are all closely related fundamental properties of index functions, ensuring that this function will behave like an average. The idea of the “*identity* test” was introduced by Laspeyres and explains that: if *no* price changes the price index function should be 1 (unity). The example which is quoted most often of an index violating

[68] Dutot’s index $\sum p_t / \sum p_0$ as an example satisfies quantity dimensionality and price dimensionality but not commensurability.

(strict) identity, is the value ratio (index) V_{0t}, due to different quantities in the numerator and the denominator (in fact they are likely to differ since we may assume in most practical applications that different prices in 0 and t will also lead to different quantities $q_{it} \neq q_{i0}$). Often a distinction is made between strict and weak identity in the following manner:

$$P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t) = 1 \qquad \text{strict identity,} \tag{2.2.5}$$

when $\mathbf{p}_t = \mathbf{p}_0$ holds for the price vector

$$P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t) = 1 \qquad \text{weak identity,} \tag{2.2.5a}$$

where $\mathbf{p}_t = \mathbf{p}_0$ *and* $\mathbf{q}_t = \mathbf{q}_0$.
As usual the weak version is gained from the strict one by adding further assumptions (like $\mathbf{q}_t = \mathbf{q}_0$) assuming that not only prices remain unchanged but also quantities. Thus obviously strict identity implies weak identity, but the converse is not true. The value index V_{0t} satisfies weak identity, but not strict identity.
A statement to be regarded in a certain sense as the "opposite" of identity, is: if *one* price, taken in isolation (and thus also: "some", or "all" prices) rises (or declines), the index should *not* be 1. This is guaranteed by (weak) *monotonicity* (see below).
Identity can also be related to *time reversibility*: if prices rise in the interval from 0 to 1 such that $\mathbf{p}_1 > \mathbf{p}_0$, but return (decline) thereafter to their initial level, such that $\mathbf{p}_2 = \mathbf{p}_0$, then the price index should be unity ($P_{02} = 1$). It is, as an example, a serious defect of *chain indices* which is often pointed out, that such indices need not result in unity under such conditions.
According to **proportionality** an index should be λ if all prices change at the same rate λ such that $p_{it}/p_{i0} = \lambda$ for all $i = 1, 2, \ldots, n$. This requirement is intuitively appealing since it refers to an extremely simple situation. Its rationale is that an index $P(\cdot)$ should at *least* reflect a *uniform* change correctly: if $P(\cdot)$ fails in doing so, it is more likely that $P(\cdot)$ will also reflect a *mixed* change incorrectly, as it occurs in practice in general. The requirement

$$P(\mathbf{p}_0, \mathbf{q}_0, \lambda\mathbf{p}_0, \mathbf{q}_t) = \lambda \tag{2.2.6}$$

is called strict proportionality, where $\lambda \in \mathbb{R}$, and $\mathbf{p}_t = \lambda\mathbf{p}_0$, and

$$P(\mathbf{p}_0, \mathbf{q}_0, \lambda\mathbf{p}_0, \mathbf{q}_0) = \lambda \tag{2.2.6a}$$

is called weak proportionality, where $\lambda \in \mathbb{R}$, $\mathbf{p}_t = \lambda\mathbf{p}_0$, $\mathbf{q}_t = \mathbf{q}_0$.
Even indices badly biased are able to meet weak proportionality but not strong proportionality. This can easily be verified: proportionality implies identity (the case $\lambda = 1$), but not conversely. The first part of the assertion is easy to verify. To prove the second part an example is sufficient. The index function $\sum \left(\frac{p_{it}}{p_{i0}}\right)^2 \frac{p_{i0}q_{i0}}{\sum p_{i0}q_{i0}}$ fulfills identity but not proportionality.

Weak monotonicity is defined by

$$P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t) > P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_0, \mathbf{q}_t) \qquad \text{if } \mathbf{p}_t \geq \mathbf{p}_0 \tag{2.2.7}$$

and

$$P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t) < P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_0, \mathbf{q}_t) \qquad \text{if } \mathbf{p}_t \leq \mathbf{p}_0 \tag{2.2.8}$$

and implied by strict monotonicity (see below). It should be noted that the two inequalities, eq. 2.2.7 and 2.2.8, defining weak monotonicity are independent[69] as functions exist (however somewhat far–fetched with respect to the formula as well as with numerical examples), that are able to satisfy the first part (eq. 2.2.7), but not necessarily the second part (eq. 2.2.8), or the second part but not necessarily the first part of the definition respectively. But in all more or less "normal" situations, indices will satisfy both conditions of weak monotonicity[70].
The *implications of monotonicity* will be shown in more detail when special cases of monotonicity, such as additivity or multiplicativity will be discussed (see part d of this section). An important relationship should now be mentioned:

> Index functions that can be conceived as means of price relatives are always monotonically increasing (decreasing) when the price relatives rise (decrease), that is they are by implication monotonous in the *weak* sense.

Monotonicity is a most fundamental characteristic of all reasonable means. If a single price rises (declines), everything else being constant, there should be a reaction of the index formula, also showing a change in the proper direction, that is an increase or a decrease in the price level. If there were no such reaction, the function would not be reasonable. The median of price relatives is an example of an index that might violate this condition.

3. Correct indication of difference in amounts of change
Two axioms, strict monotonicity and linear homogeneity refer explicitly to a situation where the same index function is applied to two different sets of data, one expressing a higher (or lower) change in prices than the other.
Strict monotonicity applied to an index function $P(\cdot)$ is defined as follows:

$$P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t^*, \mathbf{q}_t) > P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t) \qquad \text{if } \mathbf{p}_t^* \geq \mathbf{p}_t \tag{2.2.9}$$

and

$$P(\mathbf{p}_0^*, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t) < P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t) \qquad \text{if } \mathbf{p}_0^* \leq \mathbf{p}_0. \tag{2.2.10}$$

[69] In the same manner the two parts of the definition of *strict* monotonicity, that is eq. 2.2.9 and 2.2.10 are independent.
[70] The same is true for *strict* monotonicity.

By $\mathbf{p}_t^* \geq \mathbf{p}_t$ is meant that at least one element of the non–negative vector $\mathbf{p}_t^*$ say p_{it}^* is greater than the corresponding element of $\mathbf{p}_t$ (and $\mathbf{p}_0^* \leq \mathbf{p}_0$ is defined correspondingly).

An index for which strict monotonicity holds will satisfy the following *two* conditions[71]:

1. by eq. 2.2.9: when the prices $\mathbf{p}^*$ (at least one of them) in the current period, t are (is) raised [or lowered] (compared with the prices $\mathbf{p}$) the index has to reflect this rise [or decline], and correspondingly

2. by eq. 2.2.10: when base period (0) prices are assumed to be higher [lower] in case of $\mathbf{p}^*$ than in case of prices $\mathbf{p}$ the index should be lower [higher].

Thus a price index has to indicate an increase ($P > 1$), whenever any of the prices (at least one of them, say the i–th), in the *current* period p_{it}^* is greater than p_{it}, and if at least one of the *base* period prices is lower than p_{i0}, that is if $p_{i0}^* < p_{i0}$, everything else being constant.

Note that the two inequalities (eq. 2.2.7 and 2.2.8), in the case of weak monotonicity correspond to one only (eq. 2.2.9) in the case of strict monotonicity, and that the two inequalities defining strict monotonicity (eq. 2.2.9 and 2.2.10) are independent. Furthermore: if strict monotonicity is met then also weak monotonicity, but the converse is not true. Inequality 2.2.7 follows from 2.2.9 by substituting $\mathbf{p}_t = \mathbf{p}_0$ on the right-hand side and $\mathbf{p}_t^* = \mathbf{p}_t$ on the lefthand side of eq. 2.2.9 (in this order). In the same manner we get eq. 2.2.8 by substituting in eq. 2.2.10 as follows, first $\mathbf{p}_t = \mathbf{p}_0$ (right) and then $\mathbf{p}_0 = \mathbf{p}_0^*$ (left). Hence strict monotonicity implies weak monotonicity.

In other words: weak monotonicity only requires a price index showing an increase whenever any of the prices in the current period are raised, or any of the prices in the base period are lowered. Strict monotonicity on the other hand requires the definition of a price index P_{0t}^*, showing a greater increase than another price index P_{0t}, whenever any of the price relatives in P_{0t}^* show a greater increase than the price relatives in P_{0t} due to

- higher current year prices (in t), or
- lower base period prices (in 0).

Note that strict monotonicity only requires an index P^* to display a higher (lower) rise or decline in prices as compared with another index P, nothing is said about the difference $P^* - P$.

A statement referring to the *amount* by which P_{0t}^* should differ from P_{0t}, is given by linear homogeneity and by additivity or multiplicativity (for the last two terms see sec. 2.2.d).

An index function $P(\cdot)$ satisfying

$$P(\mathbf{p}_0, \mathbf{q}_0, \lambda\mathbf{p}_t, \mathbf{q}_t) = \lambda P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t) \qquad \lambda \in \mathbb{R} \tag{2.2.11}$$

[71] In the next part of this section it will be shown that the two conditions are independent.

is said to satisfy **linear homogeneity** (or homogeneity of degree +1) in the prices of period t. Note that eq. 2.2.11 corresponds to eq. 2.2.9. The axiom corresponding to eq. 2.2.10 is

$$P(\lambda \mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t) = \frac{1}{\lambda} P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t), \tag{2.2.12}$$

which is called *homogeneity of degree -1* in base period prices.

Figure 2.2.4: Strict and weak monotonicity

Monotonicity is a statement concerning

one index only, such that $P_{0t} > 1$ or $P_{0t} < 1$ is required	the comparison of two indices, P_{0t} and P^*_{0t} such that $P_{0t} > P^*_{0t}$ or $P_{0t} < P^*_{0t}$ is required
weak monotonicity	strict monotonicity
the index function is monotonically increasing (decreasing) in the price relatives*	P_{0t} and P^*_{0t} refer to situations which are partly the same and partly different

equation (inequality)	P_{0t} and P^*_{0t}: the same is	P_{0t} and P^*_{0t}: different is
eq. 2.2.9	$\mathbf{p}_0, \mathbf{q}_0, \mathbf{q}_t$	$\mathbf{p}_t$ and $\mathbf{p}^*_t$
eq. 2.2.10	$\mathbf{p}_t, \mathbf{q}_0, \mathbf{q}_t$	$\mathbf{p}_0$ and $\mathbf{p}^*_0$

* This simply means that we should get $P > 1$ when price relatives p_t/p_0 indicate a rise irrespective of whether p_t exceeds p_0 or p_0 falls short of p_t (case of $P < 1$ analogously). For this reason *two* equations (eq. 2.2.7 and 2.2.8) correspond to one only (eq. 2.2.9) in case of strict monotonicity.

In the case of strict monotonicity, both parts of the definition (eqs. 2.2.9 and 2.2.10) have to be postulated separately. But to express the axiom of linear homogeneity it is in general sufficient to focus on eq. 2.2.11. To consider eq. 2.2.12 separately, it is not necessary for the following reason: if linear homogeneity and dimensionality is met then also homogeneity of degree -1. Furthermore:

Linear homogeneity in combination with identity implies (strict) proportionality. Again the converse is not true. The Laspeyres price index conforms with linear homogeneity and also strict proportionality since

$$\frac{\lambda p_{1t} q_{10} + \lambda p_{2t} q_{20} + \cdots + \lambda p_{nt} q_{n0}}{p_{10} q_{10} + p_{20} q_{20} + \cdots + p_{n0} q_{n0}} = \lambda \frac{p_{1t} q_{10} + p_{2t} q_{20} + \cdots + p_{nt} q_{n0}}{p_{10} q_{10} + p_{20} q_{20} + \cdots + p_{n0} q_{n0}} = \lambda P^L_{0t}.$$

But the index $P_{0t}^{Y} = \sqrt{\frac{\sum p_t^2 q_t^2}{\sum p_0^2 q_0^2}}$ proposed by Young is an example of an index satisfying linear homogeneity, but proportionality only in its *weak* form and therefore *not strict* identity, and *not* strict (or even weak) monotonicity either.
On the other hand indices even satisfying strict proportionality, but not linear homogeneity are

- the exponential mean index $P_{0t}^{EX} = \ln\left[\sum \exp\left(\frac{p_{it}}{p_{i0}}\right) \frac{p_{i0}q_{i0}}{\sum p_{i0}q_{i0}}\right]$, and
- the Vartia–I index[72].

Stuvel's pair of indices (price index and quantity index) satisfies *weak* proportionality $P(\mathbf{p}_0, \mathbf{q}_0, \lambda\mathbf{p}_0, \mathbf{q}_0) = \lambda$ and identity, but *not* linear homogeneity.
Another interesting example given by EICHHORN and VOELLER (1983) is the Marshall–Edgeworth factor antithetic price index, that is V_{0t}/Q_{0t}^{ME} which satisfies identity and proportionality[73], but not linear homogeneity.
Remember that there were *two* conditions under which proportionality would follow: linear homogeneity *and* identity. Some examples were given for proportionality being met but not linear homogeneity, but it is also possible to find index functions satisfying linear homogeneity, and yet not being capable of satisfying proportionality. An example for this is the value index V_{0t}:
V_{0t} satisfies linear homogeneity due to $\frac{\sum \lambda p_t q_t}{\sum p_0 q_0} = \lambda \frac{\sum p_t q_t}{\sum p_0 q_0}$, but since $\frac{\sum \lambda p_t q_t}{\sum p_0 q_0} \neq \lambda$ strict proportionality is not met (when $q_{it} \neq q_{i0}$). There is a similar situation in the case of chain indices.
Hence it is *not* correct to say linear homogeneity is more demanding, or "stronger" than proportionality, although there exists a certain resemblance between

- weak and strong monotonicity on the one hand (see fig. 2.2.4), and
- proportionality and linear homogeneity on the other hand (see fig. 2.2.5).

We may safely conclude "if strong monotonicity then weak monotonicity" but not "if linear homogeneity then proportionality" unless the index function also meets identity (a special case of proportionality).
It should be noticed that an index with base 0 aims at making *several* comparisons, not only one, that is a *series* of comparisons $P_{01}, P_{02}, P_{03}, \ldots$ is in general aimed at. Therefore linear homogeneity and strict monotonicity, both statements comparing two indices with one another are to be preferred to proportionality and weak monotonicity, which are requirements referring to one index only.

[72] This holds, in contrast to the Vartia–II index. OLT (1996), p. 86 erroneously states that the Vartia–II index violates linear homogeneity and the Vartia–I index violates strict proportionality.
[73] This index satisfies all four axioms of the system EV–4 but not linear homogeneity of EV–5 (cp. sec. 2.2.a).

Figure 2.2.5: Linear homogeneity, monotonicity and proportionality

A statement concerning

one index only, reflecting correctly	difference of two indices
direction of change: **weak monotonicity**	direction: **strict monotonicity**
P_{0t} is increased ($P_{0t} > 1$) whenever any of the prices in t are raised or any of the prices in 0 are lowered	$P^*_{0t} > P_{0t}$ when prices in t are $\mathbf{p}^*_t > \mathbf{p}_t$ and in 0 are $\mathbf{p}^*_0 = \mathbf{p}_0$
amount of change: **proportionality**	amount: **linear homogeneity**
if all prices rise at the same rate[a] λ, the index P_{0t} should amount to λ ($\mathbf{p}_t = \lambda \mathbf{p}_0$)[b]	$P^*_{0t} = \lambda P_{0t}$ when prices in t are $\mathbf{p}^*_t = \lambda \mathbf{p}_t$ and in 0 are $\mathbf{p}^*_0 = \mathbf{p}_0$[b]

[a] more precisely λ is the growth *factor* of prices.
[b] comparing prices $\mathbf{p}_t$ with $\mathbf{p}_0$ (proportionality) or prices $\mathbf{p}^*_t$ and $\mathbf{p}_t$ (linear homogeneity).

Strict proportionality $P(\mathbf{p}_0, \mathbf{q}_0, \lambda\mathbf{p}_0, \mathbf{q}_t) = \lambda$, implied by linear homogeneity together with identity is a characteristic that greatly appeals to common sense[74]: If all prices rise at the same rate of growth λ, the index should indicate exactly this rate of growth, i.e. it should amount to λ.

4. More details on mean value property and proportionality

The distinction between *strict* mean value property

$$\min\left(a^i_{0t}\right) \leq P_{0t} \leq \max\left(a^i_{0t}\right) \tag{2.2.13}$$

where a^i_{0t} are individual price relatives, and *weak* mean value property not requiring that P_{0t} equals $\min\left(a^i_{0t}\right)$ or $\max\left(a^i_{0t}\right)$ if and only if all price relatives are identical, is somewhat theoretical and sophisticated.

Thus a price price index should take values within the boundaries of $\min\left(a^i_{0t}\right)$, the smallest and $\max\left(a^i_{0t}\right)$, the greatest individual price relative. To add “strict” means that the boundaries themselves should be relevant *only* when $a^i_{0t} = a_{0t} = \min\left(a^i_{0t}\right) = \max\left(a^i_{0t}\right)$ for all $i = 1, \ldots, n$.

According to PFOUTS (1966) the mean value property is one of the most essential properties of an index function. The reason is:

[74] It is widely an expression of common sense, but there are chain–index–formulas that violate this fundamental condition.

The mean value property conforms with our intuitive notion of an index being a measure of a "*representative*" *aggregate* change: the change of an aggregate (as measured by an index) should be related to the change of its components (individual price relatives) such that the index is a representative (typical) relative. It is one of the inconsistencies we observe in advocating for chain indices, that a lot of care is taken for representativity of *weights* assigned to the relatives instead of ensuring representativity of the *result*.

It is again the operation of chaining which is responsible for a situation in which representative weights in each period, and a price index as a representative price relative over an interval in time are not the same.

More rigorously expressed in mathematical terms mean value property requires, a real number λ such that the price index is a convex combination of the price relatives

$$P\left(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t\right) = \lambda \min\left(a^i_{0t}\right) + (1-\lambda) \max\left(a^i_{0t}\right) \tag{2.2.14}$$

where "strict" means $0 < \lambda < 1$ whilst in case of "weak" $0 \leq \lambda \leq 1$ is admitted.

The relations between linear homogeneity, proportionality (and identity as a special case) and monotonicity on the one hand and mean value property on the other are as follows:

If (strict) monotonicity, linear homogeneity and identity are met	$\longrightarrow$	then also (strict) mean value property

Only in combination with strict monotonicity and identity linear homogeneity entails the mean value property. This is due to the fact, that

1. linear homogeneity and identity yield proportionality, and

2. monotonicity and proportionality imply the (strict) mean value property.

Thus we also get

If (strict) monotonicity and proportionality are met	$\longrightarrow$	then also (strict) mean value property

Hence the relation between proportionality, and the mean value test is closer than between linear homogeneity, and the mean value test.

Proportionality is clearly an implication of mean value property

If (strict) mean value property is met	$\longrightarrow$	then also proportionality

but again the converse is not true, as can be seen by the index P_{max} which is defined as the maximum price relative.

It is *not* true, however, that (strict) mean value property implies linear homogeneity as it does proportionality. An example of this is the exponential mean[75], undoubtedly a mean, and therefore bounded by $\min(p_{it}/p_{i0})$ and $\max(p_{it}/p_{i0})$, and satisfying proportionality, but not linear homogeneity.

How linear homogeneity differs from mean value property can best be seen in the case of the value index V_{0t}: if all prices p_t^* are λ–fold prices p_t, ($p_{it}^* = \lambda p_{it}$) the value index will also display a change by λ, irrespective of what the quantities are. Thus V_{0t} is linearly homogeneous, and also monotonously increasing in prices $p_t^* > p_t$, but V_{0t} is not a mean value of price relatives, and V_{0t} does not necessarily indicate a λ–fold change if all prices equally change λ–fold.[76]

The same applies to chain indices: the result of the chain $\bar{P}_{02}^{LC} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \frac{\sum p_2 q_1}{\sum p_1 q_1}$ can well differ from λ although all prices are $p_{i2} = \lambda p_{i0} \quad \forall i$, since $\frac{\sum p_1 q_0}{\sum p_0 q_0} \frac{\sum \lambda p_0 q_1}{\sum p_1 q_1}$ does not necessarily equal λ unless *for example* all $q_1 = q_0$ such that $\frac{\sum p_1 q_0}{\sum p_0 q_0} \frac{\sum \lambda p_0 q_0}{\sum p_1 q_0} = \lambda$.

We can also show that the mean value of price relatives interpretation, does not apply to chain indices (see sec. 3.1) which should be seen as a *serious* defect of chain indices.

d) Additivity and linearity of index functions[77]

Strict monotonicity only requires that the index function reflects an increase or a decrease when prices increase or decrease, that is the *direction* of change that has to be indicated correctly. But nothing is said about the *amount* of increase or decrease the index should reflect. An interesting *subset* of the set of monotonous index functions is known as "additive"[78] (or also "linear") index functions. We also present a generalisation of the theorem of L. v. Bortkiewicz (see sec. 2.1) applicable to linear indices of all kind.

1. Additivity and multiplicativity as special types of strict monotonicity

Assume non–negative price vectors, $\mathbf{p}_t^*$ and $\mathbf{p}_0^*$ which are defined as sums of two price vectors then the function $P(\cdot)$ is said to be **additive** (or better: **linear** in the prices) if

$$P(\mathbf{p}_0, \mathbf{p}_t^*) = P(\mathbf{p}_0, \mathbf{p}_t) + P(\mathbf{p}_0, \mathbf{p}_t^+) = A + B, \qquad (2.2.15)$$

75 The same would apply to Stuvel's indices.

76 In other words, the value index is linearly homogeneous, but does not satisfy strict proportionality, identity and the mean value test.

77 Here the properties of "additivity" and "linearity" are considered only for a fixed t in P_{0t} taken in isolation The same is true for sec. 2.3b. These properties need a re–definition in case of *successive* indices, like $P_{0t}, \ P_{0t+1}, \ \ldots$ (see sec. 3.2b).

78 As will be shown in sec. 2.3 the term "additivity" is used in index theory in many different ways.

where $\mathbf{p}_t^* = \mathbf{p}_t + \mathbf{p}_t^+$, and

$$\frac{1}{P(\mathbf{p}_0^*, \mathbf{p}_t)} = \frac{1}{P(\mathbf{p}_0, \mathbf{p}_t)} + \frac{1}{P(\mathbf{p}_0^+, \mathbf{p}_t)} = \frac{1}{C} + \frac{1}{D} \tag{2.2.16}$$

where $\mathbf{p}_0^* = \mathbf{p}_0 + \mathbf{p}_0^+$.
It is easy to verify that both indices, P^L (Laspeyres) and P^P (Paasche) satisfy both conditions, but Fisher's ideal index P^F is not additive, although both components, P^L and P^P are additive. Non–linearity of Fisher's ideal index (direct version) will be discussed in detail in sec. 2.3 . Furthermore all sorts of chain indices are not linear either.
The two conditions, eq. 2.2.15 and eq. 2.2.16 are independent: Carli's index[79] and the index of Drobisch

$$P_{0t}^{DR} = \frac{P_{0t}^L + P_{0t}^P}{2} \tag{2.2.17}$$

satisfy 2.2.15 but not eq. 2.2.16 and the unweighted harmonic mean of price relatives, satisfies 2.2.16 but not eq. 2.2.15. Fisher's index is not additive with respect to both equations.
Not being additive has been regarded as a serious defect of Fisher's index (direct version) by PFOUTS (1966). For him the notion of additivity has an intuitive appeal:

> In case of additivity (in the sense of linearity) the difference between two indices reflects a difference between two price vectors; the amount of change indicated by the price index is a simple function of the change in individual prices.

In the absence of additivity no such inference from a difference between two price indices to a difference in prices is possible.To make additivity more plausible PFOUTS presents still another argument: Additivity "states if all prices are increased by the same amount , the index of the new price should equal the old index number plus the index number of the constant amount" (p.176). This requirement might be regarded as a *weaker version of additivity*: Then the vector $\mathbf{p}_t^+$ in eq. 2.2.15 is simply a vector of constants that is $\mathbf{p}_t^+ = \begin{bmatrix} b \\ \vdots \\ b \end{bmatrix}$, and we get

$$P_{0t}^{L,t*} = P_{0t}^L + b\frac{\sum q_0}{\sum p_0 q_0} = P_{0t}^L + \frac{b}{\tilde{p}_{0,0}} \tag{2.2.15a}$$

for the index of Laspeyres, and

$$P_{0t}^{P,t*} = P_{0t}^P + b\frac{\sum q_t}{\sum p_0 q_t} = P_{0t}^P + \frac{b}{\tilde{p}_{0,t}} \tag{2.2.15b}$$

[79] Unweighted arithmetic mean of price relatives.

for the index of Paasche, where $\tilde{p}_{0,0}$ and $\tilde{p}_{0,t}$ are a sort of general "unit value" (average price over all commodities) at time 0 calculated on the basis of quantities at time 0 or t respectively. In the same way it can be shown that P^L_{0t} as well as P^P_{0t} also fulfill eq. 2.2.16. By contrast P^{DR}_{0t} satisfies eq. 2.2.15 but not 2.2.16, simply because

$$\frac{1}{P^{DR}_{0t}} = \frac{1}{\frac{1}{2}\left(P^L_{0t} + P^P_{0t}\right)} \neq \frac{1}{2}\left(\frac{1}{P^L_{0t}} + \frac{1}{P^P_{0t}}\right) = \frac{P^L_{0t} + P^P_{0t}}{2P^L_{0t}P^P_{0t}}.$$

This can also be seen with the help of a small numerical example.

Example 2.2.1

Consider two commodities with prices and quantities as follows

i	p_{i0}	p_{it}	q_{i0}	q_{it}
1	12	18	10	8
2	15	12	8	10

To see the impact of different prices in the base period (0) as well as in the current period (t) assume the following vectors 1) $\mathbf{p}^+_t = \begin{bmatrix}3\\3\end{bmatrix}$ *and 2)* $\mathbf{p}^+_0 = \begin{bmatrix}3\\3\end{bmatrix}$. *Calculate* P^L_{0t}, P^P_{0t}, P^{DR}_{0t}, *the unweighted harmonic mean of price relatives and* P^F_{0t}.

ad 1)

formula	$P(\mathbf{p}_0, \mathbf{p}_t)$	$P(\mathbf{p}_0, \mathbf{p}^+_t)$	$P(\mathbf{p}_0, \mathbf{p}^*_t)$
	(1)	(2)	(3)
Laspeyres	276/240 = 1.15	54/240 = 0.225	330/240 = 1.375
Paasche	264/246 = 1.0732	54/246 = 0.2195	318/246 = 1.2927
Drobisch	$\frac{1}{2}\left(\frac{276}{240}+\frac{264}{246}\right) = 1.1116$	$\frac{1}{2}\left(\frac{54}{240}+\frac{54}{246}\right) = 0.2226$	$\frac{1}{2}\left(\frac{330}{240}+\frac{318}{246}\right) = 1.3338$
harmonic mean[a]	$\left[\frac{1}{2}\left(\frac{12}{18}+\frac{15}{12}\right)\right]^{-1} = 1.0435$	$\left[\frac{1}{2}\left(\frac{12}{3}+\frac{15}{3}\right)\right]^{-1} = 0.2222$	$\left[\frac{1}{2}\left(\frac{12}{21}+\frac{15}{15}\right)\right]^{-1} = 1.2727$
Fisher[b]	$\sqrt{\frac{276}{240}\cdot\frac{264}{246}} = 1.1109$	$\sqrt{\frac{54}{240}\cdot\frac{54}{246}} = 0.222239$	$\sqrt{\frac{330}{240}\cdot\frac{318}{246}} = 1.3332$

The sum of row (1) and row (2) equals

[a] *2882/2277 =1.2657 instead of 2898/2277 = 1.2727*

[b] *1.333161 instead of 1.333206*

ad 2)

formula	$1/P(\mathbf{p}_0, \mathbf{p}_t)$	$1/P(\mathbf{p}_0^+, \mathbf{p}_t)$	$1/P(\mathbf{p}_0^*, \mathbf{p}_t)$
	(1)	*(2)*	*(3)*
Laspeyres	*240/276 = 0.86957*	*54/276 = 0.19565*	*294/276 = 1.06522*
Paasche	*246/264 = 0.93182*	*54/264 = 0.20455*	*300/264 = 1.13636*
Drobisch[a]	$\left[\frac{1}{2}\left(\frac{276}{240}+\frac{264}{246}\right)\right]^{-1}$ *= 0.8996*	$\left[\frac{1}{2}\left(\frac{276}{54}+\frac{264}{54}\right)\right]^{-1}$ *= 0.2*	$\left[\frac{1}{2}\left(\frac{276}{294}+\frac{264}{300}\right)\right]^{-1}$ *= 1.099641*
harmonic mean	$\frac{1}{2}\left(\frac{12}{18}+\frac{15}{12}\right)$ *= 0.958*	$\frac{1}{2}\left(\frac{3}{18}+\frac{3}{12}\right)$ *= 0.208*	$\frac{1}{2}\left(\frac{15}{18}+\frac{18}{12}\right)$ *= 1.1667*
Fisher[b]	$\left(\sqrt{\frac{276}{240}\cdot\frac{264}{246}}\right)^{-1}$ *= 0.9002*	$\left(\sqrt{\frac{276}{54}\cdot\frac{264}{54}}\right)^{-1}$ *= 0.20005*	$\left(\sqrt{\frac{276}{294}\cdot\frac{264}{300}}\right)^{-1}$ *= 1.1002*

The sum of row (1) and row (2) equals
[a] *1.099616 instead of 1.099641*
[b] *1.1002031 instead of 1.1002156*

The example shows that both indices, the harmonic mean index and Drobisch's index each violate one of the two conditions while Fisher's index P_{0t}^F violates both conditions of linearity.◀

Multiplicativity of the index function is defined as follows (see EICHHORN and VOELLER (1976), p. 14):

$$P(\mathbf{p}_0^*, \mathbf{p}_t^*) = P(\mathbf{K}\mathbf{p}_0, \mathbf{L}\mathbf{p}_t) = P(\mathbf{p}_0, \mathbf{p}_t)\cdot\phi\,(\kappa_1, \ldots, \kappa_n, \lambda_1, \ldots, \lambda_n) \tag{2.2.18}$$

where **K** and **L** are diagonal matrices

$$\mathbf{K} = \begin{bmatrix} \kappa_1 & \ldots & 0 \\ & \ddots & \\ 0 & & \kappa_n \end{bmatrix} \quad \text{and} \quad \mathbf{L} = \begin{bmatrix} \lambda_1 & \ldots & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

and Φ is a function depending on the real numbers $\kappa_1, \kappa_2, \ldots, \kappa_n, \lambda_1, \lambda_2, \ldots, \lambda_n$ only such that ϕ is a positive real number. An example of an index function satisfying multiplicativity is the Cobb–Douglas–Index (see eq. 4.2.1). The logarithmic Laspeyres also known as Jöhr's index (see eq. 2.3.1), can however violate multiplicativity. ACZÉL and EICHHORN (1974) (see also EICHHORN / VOELLER (1976), p.18 – 21) proved the following theorem on additive indices: an index function $P(\cdot)$ satisfying the conditions of additivity (eq. 2.2.15 and 2.2.16) necessarily

1. meets the following condition

$$P\left[\mathbf{p}_0, \left(\frac{M}{N}\right)\mathbf{p}_t\right] = \frac{1}{N}P(\mathbf{p}_0, M\mathbf{p}_t) = \frac{M}{N}P(\mathbf{p}_0, \mathbf{p}_t), \tag{2.2.19}$$

   which implies linear homogeneity, homogeneity of degree −1 and (price) dimensionality (the converse is not true)

2. must have the following form

$$P^A = \frac{\mathbf{a}'\mathbf{p}_t}{\mathbf{b}'\mathbf{p}_0} \tag{2.2.20}$$

   which is an **additive index**, the ratio of two inner products[80], like the Laspeyres form ($\mathbf{a} = \mathbf{b} = \mathbf{q}_0$) or the Paasche form ($\mathbf{a} = \mathbf{b} = \mathbf{q}_t$) but not for example Fisher's ideal index. Eq 2.2.19 and eqs. 2.2.16 / 2.2.17 are equivalent.

Fisher's ideal index P^F does not fulfill the conditions of additivity (2.2.16 / 2.2.17) because P^F is not a ratio of inner products (2.2.19). The same is true for the logarithmic Laspeyres index, or for the quadratic mean index, defined as follows

$$P_{0t}^{QM} = \sqrt{\sum \left(\frac{p_t}{p_0}\right)^2 \frac{p_0 q_0}{\sum p_0 q_0}}. \tag{2.2.21}$$

On the other hand the indices of Laspeyres, Paasche, and Walsh, or even the value index are additive.
The price index P_{0t}^{QM} will be referred to in sec. 2.3 because of its interesting aggregation properties similar to additive indices though P_{0t}^{QM} not being additive in the sense defined above.

2. Generalization of Bortkiewicz's theorem for additive indices
The theorem of Bortkiewicz as introduced in sec. 2.1 is a special case of a more general law of the ratio of two additive indices (see fig. 2.2.6). The interesting result is eq. 2.2.21.
It can easily be seen that the *special case* of sec. 2.1 was given as follows: $X_0 = P_{0t}^L$, $X_t = P_{0t}^P$ and $\bar{Y} = Q_{0t}^L$. Only in this case the coefficients of variation, V_x and V_y are symmetrically defined, one representing the relative dispersion of price relatives, and the other the relative dispersion of quantity relatives.
The theorem (eq. 2.2.21) applied to the comparison of P_{0t}^L (Laspeyres) and P_{0t}^W (Walsh) gives:
$x_t = p_t$, $x_0 = p_0$, $y_0 = q_0$ and $y_t = \sqrt{q_0 q_t}$, hence the relevant variances are

- the variance of the relatives $\frac{x_t}{x_0} = \frac{p_t}{p_0}$ relative to the mean $\bar{X} = P_{0t}^L$, and

[80] **a** and $\mathbf{p}_t$ do not need to contain the same number of commodities as **b** and $\mathbf{p}_0$.

- the variance of the $\frac{y_t}{y_0} = \sqrt{\frac{q_t}{q_0}}$ measured around $\bar{Y} = \frac{\sum p_0 \sqrt{q_0 q_t}}{\sum p_0 q_0}$.

The extent to which Walsh's index, P_{0t}^{W} turns out to be greater or smaller than Laspeyres' index, P_{0t}^{L} depends on these variances in combination with the covariance between $\frac{p_t}{p_0}$ and $\sqrt{\frac{q_t}{q_0}}$ which is given by $\frac{\sum p_t \sqrt{q_0 q_t}}{\sum p_0 q_0} - P_{0t}^{L}\bar{Y}$. Thus if $\frac{p_t}{p_0}$ and $\frac{q_t}{q_0}$ are negatively correlated such that $P_{0t}^{L} > P_{0t}^{P}$ the same will be true for $\frac{p_t}{p_0}$ and $\sqrt{\frac{q_t}{q_0}}$ such that $P_{0t}^{L} > P_{0t}^{W}$. Hence we get: if $P_{0t}^{L} < P_{0t}^{P}$ then $P_{0t}^{L} < P_{0t}^{W}$ and if $P_{0t}^{L} > P_{0t}^{P}$ then also $P_{0t}^{L} > P_{0t}^{W}$.

e) Some other properties and uniqueness theorems

This section discusses some additional properties found useful at least by some authors. For Irving Fisher the most important tests were the so called reversal tests. They are however poorly motivated, and the "philosophy" behind them is controversial.

1. Fisher's reversal tests and the circular test (transitivity)

Fisher introduced three reversal tests, but the first one (the *commodity* reversal test)[81] is almost "trivial", so that nowadays it is rarely mentioned at all. The second and third ones are the two "great reversal tests" (as Fisher himself has put it), the *time*– and the *factor* reversal test. Furthermore in Fisher's theory the amount to which a pair of indices exceeds, or falls short of unity (in the case of the time reversal test), or the value ratio (factor reversal) was also a most important criterion, if not a measure of "bias". The three reversal tests were based on an argument of "fairness", or as Fisher said: "Index numbers to be fair ought to work both ways – both ways as regards ... the two times to be compared or as regard the two sets of associated elements for which index numbers may be calculated – that is, prices and quantities" (FISHER (1922), p. 62)[82].

By the *factor reversal test*, as mentioned already, the product of a price and a quantity index should equal the value index V_{0t} (expenditure ratio) $P_{0t}Q_{0t} = V_{0t}$. The formula of Q_{0t} should be gained by interchanging price vectors and quantity vectors in the formula of P_{0t} (that is Q_{0t} is the direct quantity index, otherwise we speak of the *product test* which is a weak factor reversal test)[83].

The factor reversal idea played an ignominious role in at least three ways:

1. it inspired the search for "ideal" index formulas, laborious and of only little use,

[81] According to this test (also known as symmetry axiom) the order (sequence) of the commodities ought to make no difference for the result of the index calculation.

[82] That this type of reasoning is not tenable can be seen from the following fact: while the commodity reversal test is almost trivial, there are only very few index functions that comply with the factor reversal test. Time reversibility appears to have a middle position in this respect.

[83] The so called "value index preserving test" is another weakening of factor reversibility (see next part of this subsection).

2. in combination with a concept of "bias" it is said that a formula P_{0t} which tends to the result $P_{0t}Q_{0t} > V_{0t}$ is biased upwards whilst if $P_{0t}Q_{0t} < V_{0t}$ the index allegedly has a downward bias, and from here hails much of the undeserved criticism of the Laspeyres formula, and of the admiration of Fisher's ideal index, (likewise undeserved), and finally

3. this test is often supposed to be of considerable importance, if not indispensable for "deflation" of aggregates, in order to have a value, or an increase of value decomposable unambiguously into two components, a price and a quantity component with no residual term left.

A meaningful decomposition of value into a price and a quantity component, however, can also be achieved when only the weaker product test is met[84]. The motivation to require the much stronger, and highly restrictive factor reversal test instead of the product test is unclear. In our view, as already mentioned, the only reason could be the desire to do both, inflation measurement and deflation (with a "deflator" price index) with the help of the same price index. But there are quite a few reasons why it is futile to seek a single indicator serving both purposes equally well. For all these reasons the *factor reversal* criterion is not only lacking a firm foundation, but what is even worse, the idea of this criterion is the source of much confusion about and injustice to index formulas.

The so called *time reversal test*, introduced already, requires that $P_{t0} = (P_{0t})^{-1}$, that is interchanging base time (0) and comparison (reference) time (t) should yield the inverse result. The *circular test* (*transitivity*) in combination with identity implies time reversibility (the converse is not true). Transitivity will be considered in more detail in chapter 4.

The motivation of such tests, however, is astoundingly poor. Fisher himself did not elaborate very detailed arguments in favor of his time reversal (and also factor reversal) test[85], and the circular test was always controversial in his philosophy. As will be shown in chapter 4:

> Ironically ideas like time reversibility and chainability (transitivity) having some significance in the chain index framework are *not* sufficiently motivated in case of *intertemporal* comparisons. They gain importance mainly for *interspatial* (place to place, e.g. international) comparisons, a situation to which the rationale of chain indices and of the Divisia index is *not* applicable.

[84] An index formula unable to pass the product test is for example the Törnquist index. This is a serious defect of a formula which is one of the most favored formulas in the SNA and in the "economic theory of index numbers".

[85] His ideas were mainly (1) since individual price relatives always meet these tests (see table 0.3.1) it seems desirable that they also hold for aggregative price and quantity relatives i.e. for indices, and (2) the above mentioned argument of "fairness".

Figure 2.2.6: Generalisation of Bortkiewicz's theorem

first additive index X_0	second additive index X_t
$X_0 = \frac{\sum x_t y_0}{\sum x_0 y_0}$	$X_t = \frac{\sum x_t y_t}{\sum x_0 y_t}$

weighted (weights $w_0 = x_0 y_0 / \sum x_0 y_0$ throughout) arithmetic mean of the relatives

x_t/x_0: $\bar{X} = X_0$	y_t/y_0: $\bar{Y} = \frac{\sum y_t x_0}{\sum y_0 x_0}$

variances (weights w_0) of the relatives

x_t/x_0: $s_x^2 = \sum \left(\frac{x_t}{x_0} - \bar{X}\right)^2 w_0$	y_t/y_0: $s_y^2 = \sum \left(\frac{y_t}{y_0} - \bar{Y}\right)^2 w_0$

the covariance is given by

$$s_{xy} = \sum \left(\frac{x_t}{x_0} - \bar{X}\right)\left(\frac{y_t}{y_0} - \bar{Y}\right) w_0 = \frac{\sum x_t y_t}{\sum x_0 y_0} - \bar{X}\bar{Y}$$

and the ratio of two additive indices

$$\frac{X_t}{X_0} = 1 + r_{xy} V_x V_y \quad \text{where } r_{xy} = \frac{s_{xy}}{s_x s_y}, V_x = \frac{s_x}{\bar{X}} \text{ and } V_y = \frac{s_y}{\bar{Y}} \tag{2.2.22}$$

The *country reversal test* is defined by $P_{BA} = 1/P_{AB}$ (the first subscript denoting the "base" country). The reason for this is to avoid ambiguity: there is only one estimate of the purchasing power parity between two countries, because one parity (P_{BA}) is simply the inverse of the other parity (P_{AB}). The rationale of this test seems to be clear: reaching the conclusion that A is k times as much as B it would be hard to understand why B should not be one k–th of A. But this kind of reasoning is no longer convincing once it is acknowledged that a parity has to be established on the basis of

weights belonging to either or both countries, A and B. Then it is no longer evident that the two countries, A and B should be treated symmetrically. On the contrary it is often said, that country reversibility is in conflict with "representativity" for a certain country (also called "characteristicity").
With regard to two countries, A and B there *seems* to be no reason to prefer one of the two parities, P_{AB} and P_{BA} respectively, to the other, but with regard to two points in *time*, that is in the intertemporal framework (P_{0t} and P_{t0}) it should be recognized, as already mentioned, that 0 and t are not just two points in time but rather one single, fixed period 0 and a "period" t which is a variable, a multitude of points in time. Hence there is no need for "symmetry".
A similar argument applies to the "circular test". In international comparisons many more meaningful comparisons exist than in the intertemporal framework. As an example we may make a comparison between country A and B *indirectly* via a third country C or via a third and a fourth country, such that linking P_{AC} with P_{CB} or P_{AC} and P_{CD} with P_{DB} should yield the same result as obtained by the direct comparison. There is no such need, however, to make indirect comparisons such as P_{03} with P_{35} or P_{02} and P_{27} with P_{75} compatible with the direct comparison P_{05} when the variable *time* is concerned.
There are many similarities, but also many dissimilarities between intertemporal and interregional price comparisons. An argument, that is convincing for interregional price comparison must not be valid in the case of intertemporal comparison and vice versa. To sum up:

The need for the so called "reversal tests" as well as for the requirement of transitivity (consistency between direct and all sorts of indirect comparisons) is not compelling as they are far from being well reasoned, but rather motivated mainly

- in terms of "symmetry", "fairness", or "analogy" of indices with relatives, and
- by a mistaken analogy of intertemporal and interspatial comparisons.

Unfortunately the idea that a "composite commodity" should be treated like a single commodity[86] still has its adherents these days. So HANSEN and LUCAS (1984) concluded: "Following from the purpose and intent of a composite commodity is that it should be a composite of the separate properties of its components, and that these composite properties should behave like the component properties." And this is what they call "*the analogue approach to index numbers*" originating from Fisher but still valid in their view. They call the reversal tests and circularity "*analogue tests*", and of course they not only find time reversal necessary[87], but also favor "ideal" indices as being "unbiased" (by *definition*!) and plead for circularity. It speaks volumes that they

- refrain from considering aggregation problems altogether, and

[86] And therefore the idea of "analogy" between indices and relatives.
[87] No one actually criticizes this property, but it is often ignored" (ibid. p. 27)

- see no gap between their "analogue approach", and the "economic theory approach" they found no less useful, but which is of course pointless when there is in fact only one single commodity, and finally they are
- vigorously in favor of chain indices.

They not only argued that circularity is satisfied "for every chain linked index" (p. 30) but also appeared convinced of the well known argument[88]: "...every index has to be rebased and linked fairly frequently for purely practical reasons, so there is an obvious temptation to carry this to its logical conclusion and chain on an annual basis."[89] Why is the "logical conclusion" not the beloved "instantaneous index" (p. 31) of Divisia? We most strictly disagree with all these curious "conclusions" pretending to be based on indisputable principles: "our conclusion, then, of the analogue test approach is that, rather than being artificial or arbitrary, it follows naturally from the basic inclusion in economics of composite commodities therefore, in general, the formula default should be an ideal chain–linked index." (p. 28).
The conclusions are not tenable and if they were it would impressively indicate that principles from which they pretendedly follow quite "naturally", that is Fisher's "symmetry", "fairness", or "analogy" philosophy, must be flatly wrong. Unfortunately Fisher's philosophy bears its years well, and still lives on in textbooks of statistics, articles on index theory and international conferences of price statisticians.

2. Weak time reversal test, price reversal test and a uniqueness theorem for Fisher's index
The following *weakening of the* rather restrictive *time reversal test* is interesting to consider:

$$\text{if } P_{0t} > 1 \text{ then } P_{t0} < 1 \text{ and if } P_{0t} < 1 \text{ then } P_{t0} > 1. \tag{2.2.23}$$

This requirement seems to be reasonable, not too ambiguous, and *sufficient* to postulate: it is only required that an increase in the direction $0 \to t$ should correspond to a decline in the opposite direction $t \to 0$ and vice versa. In the case of the Laspeyres price index this is equivalent to: if $P^L_{0t} > 1$ then $P^P_{0t} > 1$ and if $P^L_{0t} < 1$ then $P^P_{0t} < 1$, because $P^L_{t0} = 1/P^P_{0t}$ and $P^P_{t0} = 1/P^L_{0t}$.
It is interesting to note that this is not necessarily the case.
Two price indices P and P^* are said to satisfy the *price reversal test* if the price vectors in $P_{0t} = P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t)$ are interchanged such that $\mathbf{p}_0 \to \mathbf{p}_t$, $\mathbf{p}_t \to \mathbf{p}_0$ and the quantity vectors remain unchanged such that we obtain the index $P^*_{0t} = P(\mathbf{p}_t, \mathbf{q}_0, \mathbf{p}_0, \mathbf{q}_t)$, and

$$P_{0t}P^*_{0t} = P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t)P(\mathbf{p}_t, \mathbf{q}_0, \mathbf{p}_0, \mathbf{q}_t) = 1 \tag{2.2.24}$$

holds. Note that the time reversal test requires $P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t)P(\mathbf{p}_t, \mathbf{q}_t, \mathbf{p}_0, \mathbf{q}_0) = 1$ which is different in that not only price vectors, but also quantity vectors are interchanged. It is easy to see that the Laspeyres price index satisfies the price reversal test

[88] called "why not"– argument in chapter 6.
[89] ibid. p. 28.

since in P^L_{0t}, an isolated interchanging of price vectors leads to $P^{*L}_{0t} = \frac{\sum p_0 q_0}{\sum p_t q_0}$ and in fact $P^L_{0t} P^{*L}_{0t} = 1$.

On the basis of this test VOELLER (1976) formulated a *uniqueness theorem*[90] which states that Fisher's ideal index P^F_{0t} is the only index satisfying the following three reversal tests simultaneously: 1. time reversal, 2. factor reversal, and 3. price reversal test. However, the problem is that the meaning, and the usefulness of the price reversal test (taken in isolation, let alone considered in combination with the other two reversal tests) is open to debate.

The meaning of a uniqueness theorem is as follows: Fisher's ideal index can be *characterized* as being the *only* index that simultaneously satisfies all three reversal tests. The interpretation becomes even more complicated, as according to another uniqueness theorem of Voeller, Fisher's ideal is also the *only* index that satisfies the

1. time reversal, 2. factor reversal, and 3. quantity reversal test simultaneously,

where quantity reversibility requires $P_{0t} = P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t) = P^*_{0t} = P(\mathbf{p}_0, \mathbf{q}_t, \mathbf{p}_t, \mathbf{q}_0)$, a price index that is invariant upon interchanging of quantities (or in other words: the index should be symmetric with respect to the quantity vectors). Some authors are inclined to conclude as follows: more uniqueness theorems for a certain index formula is equivalent to superiority of the formula in question. In our view this is by no means convincing, on the contrary, because as a rule at least for one of the axioms combined in a theorem an interpretation is difficult to find, and the relation between a number of such uniqueness theorems is obscure.

3. Proportionality with respect to quantity indices, the value index preserving test

A test called "value index preserving test"[91] by VOGT (1978a), (1979), (see also BALK (1995)) requires that a price index should be P^L_{0t} if no quantity changes occur and therefore $Q_{0t} = 1$.

It is an implication of proportionality ($Q(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \lambda\mathbf{q}_0) = \lambda$ when $\mathbf{q}_t = \lambda\mathbf{q}_0$) and hence also identity ($\lambda = 1$) applied to a quantity index given that the quantity index in combination with a price index satisfies the product test. Thus

$$P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \lambda\mathbf{q}_0) \underbrace{Q(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \lambda\mathbf{q}_0)}_{=\lambda=1} = V_{0t} = \frac{\sum p_t q_0}{\sum p_0 q_0} = P^L_{0t} = \frac{\sum p_t q_t}{\sum p_0 q_t} = P^P_{0t}. \tag{2.2.25}$$

> Note that the *direct* Fisher price index as a deflator will satisfy this test, but the *chain* variant of this index will *not*, which will give rise to some interesting insights into deflation using chain indices (see chapter 5).

[90] see also EICHHORN and VOELLER (1983), p. 443f, BALK (1995) (there "theorem 6").
[91] Wertindextreue–Test.

4. Value dependence test and another uniqueness theorem for Fisher's ideal index
The idea of the so called "value dependence test" (not to be mixed up with the "value index preserving test") is not easy to understand, because the requirement expressed in it seems to be rather arbitrary and not intuitively convincing. By this test the function

$$P_{0t} = P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t) \qquad \mathbb{R}^{4n}_{++} \to \mathbb{R}_{++} \tag{2.2.26}$$

and the following function

$$P_{0t} = f\left(\sum p_0 q_0, \sum p_0 q_t, \sum p_t q_0, \sum p_t q_t\right) \tag{2.2.27}$$

$f(a, b, c, d)$ should yield the same result.
By virtue of this test the index formula can be expressed as some kind of combination of four "value" terms. For some index functions capable of passing this test see tab. 2.2.1. Functions that are unable to pass this test are for example the Logarithmic Laspeyres or Walsh index. The product of P^F and Q^F is d/a, which is the value index $V_{0t} = \sum p_t q_t / \sum p_0 q_0$.

> Since eq. 2.2.26 does not apply to chain indices but rather for example
>
> $$\bar{P}_{02} = \bar{P}_2(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_1, \mathbf{q}_1, \mathbf{p}_2, \mathbf{q}_2) \qquad \mathbb{R}^{6n}_{++} \to \mathbb{R}_{++} \tag{2.2.28}$$
> $$\bar{P}_{03} = \bar{P}_3(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_1, \mathbf{q}_1, \mathbf{p}_2, \mathbf{q}_2, \mathbf{p}_3, \mathbf{q}_3) \qquad \mathbb{R}^{8n}_{++} \to \mathbb{R}_{++} \tag{2.2.28a}$$
>
> and so on eq. 2.2.27 does not apply either. Thus the value dependence test is not applicable to chain indices (in contrast to a link for example).

Table 2.2.1: Some examples for indices satisfying the value dependence test

Formula	f in case of price index	f in case of quantity index
Laspeyres	$P^L_{0t} = \frac{c}{a}$	$Q^L_{0t} = \frac{b}{a}$
Paasche	$P^P_{0t} = \frac{d}{b}$	$Q^P_{0t} = \frac{d}{c}$
Marshall and Edgeworth	$P^{ME}_{0t} = \frac{c+d}{a+b}$	$Q^{ME}_{0t} = \frac{b+d}{a+c}$

The two Dutch index theoreticians J. van Yzeren and B. M. Balk have shown that Fisher's index is the *only* formula that simultaneously satisfies the following three tests: factor reversal, linear homogeneity, and value dependence.
However, the problem of this finding is again to find some intuitively appealing interpretation that could be given to this result and to demonstrate the specific advantages of Fisher's formula due to the fact that it is the *only* formula that satisfies the three tests.

2.3 Deflation and aggregation

Price indices are compiled for two different purposes, firstly to provide measures of the year to year movement of specific price *levels* (e.g. the rise or decline of consumer prices) and secondly to "deflate" aggregates.[92] The focus of this section is to examine deflation in more detail and in combination with some aggregation properties, for which a confusing variety of names are known, as for example "consistency in aggregation", "structural consistency" or "additivity" or so. In a sense this section is also an extension and continuation of the axiomatic approach of sec. 2.2. Another purpose is to shed some light on aspects apparently overlooked when the so called "ideal index" of Fisher, as direct or chain index even, is recommended for the purpose of deflation. The suitability of such formulas, and hence the wisdom of such recommendations is particularly questionable in the context of deflation (see chapter 5).

a) Objectives and types of deflation

The operation called "deflating" or "deflation" is usually defined as isolating the volume component (that is quantity and quality) in an aggregate. In this sense we have already introduced the concept "deflation" in sec. 0.2b: the task is to derive an aggregate at constant prices (quantities q_{it} valued at constant prices p_{i0}, or "volume" $\sum p_0 q_t$) from an aggregate at current prices (known as "value" $\sum p_t q_t$).[93] We are dealing with essentially the same problem as if quantities referring to country A (denoted by q_{iA}) are to be valued at prices p_{iB} (another country B) instead of prices p_{iA} (the same country A). Therefore it may be useful to regard deflation as the intertemporal variant of a much wider concept, which might be called "*revaluation*":

[92] We should regard the two tasks, measurement of price levels and deflation of aggregates respectively as two distinct topics in index theory, each of which has a "logic" of its own. As already pointed out repeatedly it is our position that there is no need to strive for an index formula capable of doing both jobs at a time. In practice we will always have different indices due to differences in scope, definitions, and selections of goods and services to which an index applies, which should serve as an inflation measure or as a deflator respectively. Hence it is *pointless* to require *factor reversibility*. For decades already (since Fisher's days) index theory has unduly paid much attention to this property, and it is difficult to understand why it still continues to do so in our days.

[93] It should be noticed that the model of only two components in values and value change, that is a quantity and a price component may be regarded as too simplified. Furthermore it is important that both aggregates, values and volumes, should refer to the *same* collection of n commodities, that is the summation should take place over the same $i = 1, \ldots, n$ commodities in both cases.

Generally the aim of revaluation is to compare values under different regimes of prices. Each revaluation is somewhat fictitious: quantities and prices that are not observed at the same time or at the same place are combined. When pricing of a constant basket (assigning prices p_t to quantities q_0) is found inacceptable then – strictly speaking – a "volume", in which quantities q_t are multiplied by prices other than p_t is doubtful as well.

Valuation at constant prices of aggregates observed at current prices can be done in two fundamentally different ways, depending on what is intended to be calculated, "volumes" or "real" aggregates (see fig. 2.3.1). Both types of deflation, that is the isolation of a "quantity component" ("in volume terms") as well as "deflating" in the sense of adjusting for inflation ("in real terms")[94], are particularly important in the context of National Accounts (NA).

Unlike the "volume" approach the "real income" approach is not limited to aggregates representing sales or purchases of (a certain quantity of) goods or services, that is to *commodity flows*(CFs)[95], but is also applicable to *non commodity flows* (NCFs), that is aggregates lacking an observable well defined "quantity", by definition (for example income, tax and interest payments, bank deposits, balancing items in NA and the like).

The most prominent example of an NCF is value added, which is a balancing item in production accounts. The so called *double deflation* method consists in applying the volume oriented method to this type of aggregates, defined as differences between two CF aggregates[96].

The following considerations are restricted to the "volume approach" and to CFs so that there is no problem in defining (or even summating over) "quantities" of goods "belonging" to the aggregate in question. A distinction is sometimes made between a

- *direct* method of calculating volumes or volume (= quantity) indices by estimation or extrapolation of some relevant indicators of *quantity*, and
- an *indirect* approach, by repricing each commodity (with p_{i0} instead of p_{it}) individually or (of course more conveniently) collectively (in a global manner), using an appropriate comprehensive (covering a group of commodities) *price* index (also known as the "deflator").

The traditional deflator is a direct Paasche price index (P^P_{0t}) but many other formulas are recommended as well[97]. The indirect method of deflation (using price indices)

94 The task here is to calculate the amount of money, the value would represent, if the purchasing power of money had been kept constant, or in other words, if inflation had not taken place. Note that a correct accounting of inflation would require much more, however. It is far from clear whoever is the "winner" or the "looser" when inflation had taken place.

95 The term "flow" here always comprises of stocks also.

96 The justification of this method is highly controversial, since double deflation may yield somewhat awkward results (e.g. negative real value added).

97 As already mentioned international recommendations (U.N. and European Union) prefer the direct or even the chain version of Fisher's index. This will be discussed in more detail in sec. 2.3f and in chapter

may be preferable to the direct method (observing directly the relevant quantities) because:

1. statistics on values and prices may be more readily available than those on quantities,

2. as price relatives tend to vary less than quantity relatives, the sampling error associated with the indirect method is likely to be smaller, and

3. it is in general easier to deal with new products in the framework of price indices than to account for such phenomena correctly in quantity indices.

In sec. 2.3e we will propose some criteria to distinguish between "good" and "bad" deflation. The idea is to make indirect deflation consistent with direct deflation.

Moreover the coexistence of CFs and NCFs, and therefore of both methods, the "volume approach" and the "real income approach"[98] in the practice of NA strongly requires consistency of both methods with one another.

In the deflation theory we are used to distinguishing between two components of value only, a price and a quantity[99] component. According to some authors (for example SCHIMMLER (1973)) this is insufficient, and components like "structural changes in physical quantities", "changes in quality structure" (usually accounted for in the volume [quantity] component), or "change in price structure" should also be measured in order to avoid reflection of "impure" effects.[100]

But this gives rise to a number of problems:

- The extent to which a structural component can be made visible is a question of the *level of disaggregation*. An index A, with aggregates broken down into n subaggregates A_i $(i = 1, 2, \ldots, n)$ will be more influenced by "structural" changes than B, an index in which only m subaggregates B_j $(j = 1, 2, \ldots, m$ and $m < n)$ are distinguished.

- It will be difficult to conceptualize a "structural" component distinct from the "level" component. Does for example "structure" with respect to qualities or prices permit distinctions like "more" or "less" change?

5.

98 There are also some concepts directly related to the comparison of the two methods, for example the so called terms of trade effect.

99 Also called "volume component" by some authors because it also includes a quality change (when price indices make adjustments for them).

100 For Schimmler the traditional approach to deflation or "two–component system", as he calls it is unable to depict pure effects.

b) Aggregative properties of index functions and volumes (overview)

For a long time index theory has not paid much attention to aggregative properties of index functions. The reason is presumably that price level measurement used to be the only, or at least the dominant purpose for which price indices were needed. As *deflation* became a more and more important field of price statistics, more emphasis was put on aggregation problems.

Favorable aggregative properties are also desirable for *analytical purposes*, i.e. it should easily be possible to decompose or aggregate price indices. A lot of analysis is done by compiling an index in various experimental versions, for example by excluding or including certain series of prices, because such calculations will give some insight into the structure of prices, as well as the causes of an increase or decrease of the overall price level.

Figure 2.3.1: Alternative procedures of deflation

Deflating aggregates*

in volume terms	**in real terms**
"at constant prices"	at constant purchasing power of money
Volume–approach (quantity interpretation of the result intended); isolation of the quantity component	Real–income–approach (adjusting income for inflation): the aim is to estimate the effect of inflation
Applicable to value changes only, that can be decomposed into price– and quantity–changes, i.e. to commodity flows (CFs)	Applicable also to values that do not have price and quantity dimensions on their own, i.e. to non commodity flows (NCFs)
Deflate aggregates $A_1, A_2, \ldots, A_m$ (representing CFs) by using m price indices P_i $(i = 1, \ldots, m)$ for the m aggregates (P_1 refers to the commodities in A_1 and P_2 to A_2 and so on)	Deflate aggregates $A_1, A_2, \ldots, A_m$ (all or many of them NCFs) using *one single* deflator that accounts for the general development of the purchasing power

the typical problem of applying the method is

Double deflation (i.e. deflating $C = A - B$ by deflating aggregates A and B separately)	choice of *the* general (universal) deflator, that has to be applied to all aggregates

* stocks or flows, see text for abbreviations CF and NCF.

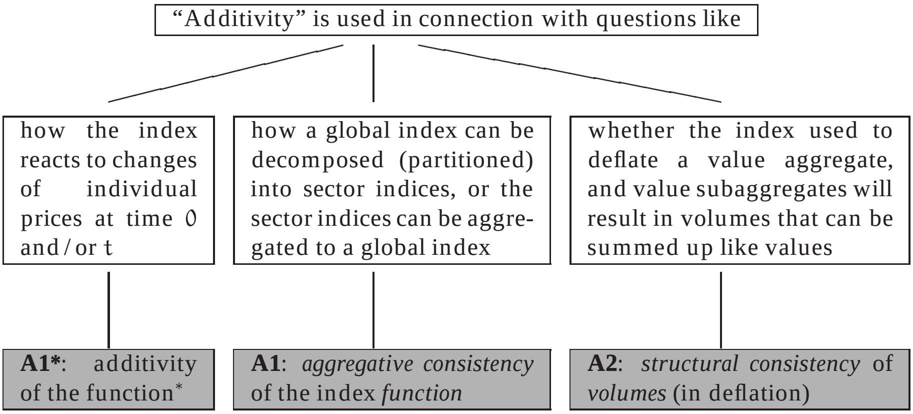
100 See footnote 77 on page 67.

The term "additivity" is used in various meanings, and the relationship between these types of "additivity" is by no means self–explanatory (See fig. 2.3.2). We may distinguish between additivity

- of the index function in the context of aggregation or disaggregation, and
- additivity with respect to properties of aggregates resulting from using this function for deflation.

Additivity in the sense of **A1**, called **aggregative consistency**, refers to the suitability of the index function for analytical purposes. It focuses on the nature of the relationship between a *global* (overall) index P and the sub–indices, or *sector* indices $(P_1, \ldots, P_K)$ the global index is made off. The function f in $P = f(P_1, \ldots, P_K)$ is not necessarily linear (= additive in the sense of **A1***).

Figure 2.3.2: The usage of the terms "additivity" in index theory



* The notion **A1*** also called *linearity* (see sec. 2.2d) is a special case of **A1**

It is sufficient for **A1**, that the function f is the *same* type of function, by which

- price relatives are combined (as a rule, by taking a weighted average) to sector indices P_k $(k = 1, \ldots, K)$, and then sector indices are combined to calculate the global index P, such that $P_{0t} = f(P^k_{0t})$, $k = 1, \ldots, K$.
- or by which price relatives are directly aggregated to the global index.

The property **A1**, also known as "consistency in aggregation" is of considerable importance for the analytical usefulness of an index for example as a measure of a price "level":

> Additivity in the sense of **A1** (= aggregative consistency) greatly facilitates the work done by analysts, which consists to a great deal, in examining the time path of, and correlation between indices of different composition. Consequently the lack of aggregative consistency, as in the case of chain indices is a serious defect for the analytical usefulness of an index.

The second aspect (**A2**) deals with deflation. **Structural consistency**[101] **in volumes** (deflated values), means that it is ensured that the same definitional equations satisfied by values should also hold for volumes derived from values, by using the appropriate sectoral and global price indices. It turns out that this is a highly restrictive requirement, and it can easily be shown (though) that this condition is only met by direct Paasche price indices.[102]

It is noteworthy to recognize that none of the three aggregational properties **A1***, **A1**, and **A2** will be satisfied by Fisher's "ideal index".[103]

c) Aggregative consistency of the index function (A1) and the equality test

Consistency in aggregation, also frequently simply called "additivity", is related to the aggregation of sector–indices to global indices. In official statistics the *global* (total) aggregate, is often subdivided into *sectors* (subaggregates). Consider n commodities that can be classified into K nonoverlapping (distinct), and exhaustive (all n goods comprising, such that $n = n_1 + n_2 + \cdots + n_K$) groups or "sectors" ($k = 1, \ldots, K$). For such a breakdown of an aggregate into K "sectors", or subaggregates a classification or "nomenclature" is usually needed. The total structure of an aggregate in practice contains several levels, four or five for example.

> The index function $P_{0t} = f(\cdot)$ is called *aggregative consistent* if
> 1. the function f of n price relatives a^i_{0t} used to derive the global index P_{0t} and the function f used to
> 2. calculate K sector indices $P^1_{0t}, \ldots, P^K_{0t}$ of $n_1, \ldots, n_K$ price relatives ($n = n_1 + \cdots + n_K$), and to
> 3. aggregate over the K sector indices to get the global index $P_{0t} = f(P^1_{01}, \ldots P^K_{0t})$
>
> is the same type of function throughout.

It can be shown that index formulas, satisfying **A1*** (like the Laspeyres, Paasche, Marshall–Edgeworth index and many other index formulas) will also satisfy **A1**, but

[101] The SNA also calls this "additivity".

[102] Surprisingly this is very little known, let alone the significance of this result.

[103] As shown below Fisher's index is not even capable of meeting the "weak" variant of **A1**, that is the "equality test".

not vice versa. There are index functions that are aggregatively consistent without being linear, like the *Logarithmic Laspeyres* index (very much advocated by W. A. Jöhr),

$$P^J_{0t} = \prod \left(\frac{p_{it}}{p_{i0}}\right)^{g_i} \quad \text{or} \quad \ln\left(P^J_{0t}\right) \sum \ln(a^i_{0t}) g_i, \tag{2.3.1}$$

where $a^i_{0t} = p_{it}/p_{i0}$ and $g_i = p_{i0}q_{i0}/\sum p_{i0}q_{i0}$ or the *Quadratic mean* index given by eq. 2.2.21, that is $P^{QM}_{0t} = \sqrt{\sum \left(a^i_{0t}\right)^2 g_i}$.

It can easily be seen (ex. 2.3.1 will demonstrate this), that the global index P^L (Laspeyres) is a weighted arithmetic mean of price relatives (with weights $g_i = p_{i0}q_{i0}/\sum p_{i0}q_{i0}$), but P^L can also be calculated as weighted arithmetic mean of the sector indices P^{Lk}_{0t} $(k = 1, \ldots, K)$ with weights defined analogously. The same is also true for the Paasche index P^P, with weights, $w_i = p_{i0}q_{it}/\sum p_{i0}q_{it}$.

But in the case of Fisher's index this is *not* possible. Fisher's ideal index P^F is not linear (additive in the sense of **A1***), and not aggregatively consistent (**A1**) either. We are unable to calculate the global index, P^F from K sector indices P^{Fk} and from n price relatives in *the same way*. With n individual price relatives we get

$$P^F_{0t} = \sqrt{\left(g_1 a^1_{0t} + g_2 a^2_{0t} + \cdots + g_n a^n_{0t}\right)\left(w_1 a^1_{0t} + w_2 a^2_{0t} + \cdots + w_n a^n_{0t}\right)}. \tag{2.3.2}$$

The result of the first expression (in brackets) is P^L_{0t}, and the result of the second expression is P^P_{0t}. But using K Fisher type sector indices the formula for P^F_{0t}, when sector indices are aggregated, instead of individual price relatives is *not* given by

$$\sqrt{\left(g_1 P^{F1}_{0t} + g_2 P^{F2}_{0t} + \cdots + g_K P^{FK}_{0t}\right)\left(w_1 P^{F1}_{0t} + w_2 P^{F2}_{0t} + \cdots + w_K P^{FK}_{0t}\right)} \tag{2.3.3}$$

where "sector Fisher indices" are defined by $P^{Fk}_{0t} = \sqrt{P^{Lk}_{0t} P^{Pk}_{0t}}$, but rather by

$$P^F_{0t} = \sqrt{\left(g_1 P^{L1}_{0t} + g_2 P^{L2}_{0t} + \cdots + g_K P^{LK}_{0t}\right)\left(w_1 P^{P1}_{0t} + w_2 P^{P2}_{0t} + \cdots + w_K P^{PK}_{0t}\right)} \tag{2.3.4}$$

where $P^{Fk}_{0t} = \sqrt{P^{Lk}_{0t} P^{Pk}_{0t}}$. Thus P^F_{0t} is *not* aggregatively consistent. The following example will demonstrate this.

Example 2.3.1

For the sake of simplicity we assume $K = 2$ and call the two sectors A and B. Prices and quantities of four commodities (goods) 1, 2, 3 and 4, classified into the two groups (or sectors) A and B are given as follows:

		base period (0)			*current period (t)*			
	good	*price p_0*	*quantity q_0*	p_0q_0	*price p_t*	*quantity q_t*	p_tq_t	p_0q_t
A	*1*	*10*	*16*	*160*	*15*	*12*	*180*	*120*
A	*2*	*16*	*15*	*240*	*18*	*14*	*252*	*224*
B	*3*	*20*	*10*	*200*	*24*	*12*	*288*	*240*
B	*4*	*15*	*20*	*300*	*21*	*15*	*315*	*225*
sum (Σ)				*900*			*1035*	*809*

The following index formulas should be used to calculate both, sector indices and the global index: the formula of Laspeyres, Paasche, Drobisch $\left(P_{0t}^{DR} = \frac{1}{2}(P_{0t}^{L} + P_{0t}^{P})\right)$*, Fisher, logarithmic Laspeyres* $\left(P_{0t}^{J}\right)$*, and the Quadratic mean price index* $\left(P_{0t}^{QM}\right)$*. It should also be shown how the global index can be derived from the sector indices.*
Solution: The price relatives a_i *and the weights* $g_i = p_0q_0 / \sum p_0q_0$ *and* $w_i = p_0qt / \sum p_0q_t$ *are*

good	a_i	g_i global	g_i within sector	w_i global	w_i within sector
1 (A)	1.5	0.1777	0.4 = 160/400	0.1483	0.3488
2 (A)	1.125	0.2667	0.6 = 240/400	0.2768	0.6511
3 (B)	1.2	0.2222	0.4 = 200/500	0.2967	0.5161
4 (B)	1.4	0.3333	0.6 = 300/500	0.2781	0.4839

We get the following sector indices

sector price index	sector A	sector B
Laspeyres	$P_{0t}^{LA} = 510/400 = 1.275$	$P_{0t}^{LB} = 660/500 = 1.32$
Paasche	$P_{0t}^{PA} = 432/344 = 1.2558$	$P_{0t}^{PB} = 603/465 = 1.2968$
Drobisch	$P_{0t}^{DR\ A} = 1.2654$	$P_{0t}^{DR\ B} = 1.3084$
Fisher	$P_{0t}^{FA} = 1.2654$	$P_{0t}^{FB} = 1.3083$
Log. Laspeyres	$P_{0t}^{JA} = 1.2622$	$P_{0t}^{JB} = 1.3163$
Quadratic mean	$P_{0t}^{QM\ A} = 1.28817$	$P_{0t}^{QM\ B} = 1.32361$

With weights g_i *and* w_i *given by*

weights*	sector A	sector B
weights g	$g_A = 400/900 = 0,4444$	$g_B = 500/900 = 0,5556$
weights w	$w_A = 344/809 = 0.4252$	$w_B = 465/809 = 0.5748$

* *expenditure shares*

Laspeyres' (Paasche's) global index can be calculated as weighted arithmetic means of sector indices. Therefore the resulting global indices are:
Laspeyres $P_{0t}^{L} = 0,4444 \cdot 1.275 + 0,5556 \cdot 1.32 = 1.3 = 1170/900$ *and Paasche* $P_{0t}^{P} = 0,4252 \cdot 1.2558 + 0,5748 \cdot 1.2968 = 1.2794 = 1035/809$.
The logarithmic Laspeyres sector indices are $P_{0t}^{JA} = 1.5^{0.4} \cdot 1.125^{0.6} = 1.2622$ *and* $P_{0t}^{JB} = 1.2^{0.4} \cdot 1.4^{0.6} = 1.3163$*, and the global index can be expressed as a simple function of the sector indices:* $1.2622^{0.4444} \cdot 1.31628^{0.5555} = 1.2919$.
This aggregation of the two sector indices yields the same result, as a direct calculation using all four price relatives $P_{0t}^{J} = 1.5^{0.1777} \cdot 1.125^{0.2667} \cdot 1.2^{0.2222} \cdot 1.4^{0.3333} = 1.2919$.
This demonstrates that the formula P_{0t}^{J}*, though not linear, is additivly consistent, which rests on the fact that the weights are related to one another as follows:* $g_1 = 0.1777 = 0.4 \cdot 0.4444$, $g_2 = 0.2667 = 0.6 \cdot 0.4444$, $g_3 = 0.2222 = 0.4 \cdot 0.5555$, *and* $g_4 = 0.3333 = 0.6 \cdot 0.5555$*. A similar argument applies to the quadratic mean index (QM–index). Using individual price relatives we get* $P_{0t}^{QM} = \sqrt{\frac{400}{900} \cdot \left(1.5^2 \cdot \frac{160}{400} + 1.125^2 \cdot \frac{240}{400}\right) + \frac{500}{900} \cdot \left(1.2^2 \cdot \frac{200}{500} + 1.4^2 \cdot \frac{300}{500}\right)} =$

1.307988 *which leads to the same result as the aggregation of the global QM–index, using squared sectoral QM–indices (the terms in brackets), so that we have*

$$P_{0t}^{QM} = \sqrt{\frac{4}{9}\left(P_{0t}^{QM\ A}\right)^2 + \frac{5}{9}\left(P_{0t}^{QM\ B}\right)^2}.$$

*This clearly shows that the QM–index is aggregatively consistent (**A1**) though not additive (**A1***).*

Things are much less convenient in the case of Drobisch's and Fisher's index. The geometric mean of the Laspeyres and the Paasche index, gives us the global Fisher index $P_{0t}^{F} = \sqrt{1.3 \cdot 1.2794} = 1.2896$. *But no simple formula (like a weighted geometric mean) exists, by which the* K *sector indices of Fisher* P_{0t}^{Fk} *(*$k = 1, \ldots, K$*) could be "combined" to arrive at this global index of Fisher. An aggregation can only be performed, according to eq. 2.3.4, with the* K *sector indices of the Laspeyres–type (*P_{0t}^{Lk}*), and of the Paasche–type (*P_{0t}^{Pk}*). The same is true for Drobisch's index (see below).* ◀

As stated above, the example shows that aggregative consistency (**A1**) is not confined to index formulas, that can be expressed as *linear* functions of the price relatives (**A1***). Nor is it necessary that index functions have a *dual interpretation*: in addition to the weighted mean of price relatives interpretation the ratio of expenditures interpretation. Neither P^{J} nor P^{QM} possess such a "ratio of expenditures" interpretation. Both indices are non–linear index functions and yet both functions are aggregatively consistent. But

> Fisher's ideal index, does not meet *any* of these requirements: additivity, dual interpretation and aggregative consistency. It is also a misunderstanding to assume, that if an index can be expressed as means of price relatives, then the index function is also aggregatively consistent.

To show that this is *not* true we best consider the index of Drobisch $P_{0t}^{DR} = \frac{1}{2}\left(P_{0t}^{L} + P_{0t}^{P}\right)$ which is, in terms of individual price relatives ($i = 1, \ldots, n$),

$$P_{0t}^{DR} = \frac{1}{2}\left[\left(\sum_{i=1}^{n} \frac{p_{it}}{p_{i0}} g_i\right) + \left(\sum_{i=1}^{n} \frac{p_{it}}{p_{i0}} w_i\right)\right] = \frac{1}{2}\left(P_{0t}^{L} + P_{0t}^{P}\right) = \sum_{i=1}^{n} \frac{p_{it}}{p_{i0}} \frac{g_i + w_i}{2}.$$

Thus P^{DR} is clearly an arithmetic mean in terms of price relatives with weights $(g_i + w_i)/2$. With n commodities grouped into two sub–indices, such that $j = 1, \ldots, m$ belongs to group A, and $k = m+1, \ldots, n$ to group B respectively. We have

$$P_{0t}^{DR} = \frac{1}{2}\left[\left(\sum_{j}^{m} \frac{p_{jt}}{p_{j0}} g_j^*\right) \sum_{j}^{m} g_j + \left(\sum_{k}^{n} \frac{p_{kt}}{p_{k0}} g_k^*\right) \sum_{k}^{n} g_k + \right.$$
$$\left. \left(\sum_{j}^{m} \frac{p_{jt}}{p_{j0}} w_j^*\right) \sum_{j}^{m} w_j + \left(\sum_{k}^{n} \frac{p_{kt}}{p_{k0}} w_k^*\right) \sum_{k1}^{n} w_k\right],$$

where $g_j^* = g_j / \sum_j g_j$, and g_k^*, w_j^*, and w_k^* are defined correspondingly. We may also write

$$P_{0t}^{DR} = \frac{1}{2}\left[P_{0t}^{LA} g_A + P_{0t}^{LB} g_B + P_{0t}^{PA} w_A + P_{0t}^{PB} w_B\right]$$

with two sub–indices only. We have $g_A + g_B = 1$ and $w_A + w_B = 1$ and by definition $P_{0t}^{DRA} = \frac{1}{2}\left(P_{0t}^{LA} + P_{0t}^{PA}\right)$, and P_{0t}^{DRB} defined analogously, we get

$$P_{0t}^{DR} = P_{0t}^{DRA} + P_{0t}^{DRB} - \frac{1}{2}\left[P_{0t}^{LA} g_B + P_{0t}^{LB} g_A + P_{0t}^{PA} w_B + P_{0t}^{PB} w_A\right]$$

which is in general *not* to be equated to

$$P_{0t}^{DRA}\left(\frac{g_A + w_A}{2}\right) + P_{0t}^{DRB}\left(\frac{g_B + w_B}{2}\right)$$

unless we would have weights $g_A = g_B = w_A = w_B = 1/2$. Hence in general the correct global index:

$$\frac{1}{2}\left[\left(P_{0t}^{LA} g_A + P_{0t}^{PA} w_A\right) + \left(P_{0t}^{LB} g_B + P_{0t}^{PB} w_B\right)\right]$$

will differ from

$$\frac{1}{2}\left[\frac{1}{2}\left(P_{0t}^{LA} + P_{0t}^{PA}\right) + \frac{1}{2}\left(P_{0t}^{LB} + P_{0t}^{PB}\right)\right].$$

In ex. 2.3.1 we have $P_{0t}^{DRA}\left(\frac{g_A + w_A}{2}\right) + P_{0t}^{DRB}\left(\frac{g_B + w_B}{2}\right) = 1.2897023$ whereas $\frac{1}{2}\left(P_{0t}^{L} + P_{0t}^{P}\right)$ yields a (slightly)[104] different result, amounting to 1.289678. In conclusion:

> Aggregative consistency is *not* automatically satisfied, when an index function can be expressed as a weighted mean of relatives.

Whilst additivity (linearity) can be regarded as a more restrictive condition than aggregative consistency the so called "**equality test**" is a weak version of aggregative consistency of the index function P. The function P satisfies the equality test if

$$P_{0t} = f(P_{0t}^1, P_{0t}^2, \dots, P_{0t}^K) = f(\lambda, \lambda, \dots, \lambda) = \lambda. \tag{2.3.5}$$

Given, all sector indices P_{0t}^k $(k = 1, 2, \dots, K)$ equal λ, then the global index should yield $P_{0t} = \lambda$. This seems to be a rather weak condition, and for a long time there was

[104] because the weights g_A and w_A do not differ much from $1/2$ in the example.

not a single index function known, to satisfy aggregative consistency, and yet being unable to meet equality.
But the example of the Vartia–I index[105] P_{0t}^{VI} showed, that equality (violated by P_{0t}^{VI}) is not a mere implication of aggregative consistency (satisfied by P_{0t}^{VI}). By contrast a situation where equality is met but **A1** is violated, is more difficult to understand. It can easily be verified that the equality test and aggregative consistency are *both*

- *satisfied* by linear index functions, such as the Laspeyres and the Paasche index, and
- violated by Fisher's ideal index.

A simple example may be adequate in order to demonstrate that Fisher's index indeed fails the equality test, a fact not easy to understand at first glance (the same is true for the Drobisch–Index).

Example 2.3.2
Consider two commodities and weights $g_1 = 0.6$, thus $g_2 = 1 - g_1 = 0.4$ to calculate P_{0t}^{L} using sectoral indices, and corresponding weights $w_1 = 0.4$, and $w_2 = 0.6$ to calculate P_{0t}^{P}. Then assume $P_{0t}^{L1} = 1.25$ and $P_{0t}^{P1} = 1.2$, and likewise $P_{0t}^{L2} = 2$ and $P_{0t}^{P2} = 0.75$ such that $P_{0t}^{F1} = P_{0t}^{F2} = \sqrt{1.5}$. But the global index P_{0t}^{F} is not *$\sqrt{1.5}$, as required by the equality test but rather $\sqrt{1.55 \cdot 0.93} = \sqrt{1.4415}$, because the global indices are $P_{0t}^{L} = 1.55$ and $P_{0t}^{P} = 0.93$.*
Note that we gain the global indices P_{0t}^{L}, P_{0t}^{P}, and thus also P_{0t}^{F} from the sectoral ones using a weighted arithmetic mean*, whilst the sectoral indices P_{0t}^{F} are gained from the sectoral Laspeyres and Paasche indices by taking an* unweighted geometric mean*. This is the reason why P_{0t}^{F} is unable to meet the equality test.* ◀

d) Structural consistency of volumes (A2) and double deflation

It is now our aim to find the conditions under which structural consistency of volumes (**A2** in fig. 2.3.2) is met. They turn out to be highly restrictive, and a strong case for (direct) Paasche indices. The demonstration is extremely simple, yet apparently (and surprisingly)

- not too well known, and more importantly
- its implication for the choice of a deflator–formula is widely ignored.

Let $V_1, V_2, \ldots, V_K$ denote *values* (aggregates at current prices) referring to *sub*–aggregate 1 through K, and V_T to the *total* (T) aggregate respectively, such that by definition

$$V_1 + V_2 + \cdots + V_K = \sum V_k = V_T \qquad (k = 1, 2, \ldots, K). \tag{2.3.6}$$

[105] Also known as index of Montgomery (Stuvel (1989), Balk (1995)).

Each volume is defined by dividing a value by its corresponding price index (deflator), $P_1, P_2, \ldots, P_K$. To satisfy structural consistency of volumes the following equation has to hold for P_T, the "total deflator"

$$\frac{V_1}{P_1} + \cdots + \frac{V_K}{P_K} = \frac{V_T}{P_T}. \tag{2.3.7}$$

Next consider value shares (or "weights") w_k to describe the fact that the total value V_T is broken down into K subaggregates' values

$$\frac{w_1 V_T}{P_1} + \cdots + \frac{w_K V_T}{P_K} = \frac{V_T}{P_T}, \qquad \text{where } w_k = \frac{V_K}{V_T}, \tag{2.3.7a}$$

and after dividing both sides of the equation by V_T

$$w_1 \frac{1}{P_1} + \cdots + w_K \frac{1}{P_K} = \frac{1}{P_T} \tag{2.3.8}$$

which simply means that P_T has to be a *weighted harmonic mean* of sectoral indices (deflators) with weights being value shares $(p_t q_t / \sum p_t q_t)$ that is $w_k = V_k/V_T$. This reduces the class of admissible indices to a unique index formula: the deflators have got to be Paasche indices. Starting with the lowest level of aggregation, that is individual price relatives up to the highest level of the overall index, a Paasche index always is a harmonic mean of sectoral indices (or relatives) of the preceding level of aggregation. Hence the following *uniqueness theorem* holds and can easily be verified:

> The only deflator price index capable of producing structurally consistent volumes at all levels of aggregation, is the direct Paasche price index.

Only if the corresponding Paasche price indices are introduced for $P_1, P_2, \ldots, P_K$ eq. 2.3.7 will hold. If on the other hand, the price indices $P_1, P_2, \ldots, P_K$ in eq. 2.3.7 are replaced by Fisher indices $P_k^F = \sqrt{P_k^L P_k^P}$ $(k = 1, \ldots, K)$ we will get

$$\frac{V_1}{P_1^P}\sqrt{\frac{P_1^P}{P_1^L}} + \cdots + \frac{V_K}{P_K^P}\sqrt{\frac{P_K^P}{P_K^L}} \neq \frac{V_T}{P_T^P}\sqrt{\frac{P_T^P}{P_T^L}}, \tag{2.3.9}$$

with the result that the left hand side (LHS), will usually differ from the right hand side (RHS). To see this consider $K = 2$. With Paasche price indices we get the following result

$$\frac{V_1}{P_1} + \frac{V_2}{P_2} = \frac{V_T}{P_T} \tag{2.3.8a}$$

or solved for P_T

$$P_T = \frac{V_1 + V_2}{\frac{V_1}{P_1} + \frac{V_2}{P_2}} = \frac{1}{\frac{w_1}{P_1} + \frac{w_2}{P_2}}. \tag{2.3.8b}$$

The LHS of eq. 2.3.9 yields

$$\frac{w_1 P_2^F + w_2 P_1^F}{P_1^F P_2^F} V_T \tag{2.3.9a}$$

by contrast to the RHS:

$$V_T \frac{1}{P_T^F} \tag{2.3.9b}$$

On the other hand

$$\frac{1}{P_T^F} = \sqrt{\frac{P_1^P P_2^P}{w_1 P_2^P + w_2 P_1^P} \frac{1}{g_1 P_1^L + g_2 P_2^L}}, \tag{2.3.9c}$$

where

$$g_k = \frac{\sum_j p_{j0} q_{j0}}{\sum_k \sum_j p_{j0} q_{j0}} = \frac{V_k^B}{\sum_k V_k^B}$$

(g_1 and g_2 are shares of the sub–aggregates referring to *base period* values) which will differ from the corresponding term in eq. 2.3.9a. Because the total Fisher index is not a harmonic mean of sectoral Fisher indices, as required in eq. 2.3.8a (P_{0t}^F ist not additivly consistent) deflation with P_{0t}^F is not *structurally* consistent either.
We now turn to double deflation, where we again encounter a harmonic mean:

> Deflation of Y where Y is a *difference* $Y = O - I$ by deflating O and I separately is called *double* deflation (or *indirect* deflation). An example for the application of this procedure is *value added* (Y), defined as difference between O (output) and I (input).

Let $O(s, k)$ denote output with quantities relating to time s, and prices relating to time k. Input I and value added Y will receive equivalent symbols. Then for values (= at current prices)

$$Y(t, t) = O(t, t) - I(t, t) \tag{2.3.10}$$

holds by definition. Now define value added at constant prices $Y(t, 0)$ as follows:

$$Y(t, 0) = \frac{O(t, t)}{P_{0t}(O)} - \frac{I(t, t)}{P_{0t}(I)} = O(t, 0) - I(t, 0). \tag{2.3.11}$$

Value added at constant prices gained this way is said to be derived by double deflation. Thus the central idea of this method is to require volume (deflated, constant price) aggregates to satisfy the same equations as nominal (current price) aggregates, and to deflate output and input separately using deflators $P(O)$ and $P(I)$ respectively.

The implicit VA–deflator is then given by $P_{0t}^{imp}(Y) = \frac{Y(t,t)}{Y(t,0)}$. Let $P_{0t}^{P}(O)$ be the Paasche price indices of output and correspondingly $P_{0t}^{P}(I)$ the price indices of input, such that $Y(t,0) = \frac{O(t,t)}{P_{0t}^{P}(O)} = \frac{I(t,t)}{P_{0t}^{P}(I)}$. Let i denote the input quota $i = I(t,t)/O(t,t)$, that is the percentage of output received as input from other firms at time t, both prices and quantities referring to period t. The implicit deflator (price index) of value added is then given by

$$P_{0t}^{imp}(Y) = \frac{Y(t,t)}{Y(t,0)} = \frac{P_{0t}^{P}(O) \cdot P_{0t}^{P}(I) \cdot (1-i)}{P_{0t}^{P}(I) - i \cdot P_{0t}^{P}(O)}, \qquad \text{where } i = \frac{I(t,t)}{O(t,t)}. \tag{2.3.12}$$

The following definition of the implicit deflator is perhaps more lucid:

$$\frac{1}{P_{0t}^{P}(O)} = i\frac{1}{P_{0t}^{P}(I)} + (1-i)\frac{1}{P_{0t}^{imp}(Y)} \tag{2.3.12a}$$

Hence the *output*–deflator is a weighted harmonic mean of the *input*–deflator $P_{0t}^{P}(I)$, and the implicit *value added*–deflator $P_{0t}^{imp}(Y)$, with weights i and $(1-i)$ respectively. But unfortunately some awkward results are possible as eqs. 2.3.12 and 2.3.12a reveal

- a negative "real" value added (when the denominator in eq. 2.3.12 is negative)
- $P_{0t}^{imp}(Y) = 0.9 < 1$ although both, output *and* input prices increase $P_{0t}^{P}(I) = 1.4$ and $P_{0t}^{P}(O) = 1.2$ (assume $i = 0.7$), and it is of course also possible that $P_{0t}^{imp}(Y) = 1.56 > 1$ with declining output– *and* input prices ($P_{0t}^{P}(O) = 0.8, P_{0t}^{P}(I) = 0.9$ and $i = 0.7$ as before).

Deflation by substituting the relevant *Fisher* price indices in eq. 2.3.11 we get volumes for which

$$\frac{1}{P_{0t}^{P}(O)} = i\frac{1}{P_{0t}^{P}(I)}R_1 + (1-i)\frac{1}{P_{0t}^{imp(F)}(Y)}R_2,$$

where $R_2 = \sqrt{\frac{P_{0t}^{P}(O)}{P_{0t}^{L}(O)}}$, and $R_1 = R_2\sqrt{\frac{P_{0t}^{P}(I)}{P_{0t}^{L}(I)}}$ hold, an equation for which a straightforward interpretation does not appear to be at hand.

e) Volumes suitable to serve as substitutes for "quantity"

In our view a deflator–index P_{0t}, which is not necessarily the direct Paasche index (P_{0t}^{P}), should meet the following criteria, expressly conceived in order to gain volumes capable of an interpretation in terms of "quantity":[106]

[106] Our considerations are therefore restricted to the volume approach and to CFs, that is to situations in which there is no problem in defining "quantities" of goods "belonging" to the aggregate in question,

1. Typology

To examine under which conditions the result of deflation could be interpreted as a measure of (or substitute for) "quantity"[107] we distinguish three situations concerning the change in prices and quantities in period t relative to period 0:

a) no change (constant) $p_{it} = p_{i0}$ and $q_{it} = q_{i0}$,

b) the prices (and / or quantities) change at the *same* rate ($p_{it} = \lambda p_{i0}$ and / or $q_{it} = \omega q_{i0}$); note that case a) is simply a special case of case b) ($\lambda = \omega = 1$),

c) a change with different rates will take place.

Hence we have the following nine situations (but actually only four [non–shaded] cases):

Prices	Quantities		
	2a) constant	2b) same rate ω	2c) different rates
1a) constant			case 1
1b) same rate λ	ex. 5.2.1		case 2
1c) different rates			case 3

How should "volumes" resulting from deflation reflect such situations? It appears straightforward to require the following properties:

if	then	case	name of criterion
all quantities change uniformly at the same rate (ω), and / or all prices remain constant ($\lambda = 1$)	volume (or a "quantity") index should reflect the *same* rate of change ω.	all shaded cases	Proportionality in prices and quantities [**PPQ**]
all prices change uniformly at the same rate (λ) while quantities change at different rates	volumes should rise / decline as quantities rise / decline	case 2	Reflection of quantity movement [**RQM**]
both, prices and quantities change with different rates	volumes should be linear in quantities*	case 3	Pure quantity comparison [**PQC**]

* or less demanding: volumes should be weighted averages of quantities

and in which "direct" measurement using a quantity index would be possible. We assume that if it were possible to observe quantities directly, and to aggregate them to a meaningful total quantity $\sum q_{i0}$, or $\sum q_{it}$ so that indicators of rise or decline of quantity $M_{0t} = \sum q_{it} / \sum q_{it}$, and of the structure of quantities $q_{i0} / \sum q_{i0}$ and $q_{it} / \sum q_{it}$ can be provided.

107 As already mentioned, it is sometimes said that volume reflects both quantity and quality.

2. Proportionality in prices and in quantities (PPQ)

Whilst the nominal aggregate (value) reflects both, change in quantities as well as in *prices*, a volume (and a quantity index) should be affected by changes in quantities only. Therefore:

1. There should be no problem of deflation, when all *prices* remain constant (row 1a, in particular case 1): under such conditions it is reasonable to require

$$P_{0t} = 1 \Rightarrow \sum p_t q_t = \sum p_0 q_t \\ \Rightarrow Q_{0t} = V_{0t} / P_{0t} = V_{0t} = Q^L_{0t}, \tag{2.3.13}$$

   and, because identity is simply a special case of proportionality in prices, ($\lambda = 1$) there should also be no problem once all prices change *in proportion* (that is $p_{it} = \lambda p_{i0}$ for all $i = 1, \ldots, n$)[108]

$$P_{0t} = \lambda \Rightarrow \sum p_t q_t / \lambda = \sum p_0 q_t \\ \Rightarrow Q_{0t} = V_{0t} / \lambda = Q^L_{0t}. \tag{2.3.13a}$$

2. Provides that the task of deflation is to "isolate" in some way the quantity component of volume in the case of uniform change (at the same rate ω) or constancy ($\omega = 1$) of *quantities* deflation should result in a "volume index" as follows:

$$\frac{V_{0t}}{P_{0t}} = \frac{\omega \frac{\sum p_t q_0}{\sum p_0 q_0}}{P_{0t}} = \frac{\omega P^L_{0t}}{P_{0t}} = \omega, \tag{2.3.14}$$

   and in an "absolute volume"

$$\frac{\sum p_t q_t}{P_{0t}} = \frac{\omega \sum p_t q_0}{P_{0t}} = \frac{\omega P^L_{0t}}{P_{0t}} \sum p_0 q_0 = \omega \sum p_0 q_0, \tag{2.3.15}$$

   whatever the change of prices may be.

It is important to see that all these stipulations might be viewed as plain common sense:

> To ensure the first result (eqs. 2.3.13 and 2.3.13a) it is sufficient that the "**deflator**" price index meets strict **proportionality in prices**. It can be shown that deflation with a *chain* Fisher price index, as opposed to a *direct* Fisher price index, can violate this requirement (see ex. 5.2.1).

[108] We may say that in such a case, there is no reason why the movement of volumes should differ from the movement of quantities. The principle **PQC** (see below) will follow exactly this idea.

Under the conditions of no. 1 (strict proportionality in prices), we get $P_{0t}^F = P_{0t}^P = P_{0t}^L$. Hence eqs. 2.3.13 and 2.3.13a will hold when a direct Fisher price index, or a direct Paasche price index is used as a deflator. Under the conditions of no. 2 it is necessary that $P_{0t} = P_{0t}^L$ in order to obtain a volume (or quantity) index equaling ω, and the absolute volume which is simply the product of ω, and base period volume $\sum p_0 q_0$. To stipulate $P_{0t} = P_{0t}^L$ is tantamount to the *value index preserving test* according to which, a price index should be

$$P_{0t} = P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \omega \mathbf{q}_0) = \frac{V_{0t}}{Q_{0t}} = \frac{V_{0t}}{\omega} = \frac{\sum p_t \omega q_0}{\omega \sum p_0 q_0},$$

or simply $P_{0t} = P_{0t}^L$, which is an implication of *proportionality* (and hence also identity) applied to a *quantity* index Q_{0t}, if Q_{0t} in combination with a price index satisfies the product test. Note that we again get $P_{0t} = P_{0t}^L = P_{0t}^P$ because

$$P_{0t}^F = \sqrt{\frac{\sum p_t q_0 \sum p_t \omega q_0}{\sum p_0 q_0 \sum p_0 \omega q_0}}.$$

Therefore using a *direct* Fisher price index $P_{0t} = P_{0t}^F$, as a deflator will ensure *proportionality in quantities*, not only in the special case of identity ($Q = \omega = 1$), but also in the case of $\omega \neq 1$.

Using a Fisher *chain* price index[109], however, in the case of $t = 2$ for example to deflate V_{02} or $\sum p_2 q_2$ we have

$$P_{02} = \bar{P}_{02}^{FC} = \sqrt{\frac{\sum p_1 q_0 \sum p_1 q_1 \sum p_2 q_1 \sum p_2 q_2}{\sum p_0 q_0 \sum p_0 q_1 \sum p_1 q_1 \sum p_1 q_2}},$$

and even the restrictive assumption $q_2 = q_0$ (*no* quantity change) would give

$$\bar{P}_{02}^{FC} = \sqrt{\frac{\sum p_2 q_0 \sum p_2 q_1}{\sum p_0 q_0 \sum p_0 q_1}} = \sqrt{P_{02}^L \left(P_{01}^P P_{12}^L\right)},$$

which in general will differ from P_{02}^L. In conclusion: Proportionality in quantities (eqs. 2.3.14 and 2.3.15) is equivalent to the value index preserving test (see sec. 2.2e), and is ensured in the case of a *direct* Paasche or Fisher price index as deflator, but not in the case of deflation with a *chain* Fisher price index.

3. Reflection of differential quantity movement (RQM)

(prices change uniformly, case 2)

Whenever quantities change with *different* rates the structure of volumes cannot, and should not remain constant, even though prices may change uniformly. It seems to

[109] In fact international recommendations (U.N. and European Union) prefer the chain version of Fisher's index as a deflator index. This will be discussed in more detail in chapter 5.

be a desirable result of deflation that volumes (global or broken down into subaggregates) rise above (or below) the average, as quantities do so, because there is no differential price movement to disturb this relation between quantities and volumes. The criterion **RQM** means:

> When all prices change at the same rate, the change of volumes should equal the change of quantities. Or in other words: dynamics and structure of volumes should be equal to dynamics and structure of quantities, when prices change uniformly.

To treat case 2 in analogy with case 1 (identity in prices) redounds to (according to eq. 2.3.13a):

- absolute volumes equaling $\sum p_t q_t / \lambda = \sum p_0 q_t$, and
- a quantity index amounting to $Q_{0t} = V_{0t} / \lambda = Q^L_{0t}$.

Hence **RQM** requires, that absolute volumes and a change therein, as measured by a quantity (or "volume") index, should depend on quantities (vectors $\mathbf{q}_0$ and $\mathbf{q}_t$) only, or more distinct: they should be linear combinations of quantities q_t, or quantity relatives q_t / q_0 respectively, with constant weights. Again it does not come as a surprise, that deflation using a *direct* Paasche or a Fisher price index meets **RQM**, whereas deflation with a *chain* Fisher price index does not.

4. Pure quantity comparison (PQC)

(when prices change differently, case 3)

When prices change differently, it might be desirable to keep volumes in close contact with quantities, by requiring, that it is guaranteed:

- (as a minimum) that volumes, or the index Q_{0t} should reflect an *average* of quantity movement (mean value property of Q_{0t}), or
- (more restrictive) that **RQM** still holds in case 3, as it does in case 2.

The criterion mentioned last, to be called **PQC** requires that:

> As (change in) volumes reflect (change in) quantities, in the case of a *uniform* change in prices (eq. 2.3.13a) as they should do in the case of a differential (non–uniform change) movement of prices.

This highly restrictive property of volumes, appears justified for a number of reasons:

1. **PQC** consists in eliminating the effect a changing *structure* of prices has on volumes,

2. **PQC** allows the interpretation of volumes as measures of the aggregated "quantity", reflecting a "*pure*" quantity movement *because* the sums $\sum q_{it}p_{it}$ and $\sum q_{it}p_{i0}$, are different due to different prices, p_{it} and p_{i0} respectively, and each element in the sequence $\sum q_{i0}p_{i0}, \sum q_{i1}p_{i0}, \sum q_{i2}p_{i0}, \ldots$ differs from the other only by the set of quantities.

3. We saw "proportionality in prices and in quantities" (**PPQ**) as plain common sense, this also implied "reflection of **differential** quantity movement" (**RQM**) and now we simply go one step further in requiring **RQM**, not only when prices change uniformly.

4. **PQC** is equivalent to requiring linearity (additivity) in quantities (q_{it}) of the quantity index resulting from deflation (with prices p_{i0} acting as constant weights).

Obviously **PQC** can be met (only) by using P^P_{0t} as a deflator, resulting in Q^L_{0t} as the quantity index. In case of non–uniform change of prices P^F_{0t} will no longer equal P^P_{0t}, and deflation with P^F_{0t} as a deflator results in Q^F_{0t}, not in Q^L_{0t}. According to eq. 2.1.18

$$Q^F_{0t} = Q^L_{0t}\sqrt{1 + r_{ab}V_aV_b} \qquad (2.3.16)$$

holds, and Q^F_{0t} may well differ from Q^L_{0t}, depending on the correlation r_{ab} between price– (a) and quantity relatives (b), and their respective coefficients of variation (V). The implications of this will be shown in the following numerical example.

Example 2.3.3

The following data is taken from ex. 2.3.1 (the first two commodities, q_{2t} slightly modified). We assume alternative prices p_t to show how "Fisher deflation" will differ from "Paasche deflation".

	general assumptions			*alternative prices* p_t		
good	p_0	q_0	q_t	p_{tA}	p_{tB}	p_{tC}
1	10	16	12	15	12	18
2	16	15	17.5	24	28.8	19.2

In variant A we have $V_{0t} = P^P_{0t} = 1.5 \Rightarrow Q^L_{0t} = 1$*. The result* $Q^L_{0t} = 1$ *will be the same in all three variants, because* Q^L_{0t} *is the weighted mean of the two quantity relatives* $12/16 = 3/4$ *and* $17.5/15 = 7/6$*, and weights* $p_0q_0 / \sum p_0q_0$*, such that* $Q^L_{0t} = \frac{3}{4}\cdot 0.4 + \frac{7}{6}\cdot 0.6 = 1$*. According to eq. 2.1.15a and under these conditions (*$Q^L_{0t} = 1$*), for the covariance* C *between price relatives and quantity relatives we get:* $C = V_{0t} - P^L_{0t}Q^L_{0t} = V_{0t} - P^L_{0t}$ *and for the quantity index of Fisher* $Q^F_{0t} = \sqrt{Q^P_{0t}} = \sqrt{1 + C/P^L_{0t}}$*. In all three variants we have* $Q^L_{0t} = 1$*.*

variant	p_{1t}/p_{10}	p_{2t}/p_{20}	C	V_{0t}	P^L_{0t}	Q^P_{0t}	$Q^F_{0t} = \sqrt{Q^P_{0t}}$
A	1.5	1.5	±0	1.5	1.5	1	1
B	1.2	1.8	+ 0.06	1.62	1.56	1.038	1.019
C	1.8	1.2	- 0.055	1.38	1.44	0.962	0.981

Hence under these conditions ($Q^L_{0t} = 1$) we can conclude: $P^L_{0t} > P^P_{0t} \Rightarrow Q^L_{0t} > Q^F_{0t}$ *(negative covariance), and* $P^L_{0t} < P^P_{0t} \Rightarrow Q^L_{0t} < Q^F_{0t}$ *(positive covariance) although the two quantity relatives are the same throughout all three variants of the example. This is also the reason why* Q^L_{0t} *where weights assigned to quantity relatives are constant should be preferred to* Q^F_{0t}*: the same quantity relatives should result in the same quantity index.* ◀

According to the theorem of Bortkiewicz the following relationship between P^F deflated aggregates, and P^P deflated aggregates is given (on all levels of aggregation) by

$$\text{volume(Fisher)} = \text{volume(Paasche)}\sqrt{\frac{V_{0t}}{V_{0t} - C}}.\qquad (2.3.17)$$
[110]

5. Consistent updating of volumes

Another attempt will now be made to justify the principle just stated: the same quantity relatives should result in the same quantity index. Volumes as (constantly weighted) linear combinations of quantities not only appear reasonable in view of the standard interpretation of volumes as substitutes for (aggregate) quantities, otherwise unobservable, but also to ensure consistency when volumes are updated on the basis of quantities.

As a starting point for the following considerations assume

1. some base period values (= volumes) given like V^0_k for a subaggregate k or for the k–th group of commodities ($k = 1, \ldots, K$), and
2. it might be desirable to "update" these aggregates to a current volume (with current quantities) using suitable quantity indices Q_k, such that

$$V^0_1 Q_1 + \cdots + V^0_K Q_K = \left(V^0_1 + \cdots + V^0_K\right) Q_T = Q_T \sum_k V^0_k = Q_T V^0. \qquad (2.3.18)$$

Note that such a task is most relevant in official statistics. An aggregate at *current* prices, as well as at "*constant*" prices is frequently broken down into a number of subaggregates, as for example private consumption (V^0_1) or investment (V^0_2) at period 0 (where quantities *and* prices are relating to 0). Hence we frequently have data like:

$$V^0_1 = \sum p_{1,0} q_{1,0}, \ldots, V^0_K = \sum p_{K,0} q_{K,0}$$

and the overall base period volume

$$V^0 = \sum_k V^0_k = \sum_k \sum p_{k,0} q_{k,0},$$

[110] Hence if the covariance is negative or positive, a volume gained by deflation with a Fisher index tends to fall short of or exceed a volume derived from deflation with a Paasche index.

and we aim at corresponding current period volumes, say V_k^t, to be gained by multiplying V_k^0 by a quantity index in a consistent manner (i.e. according to eq. 2.3.18). This is not only useful but sometimes virtually necessary, as for example when the establishment of an appropriate deflator price index (for "indirect" deflation) is difficult, whereas we may be able to dispose of an independent estimate of quantities (to make a "direct" deflation). According to the SNA recommendations a legitimate alternative to arrive at "deflated" value added (VA) is the *direct* approach, by observing *quantities*.[111]

The only total–aggregate (Q_T) quantity index permitting the type of consistent "updating" of base period (subaggregate) volumes $\sum p_0 q_0$ to current period volumes $\sum p_0 q_t$ as described in eq. 2.3.18, has got to be an arithmetic mean of $Q_1, Q_2, \ldots$ with weights $g_k = V_k^0 / V^0$ where $\sum g_k = 1$. This simply follows from eq. 2.3.18 upon dividing both sides by V^0.

$$g_1 Q_{0t}^{L1} + \cdots + g_K Q_{0t}^{LK} = Q_{0t}^{L}. \tag{2.3.18a}$$

Now assume Paasche quantity indices, then a linear combination of such indices, that is

$$w_1 Q_{0t}^{P1} + \cdots + w_K Q_{0t}^{PK} = Q_{0t}^{P} \tag{2.3.19}$$

exists, but with *variable* weights $w_k = \sum_i p_{k,it} q_{k,i0} / \sum_k \sum_i p_{k,it} q_{k,i0}$. In the case of Fisher quantity indices no such equation exists. We rather have

$$\sum_k w_k g_k Q_{0t}^{Fk} + \sum_{i \neq j} g_i w_j Q_{0t}^{Li} Q_{0t}^{Pj} = Q_{0t}^{F}, \tag{2.3.20}$$

because the Fisher indices are *not* linear.

Another justification of the **PQC** (pure quantity comparison) principle requiring absolute volumes (and a quantity index Q_{0t} respectively) built as **linear** combinations of quantities (or quantity relatives) **with constant weights** (p_0 or $p_0 q_0 / \sum p_0 q_0$ respectively) is given by the need to consistently update K base period volumes to current period volumes using quantity indices, and to ensure consistency between direct and indirect deflation. Hence there are two equivalent principles:

1. **structural consistency of volumes** gained by indirect deflation $\Rightarrow$ deflator price index as *harmonic* mean (Paasche), and
2. (structurally) **consistent updating of volumes** $\Rightarrow$ quantity index as *arithmetic* mean (Laspeyres) displaying a movement of volumes affected by quantities only.

[111] In the case of VA it is of course questionable if this is really an alternative, since VA is a NCF, that is there is no quantity dimension by definition.

Both conditions are *not* met in the case of Fisher indices[112], that is there are neither weights ω_k such that $\sum_k \omega_k Q_{0t}^{Fk} = Q_{0t}^F$ holds for the Fisher quantity indices, nor are there weights ω_k such that

$$\left(\sum_k \omega_k \frac{1}{P_{0t}^{Fk}}\right)^{-1} = P_{0t}^F$$

holds for the Fisher price indices.
On the other hand the pair P_{0t}^P, Q_{0t}^L satisfies exactly these conditions, and both indices are linear, and hence also "aggregatively" consistent (additivity **A1** in fig. 2.3.2) whilst the pair P_{0t}^F and Q_{0t}^F is not even able to pass the (compared with **A1**) weaker equality test.

[112] Note that the SNA doesn't recommend using the *direct* Fisher form for deflation but the *chain* version of these indices, that is chain Fisher price–, and chain Fisher volume indices, which is even worse. This will be discussed in great detail in sec. 5.2. According to the SNA "the preferred measure of year to year movements of real GDP is a Fisher volume index" chained together, and "the preferred measure of year to year inflation for GDP is therefore a Fisher price index".

3 Analysis of properties of chain indices

In general properties of direct indices are easily inferred from the formula used, while those of chain indices are often "rather obscure" (SZULC (1983), p. 539f.). It is therefore useful to be able to dispose of some tools in order to assess the "behavior" of chain indices. In what follows we will aim at presenting some results concerning the properties of chain indices, as well as some tools useful for analyzing chain indices. For the sake of simplicity in this chapter, and in the following ones we will consider short chains of $\bar{P}_{0t}^{LC}$, only with some very few goods (mostly two).

3.1 Traditional interpretations of an index not applicable

In this section an attempt is made to give a chain index an interpretation in terms of a

- ratio of expenditures ("changing cost–of–a–budget"), and of a
- mean of (price–) relatives.

As pointed out in sec. 2.1a it is one of the advantages of the direct Laspeyres index P_{0t}^{L} to be capable of *both* standard interpretations. To the direct Fisher index P_{0t}^{F}, on the other hand, *none* of the two interpretations apply, and the same is true for chain indices (or more exactly: there is a possibility, somewhat far–fetched, however, to give them a meaning in terms of "ratio of expenditures").

a) Chain price index as a ratio of expenditures

In $P_{0t}^{L} = \sum p_t q_0 / \sum p_0 q_0$ it is common practice to call $\sum p_t q_0$ the cost of buying the base period basket q_0 at time t. In contrast with the value index $\sum p_t q_t / \sum p_0 q_0$ the quantities q_{it} in the numerator of P_{0t}^{L} may be interpreted as "reduced" to q_{i0} ($q_{i,t} \rightarrow q_{i,0}$) *individually* (for each i separately) to the *actual* level of period 0 with the help of quantity relatives. The equivalent expression in the case of $\bar{P}_{0t}^{LC}$ would be

$$\bar{P}_{0t}^{LC} = \frac{\sum p_t q_0^{LC}}{\sum p_0 q_0}, \qquad \text{where } q_{i,0}^{LC} = \frac{q_{i,t-1}}{\bar{Q}_{0,t-1}^{PC}} \tag{3.1.1}$$

where quantities $q_{i,0}^{LC}$ in the numerator could be interpreted as a set of quantities $q_{i,t-1}$ transformed, or "reduced" *globally* (or *collectively*) to a *fictitious* level of period 0 by dividing them by $\bar{Q}_{0,t-1}^{PC}$. The *level* is traced back, not the structure. Hence:

> There is an interpretation in terms of changing "cost of a (fixed) budget", somewhat far–fetched, however: The structure (with respect to the quantities of individual commodities $i = 1, \ldots, n$) of the "budget" is determined by quantities $q_{i,t-1}$ and these quantities are "reduced" *globally* to a level q_{i0} at time 0 using a chain–quantity index.

The interpretation is not only, as indicated, somewhat far–fetched, but it should also be noted that:

1. in contrast to the usual interpretation of P^L_{0t}, the quantities in the numerator and in the denominator of $\bar{P}^{LC}_{0t}$ (eq. 3.1.1) are *not* the same. They rather are q^{LC}_{i0} in the numerator, and q_{i0} in the denominator, thus it is *not* a "fixed" budget under two different price regimes, and the *structure* of the quantities q^{LC}_{i0} (determined by the structure $q_{i,t-1}$) will in general differ from the structure of q_{i0};

2. "expenditures" $p_0 q^{LC}_0 = p_0 \frac{q_{t-1}}{Q^{PC}_{0,t-1}}$ do not add up to $\sum p_0 q_0$, and the sum of the expenditure ratios (not shares) is given by

$$\sum \frac{p_0 q^{LC}_0}{\sum p_0 q_0} = \frac{\sum p_0 q_{t-1}}{Q^{PC}_{0,t-1} \sum p_0 q_0} = \frac{Q^L_{0,t-1}}{Q^{PC}_{0,t-1}}. \tag{3.1.2}$$

Hence it is *not* possible to describe a chain (price) index in terms of price relatives ("mean of relatives" interpretation).

This is simply an interpretation of eq. 3.1.1. The "quantities" $q^{LC}_{i,0}$ in the numerator spelled out in detail are

$$q^{LC}_{i,0} = q_{i,t-1} \frac{\sum q_0 p_1 \cdot \dots \cdot \sum q_{t-2} p_{t-1}}{\sum q_1 p_1 \cdot \dots \cdot \sum q_{t-1} p_{t-1}},$$

an expression which seems to be neither particularly meaningful and understandable, nor significantly more up to date than q_{i0}, since in addition to the most recent quantities $q_{i,t-1}$, there are also prices and / or quantities related to past periods $0, \dots, t-1$ entering the formula. In the case of a chain Fisher index the corresponding relation is

$$\bar{P}^{FC}_{0t} = \frac{\sum p_t q^{FC}_0}{\sum p_0 q_0}, \qquad \text{where } q^{FC}_{i0} = \frac{q_{it}}{\bar{Q}^{FC}_{0t}}, \tag{3.1.1a}$$

which is simply a consequence of $V_{0t} = \bar{P}^{FC}_{0t} \bar{Q}^{FC}_{0t}$. In the face of the complexity of q^{FC}_{i0} we may conclude that:

> For the sake of simplicity it appears justified to demonstrate forthwith properties of chain indices predominantly in the case of the most simple chain index design, that is in the case of the Laspeyres chain index.

It is interesting to see what the term q^{FC}_{i0} really means. Take for example $t = 2$, then by dividing the quantities q_{i2} by $\bar{Q}^{FC}_{02}$ yields quantities of the following kind

$$q^{FC}_{i0} = \sqrt{q^2_{i2} \frac{\sum p_0 q_0 \sum p_1 q_0 \sum p_2 q_1}{\sum p_0 q_1 \sum p_1 q_2 \sum p_2 q_2}} \tag{3.1.1b}$$

Table 3.1.1: Quantities underlying the expenditure in the numerator of some indices*

Direct indices		
Laspeyres	$V_{02} \rightarrow P_{02}^{L} = \frac{\sum p_2 q_0}{\sum p_0 q_0}$	$q_{i0} = q_{i2} \frac{q_{i0}}{q_{i2}}$
Fisher	$V_{02} \rightarrow P_{02}^{F} = \frac{\sum p_2 q_0^F}{\sum p_0 q_0}$	$q_{i0}^{F} = q_{i2} \sqrt{\frac{\sum p_0 q_0 \sum p_2 q_0}{\sum p_0 q_2 \sum p_2 q_2}} = \frac{q_{i2}}{Q_{02}^{F}}$
Chain indices		
Laspeyres	$V_{02} \rightarrow \bar{P}_{02}^{LC} = \frac{\sum p_2 q_0^{LC}}{\sum p_0 q_0}$	$q_{i,0}^{LC} = q_{i2} \frac{q_{i1}}{q_{i2}} \frac{\sum q_0 p_1}{\sum q_1 p_1} = \frac{q_{i1}}{Q_{01}^{PC}}$
Fisher	$V_{02} \rightarrow \bar{P}_{02}^{FC} = \frac{\sum p_2 q_0^{FC}}{\sum p_0 q_0}$	$q_{i,0}^{FC} = q_{i2} \sqrt{\frac{\sum p_0 q_0 \sum p_1 q_0 \sum p_2 q_1}{\sum p_2 q_2 \sum p_0 q_1 \sum p_1 q_2}} = \frac{q_{i2}}{Q_{02}^{FC}}$

* Starting point: value index $V_{0t} = \frac{\sum p_t q_t}{\sum p_0 q_0}$, $t = 2$. The table shows the kind of transformations made $q_{i2} \rightarrow q_{i0}$, $q_{i2} \rightarrow q_{i0}^{F}$, etc.

in the numerator of $\bar{P}_{02}^{FC}$. Such arcane "quantities" are multiplied by prices p_{i2}, giving expenditures to be compared with $\sum p_{i0} q_{i0}$. Hence if there is a "ratio of expenditures" interpretation, it is at best a bit far–fetched. The equivalent term in the case of the *direct* Fisher index is given by q_{i0}^{F} in tab. 3.1.1 .
A comparison of the underlying real or fictitious quantities again shows that a chain price index is subject to the three sources of variation as shown in sec. 1.1. It reflects:

1. *differences in prices* only when q_{i0} is also used in the numerator (case 1 of the table), as it is generally used in the denominator in all formulas of tab. 3.1.1,

2. an additional *substitution* effect moving from case 1 to case 2, determining the difference between q_{i0} and q_{i0}^{F}, and finally also

3. the *time path* in cases 3 and 4 respectively, i.e. the influence of prices and quantities in the intermediate period 1 in case of q_{i0}^{LC}, and q_{i0}^{FC} as opposed to $q_{i0} = q_{i0}^{L}$, and q_{i0}^{F} where they are irrelevant.

b) Mean value property violated, no "mean–of–relatives"–interpretation

To see this consider the most simple chain of two links only, and two commodities with weights (expenditure shares) as follows: a and $1-a$ for good 1 and 2 respectively at period 0, and correspondingly b and $1-b$ at period 1.
The direct Laspeyres index is obviously given by

$$P_{02}^{L} = \frac{p_{12}}{p_{10}}a + \frac{p_{22}}{p_{20}}(1-a) = m_1 a + m_2(1-a)$$

where m_1 and m_2 denote price relatives. This should be compared with $\bar{P}_{02}^{LC}$, which is by definition

$$\bar{P}_{02}^{LC} = \left[\frac{p_{11}}{p_{10}}a + \frac{p_{21}}{p_{20}}(1-a)\right] \cdot \left[\frac{p_{12}}{p_{11}}b + \frac{p_{22}}{p_{21}}(1-b)\right],$$

or in terms of price relatives

$$\begin{aligned} \bar{P}_{02}^{LC} &= m_1 a\,[b + g(1-b)] + m_2(1-a)\,[(1-b) + b/g] \\ &= m_1 a f_1 + m_2(1-a) f_2 \end{aligned} \tag{3.1.3}$$

where g denotes a ratio of price relatives,

$$g = \frac{p_{11}p_{22}}{p_{12}p_{21}} = \frac{p_{22}\,/p_{21}}{p_{12}\,/p_{11}}.$$

In sec. 3.2e we are going to present a more general formula of the chain Laspeyres index in terms of price relatives (see eq. 3.2.21 through 3.2.24).
Given $g > 1$ (price 2 rising more than price 1) we get $f_1 > 1$ (hence $f_2 < 1$) which associates a greater weight with commodity 1 (to price relative m_1) in the chain index as compared to the direct index. Conversely as price 1 rises less than price 2 ($g < 1, f_1 < 1$), $\bar{P}_{02}^{LC}$ will give more weight to commodity 2.
Thus we expect $\bar{P}_{02}^{LC} < P_{02}^{L}$ because of giving a relatively lower weight to the commodity, the price of which rises faster in $\bar{P}_{02}^{LC}$ as compared to P_{02}^{L}. Also note that $g = 1$ is tantamount to $\bar{P}_{02}^{LC} = P_{02}^{L}$ irrespective of the size of the weights b and $1-b$ respectively.
From eq. 3.1.3 it also follows, that the chain index is not an average of price relatives since the sum $af_1 + (1-a)f_2$ not necessarily equals unity but amounts to

$$af_1 + (1-a)f_2 = a\,(b + (1-b)g) + (1-a)\left(bg^{-1} + (1-b)\right). \tag{3.1.4}$$

It can be shown that this sum is simply $Q_{01}^{L}\,/Q_{01}^{P}$, or more generally (see eq. 3.2.23 in sec.3.2e) $Q_{0,t-1}^{L}\,/\bar{Q}_{0,t-1}^{P}$. We may also present this result in a matrix notation. P_{02}^{L} is defined as follows

$$P_{02}^{L} = \begin{bmatrix} a & 1-a \end{bmatrix} \cdot \begin{bmatrix} \frac{p_{11}}{p_{10}} & 0 \\ 0 & \frac{p_{21}}{p_{20}} \end{bmatrix} \cdot \begin{bmatrix} \frac{p_{12}}{p_{11}} & 0 \\ 0 & \frac{p_{22}}{p_{21}} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \mathbf{a}'\mathbf{P}_1\mathbf{P}_2\mathbf{1} \tag{3.1.5}$$

where $\mathbf{1}$ is a column vector with elements 1. It can easily be seen that the product of diagonal matrices $\mathbf{P}_1\mathbf{P}_2 = \mathbf{P}_{02}$ is a diagonal matrix with price relatives on the main diagonal. As the inner product $\mathbf{a}'\mathbf{1} = 1$ the sum of the weights in the bilinear form (eq. 3.1.5) is unity. Moreover we get $P^L_{03} = \mathbf{a}'\mathbf{P}_1\mathbf{P}_2\mathbf{P}_3\mathbf{1} = \mathbf{a}'\mathbf{P}_{03}\mathbf{1}$ and so on.
The corresponding equation to eq. 3.1.5 in the case of the Laspeyres *chain* index $\bar{P}^{LC}_{02} = (\mathbf{a}'\mathbf{P}_1\mathbf{1})(\mathbf{b}'\mathbf{P}_2\mathbf{1})$ is more complicated

$$P^{LC}_{02} = \begin{bmatrix} a & 1-a \end{bmatrix} \cdot \begin{bmatrix} \frac{p_{11}}{p_{10}} & 0 \\ 0 & \frac{p_{21}}{p_{20}} \end{bmatrix} \cdot \begin{bmatrix} \frac{p_{12}}{p_{11}} & \frac{p_{22}}{p_{21}} \\ \frac{p_{12}}{p_{11}} & \frac{p_{22}}{p_{21}} \end{bmatrix} \cdot \begin{bmatrix} b \\ 1-b \end{bmatrix} \tag{3.1.6}$$

or

$$P^{LC}_{02} = \begin{bmatrix} a & 1-a \end{bmatrix} \cdot \begin{bmatrix} \frac{p_{12}}{p_{10}} & \frac{p_{12}}{p_{10}}\left(\frac{p_{22}p_{11}}{p_{21}p_{12}}\right) \\ \frac{p_{12}}{p_{10}}\left(\frac{p_{21}p_{12}}{p_{22}p_{11}}\right) & \frac{p_{22}}{p_{20}} \end{bmatrix} \cdot \begin{bmatrix} b \\ 1-b \end{bmatrix} \tag{3.1.6a}$$

We can express this eq. in terms of eq. 3.1.3 and obtain

$$\begin{bmatrix} \frac{p_{11}}{p_{10}} & \frac{p_{12}}{p_{10}}\left(\frac{p_{22}p_{11}}{p_{21}p_{12}}\right) \\ \frac{p_{12}}{p_{10}}\left(\frac{p_{21}p_{12}}{p_{22}p_{11}}\right) & \frac{p_{22}}{p_{20}} \end{bmatrix} = \begin{bmatrix} m_1 & m_1 g \\ m_2 g^{-1} & m_2 \end{bmatrix}, \tag{3.1.7}$$

such that

$$\bar{P}^{LC}_{02} = \begin{bmatrix} a & 1-a \end{bmatrix} \cdot \begin{bmatrix} m_1 & m_1 g \\ m_2 g^{-1} & m_2 \end{bmatrix} \cdot \begin{bmatrix} b \\ 1-b \end{bmatrix} \tag{3.1.3a}$$

which allows verifying and generalizing eq. 3.1.3. We also see that a bilinear form exists such that

$$\bar{P}^{LC}_{02} = \begin{bmatrix} a & 1-a \end{bmatrix} \cdot \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} \cdot \begin{bmatrix} b+(1-b)g \\ bg^{-1}+(1-b) \end{bmatrix} \tag{3.1.8}$$

where

$$\begin{bmatrix} b+(1-b)g \\ bg^{-1}+(1-b) \end{bmatrix} = \begin{bmatrix} 1 & g \\ g^{-1} & 1 \end{bmatrix} \cdot \begin{bmatrix} b \\ 1-b \end{bmatrix} \tag{3.1.8a}$$

From this also follows eq.3.1.4 for the sum of weights, because

$$\begin{bmatrix} a & 1-a \end{bmatrix} \begin{bmatrix} b+(1-b)g \\ bg^{-1}+(1-b) \end{bmatrix} = a\,(b+(1-b)g)+(1-a)\left(bg^{-1}+(1-b)\right), \tag{3.1.4a}$$

and it can easily be seen that upon setting $g = 1$ the sum will amount to 1, otherwise this is not guaranteed. Hence

> A chain index may well exceed the greatest individual price relative or result in a value possibly less than the smallest price relative, and thus violate mean value property. This will also entail some unfavorable aggregation properties as shown in sec. 3.2e .

In sec. 3.2e we will present some more general formulas for $\bar{P}^{LC}_{0t}$ as a weighted sum of price relatives, and for the sum of weights, that is we derive some generalisations from eq. 3.1.8 and eq. 3.1.4 (or 3.1.4a).

Example 3.1.1

Given the following prices and quantities of $n = 2$ commodities ($i = 1, 2$)

	period 0		*period 1*		*period 2*	
i	*prices*	*quantities*	*prices*	*quantities*	*prices*	*quantities*
1	*2*	*10*	*12*	*3*	*12*	
2	*5*	*4*	*7*	*10.29*	*14*	

Quantities at period 2 are irrelevant in calculating $\bar{P}^{LC}_{02}$, as well as P^{L}_{02}. The direct index is obviously given by $P^{L}_{02} = \frac{12}{2}0.5 + \frac{14}{5}0.5 = 4.4$, and this should be compared with the chain index $\bar{P}^{LC}_{02}$, given by

$$\bar{P}^{LC}_{02} = \left(6 \cdot \frac{1}{2} + \frac{7}{5} \cdot \frac{1}{2}\right)\left(1 \cdot \frac{1}{3} + 2 \cdot \frac{2}{3}\right) = 6.167 > 6$$

i.e. exceeding the greatest individual price relative ($m_2 = p_{22}/p_{20} = 6$, $m_1 = 14/5 = 2.8$). In terms of eq. 3.1.3 we have $af_1 = 0.833$ and $(1-a)f_2 = 0.4167$ such that $af_1 + (1-a)f_2 = 1.25 \neq 1$, and $6 \cdot 0.833 + 2.8 \cdot 0.4167 = 6.167$, and $g = 2$. ◀

It is not only a theoretical possibility that chain indices may violate the mean value test, but it has already been shown empirically at least once (SZULC (1983), p. 560). The (often quoted) results of the Canadian Consumer Price Index (a chain index) March 1978 were as follows:

Goods	171.1
Services	171.4
Goods and Services	170.8

In view of such findings[113] HILL (1988), p. 14 gives the following comments:

> "There is little doubt that most users find these kinds of results perverse. More seriously, it may be inconvenient or even impossible to work with chain indices within the framework of an accounting system or economic model in which different aggregates are defined, or constrained, by accounting identities when the indices for these aggregates do not respect the same identities."

[113] Quoted in SZULC (1983) as well as HILL (1988).

It should be remembered that Peter Hill was one of the chief authors of the SNA recommendations. The interesting conclusions which might be drawn from this finding are however:

> A possible violation of the mean value property may have severe consequences:
>
> 1. Aggregation *over time*
> In pursuit of the idea of using the most recent and thus (!) most "representative" *weights* in each link (period), it is by no means sure that the product, that is the chain, is "representative" in the sense of a *mean* of price relatives over the whole time interval (an average in the sense of equaling a *typical* price relative).
> 2. Aggregation *over commodities*
> For this type of aggregation the choice of weights is important. But due to the *cumulative* nature of weights in the chain approach the resulting implicit weights of the chain (as opposed to the links) will in general be such that they do not sum up to unity.

Of course the question arises, of which is more important, to have representative weights in each period in the course of time, or to have a representative result over the whole course. It should be noted that the mean value property refers to the relatives themselves, that is to p_{1t}/p_{10}, p_{2t}/p_{20}, and so on, no matter which weights are assigned to item 1, 2, ... and so on. Hence, in aggregating over *time* weights are not yet coming into play.

Which weights ("old" or "new" ones) are chosen is, however, relevant whenever an aggregation over *commodities* or *sub–indices* has to be made. We may call a structure of prices or volumes more "relevant" when "new" weights are used, but in any case a structure which is able to respect certain accounting identities is wanted. To verify the statement above, that after some time ($t \geq 2$) the implicit weights may not sum up to unity, it is useful to once more look at ex. 3.1.1: the weights developed as follows:

	commodity 1		commodity 2	
from ... to	price	weight	price	weight
$0 \to 1$	12 / 2 = 6	$a = 1/2$	7 / 5 = 1.4	$1 - a = 1/2$
$1 \to 2$	12 / 12 = 1	$b = 1/3$	14 / 7 = 2	$1 - b = 2/3$
"sum": $0 \to 2$	6	0.8333	2.8	0.4167

The interesting point responsible for the awkward result of $\bar{P}_{02}^{LC} > 6$ is the cumulative nature of inherent weights (see sec. 1.3), by which these *actual* weights are not only a function of the weights in the two periods, a and b, but are also determined by differential price movement as measured by g.

No system of weights using single weights (a or b respectively) or multiple ones in the form of a mean of weights, for example (say $(a + b)/2$ or $\sqrt{ab}$) would be capable of producing the counterintuitive result under consideration.

3.2 Axiomatic reasoning applicable with some limitations only

Axioms are a powerful tool for a priori assessment of the relative advantages or disadvantages of index formulas. In this section we will try to show that this instrument unfortunately is not applicable in the case of chain indices, at least not without some modifications. If this instrument is applied (by adequately stipulating redefined axioms), however, the performance of chain indices will appear surprisingly poor. In sec. 3.2a and b as well as in 3.2.e we try to show which aspects should be considered when axioms defined for the *bilateral* comparison (two situations taken in isolation) are to be translated into the multilateral comparison. There are mainly three aspects:

1. the impact of chaining: properties of the chain will in general differ from those of the individual links (sec.3.2a),
2. in successive links the summation will take place over a different (varying) collection of commodities (sec.3.2b), and
3. the properties of "additivity" and "linearity" are so far considered only for a *single* t in P_{0t} taken in isolation (sec.2.2d and 2.3b); they need a redefinition in the case of *successive* indices, like $P_{0t}, P_{0,\ t+1}, \ldots$ (sec.3.2e),

which should be taken into account and which would make it difficult to apply axioms in the case of chain indices.

a) Only the link is an index able to satisfy certain axioms

Many arguments advanced to justify chain indices suffer from the lack of axiomatic tools to evaluate properties of formulas. As a result doubts may arise whether a certain property of a chain formula should be regarded as an "advantage" or rather as a "disadvantage". There are two problems that need to be considered:

- Axioms as introduced and discussed in sec. 2.2 apply to the direct approach only, or at least primarily.
- As already pointed out only the link (relating two adjacent periods) is an index in the sense of a function satisfying or violating certain "axioms".

A link is defined in accordance with the general definition of eq. 0.2.1

$$P_{bt} = P(\mathbf{p}_b, \mathbf{q}_b, \mathbf{p}_t, \mathbf{q}_t) \qquad \mathbb{R}^{4n}_{++} \to \mathbb{R}_{++}$$

where the "base" b in P^C_t is $b = t - 1$.
A chain is something different. It is *not* an index and can violate axioms, despite being "made" up of links that satisfy them all. The shortest chain of *two* links is only for example a function

$$\bar{P}_{02} = \bar{P}_2(\mathbf{p}_0; \mathbf{q}_0, \mathbf{p}_1, \mathbf{q}_1, \mathbf{p}_2, \mathbf{q}_2) \qquad \mathbb{R}^{6n}_{++} \to \mathbb{R}_{++} \tag{3.2.1}$$

or correspondingly in the case of three links we have the function

$$\bar{P}_{03} = \bar{P}_3(\mathbf{p}_0; \mathbf{q}_0, \mathbf{p}_1, \mathbf{q}_1, \mathbf{p}_2, \mathbf{q}_2, \mathbf{p}_3, \mathbf{q}_3) \qquad \mathbb{R}^{8n}_{++} \to \mathbb{R}_{++} \tag{3.2.1a}$$

and so on, and it is not clear what "axioms" should be stipulated in this case.
It will be shown that it is of little or no use trying to prove the superiority of chain indices over direct indices, or to discuss the relative merits of various formulas for links in terms of axioms, even though it is not an unusual exercise[114]. The reason is simply:

> Once links are chained together to form a chain, the axiomatic properties of the links will in general wither away, and will not be imparted on the chain. Inapplicability of axiomatic considerations, however is a severe disadvantage of chain indices, and also one of the reasons for having many numerical "examples", "simulations" etc. in the case of chain indices.

The main use of some fundamental axioms is to allow a clear a priori prediction of results of index calculations, *not* depending on data but only inferred from the index function under some simple, idealized, or "stylised" conditions. When axioms are *not* applicable on the other hand there is always some indefiniteness or being "obscure", as Szulc put it (SZULC (1983)). Therefore studying results of numerical examples, and simulations is a widespread practice to overcome such difficulties in the case of chain indices.[115]
The significance of certain axiomatic properties translated into the chain index framework has always been controversial. But it appears reasonable to redefine axioms for chain indices *by analogy* to the corresponding axioms for the direct indices, at least in the case of some "simple" axioms. Strict proportionality (implying strict identity) means in the usual ("bilateral") case of only two periods being compared for example

$$P(\mathbf{p}_0, \mathbf{q}_0, \lambda\mathbf{p}_0, \mathbf{q}_t) = \lambda, \tag{3.2.2}$$

where $\lambda \in \mathbb{R}$, and $\mathbf{p}_t = \lambda\mathbf{p}_0$, and this may be "relaxed" to weak proportionality, that is

$$P(\mathbf{p}_0, \mathbf{q}_0, \lambda\mathbf{p}_0, \mathbf{q}_0) = \lambda,$$

[114] An example of this is the expertise delivered by the Italian statistician Marco Martini to Eurostat. In this work (MARTINI (1996)) an attempt is made to "prove" mathematically the superiority of chain indices over direct indices with the help of certain axioms. The following text might show that such an undertaking is highly dubious. Interestingly the original issue to be dealt with in the paper Martini 1996 was to find the best frequency for updating the weight base of an index. There was not much to be found as an answer to this question. A question where it is obviously difficult to make use of mathematics or axioms.

[115] There are also many popular prejudices concerning the behavior of chain indices, like for example that chain indices will in general display a much smoother movement than direct indices, or that the chain (Laspeyres) index will not exceed the respective direct version. It has been shown that many such prejudices are not tenable, and that "chain indices can be more biased than their direct counterparts, which is both contrary to popular belief and to the purpose of linking" (SZULC (1983), p. 555).

where $\lambda \in \mathbb{R}$, and $\mathbf{p}_t = \lambda \mathbf{p}_0$, $\mathbf{q}_t = \mathbf{q}_0$. In the case of a chain of two links (or the simplest "multilateral" case) we might require

$$\bar{P}_{02} = \bar{P}_2(\mathbf{p}_0; \mathbf{q}_0, \mathbf{p}_1, \mathbf{q}_1, \lambda\mathbf{p}_0, \mathbf{q}_2) = \lambda \qquad , \mathbf{p}_2 = \lambda \mathbf{p}_0. \tag{3.2.3}$$

But what is *weak* proportionality in this case? Does this mean $\mathbf{q}_0 = \mathbf{q}_1 = \mathbf{q}_2$? And what about the prices in period 1, that is $\mathbf{p}_1$?
It is easy to see that a Laspeyres *link* will always meet strict proportionality because

$$P_t^{LC} = \frac{\sum \lambda p_{t-1} q_{t-1}}{\sum p_{t-1} q_{t-1}} = \lambda \qquad , \text{if } p_{i,t} = \lambda p_{i,t-1} \ \forall i. \tag{3.2.3a}$$

Whereas due to

$$\bar{P}_{02}^{LC} = P_1^{LC} P_2^{LC} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \frac{\sum \lambda p_0 q_1}{\sum p_1 q_1} \neq \lambda \tag{3.2.3b}$$

this is not necessarily true for a Laspeyres *chain*. When for *all* $i = 1, \ldots, n$ commodities prices in $t = 2$ are 50% ($\lambda = 1.5$) higher than in $t = 0$ the chain index $\bar{P}_{02}^{LC}$ does not need to amount to $\lambda = 1.5$ whilst a direct index like P_{02}^{L} or virtually all other formulas, like Paasche, Fisher or so *of course* will result in 1.5. This well known shortcoming (implying also violation of *identity*) of a chain index will be demonstrated in ex. 3.2.1.
We can even go a step further and verify that not only

$$\bar{P}_{0t}^{LC} = P_1^{LC} \ldots P_t^{LC} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \cdots \frac{\sum p_t q_{t-1}}{\sum p_{t-1} q_{t-1}} \neq P_{0t}^{L} \tag{3.2.4}$$

but in general also

$$\frac{\sum p_1 q_0}{\sum p_0 q_0} \cdot \frac{\sum p_2 q_1}{\sum p_1 q_1} \cdots \frac{\sum p_t q_{t-1}}{\sum p_{t-1} q_{t-1}} \neq \frac{\sum p_2 q_0}{\sum p_0 q_0} \cdot \frac{\sum p_4 q_2}{\sum p_2 q_2} \cdots \frac{\sum p_t q_{t-2}}{\sum p_{t-2} q_{t-2}} \tag{3.2.5}$$

which shows that *transitivity* (chainability) is, perhaps surprisingly in the face of the term "chain"–index, not met – and this will be discussed in more detail in sec. 4.1 – because

1. the chain will usually differ from the corresponding direct index (a generally accepted result, discussed under the term "drift" in sec. 3.3), and
2. even for the same interval in time $(0, t)$ the result of chaining is not unique, but depends on how the interval is subdivided into subintervals.

Both features, proportionality and transitivity being violated are consequences of a more general trait of chain indices: they are "*path dependent*" which is just the opposite of transitivity. Hence we should make a distinction between "chaining" on the one hand and "chainability" on the other hand (done so in sec. 4.1). There is no use at all in writing about "transitivity by construction"[116].

[116] In MARTINI (2000) this term is widely used, giving the impression that there is no difference between direct and indirect (gained by linking) comparisons as "transitivity" implies.

With the help of some examples (ex. 3.2.1 continued through ex. 3.2.3) we try to demonstrate what a lack of some axiomatic properties really means, but before doing so it is useful to give some thought to another interesting aspect of "chaining".

b) Changes in the domain of definition, a source of an unknown error

We may ask[117] why it wouldn't be better to compile a direct Paasche index P_{0t}^{P} instead of a chain index. In pursuit of ensuring the "maximum possible representativity " P_{0t}^{P} should be better than for example $\bar{P}_{0t}^{LC}$, or $\bar{P}_{0t}^{PC}$ because P_{0t}^{P} is determined by the most recent weights only, whereas a chain is also affected by weights related to the intermediate periods $t-1, t-2, \dots$ as well. Yet a chain index is often preferred to the direct Paasche index.

By the same token the question also arises why – in some cases – a calculation of the relative $a_{0t}^{i} = p_{it}/p_{i0}$ by linking , that is by

$$\frac{p_{i1}}{p_{i0}} \cdot \frac{p_{i2}}{p_{i1}} \cdots \frac{p_{it}}{p_{i,t-1}} \tag{3.2.6}$$

is preferred to a direct computation. The Swiss Consumer Price Index (as presented by BRACHINGER (1999)) is an interesting example of an index in which such indirect computation is involved in a *number of* steps of a (multistage) compilation of the index.[118] In those cases the indirect method of calculation according to eq. 3.2.6 is usually preferred, to the direct calculation of $a_{0t}^{i} = p_{it}/p_{i0}$ even though both methods are said to be equivalent.

The reason given for preferring one method over the other is the greater "flexibility" of the indirect approach, because a commodity i may no longer exist and has to be replaced by j or so. Under such conditions, however, "equivalence" is no longer given, and the "error" E emerging now depends on the distribution of replacements with respect to their amount, and the time they take place. Consider, for example, that in $t = 2$ commodity i is replaced by j, and in $t = 3$ j is replaced by k. We then obtain

$$\frac{p_{i1}}{p_{i0}} \frac{p_{j2}}{p_{j1}} \frac{p_{k3}}{p_{k2}} = \frac{p_{i1}}{p_{i0}} \frac{p_{i2}e_{j2}}{p_{i1}e_{j1}} \frac{p_{i3}e_{k3}}{p_{i2}e_{k2}} = a_{03}^{i} \frac{e_{j2}e_{k3}}{e_{j1}e_{k2}} = a_{03}^{i} E_{03} \tag{3.2.7}$$

such that the left hand side of this equation no longer equals $a_{03}^{i} = p_{i3}/p_{i0}$, but rather the product $a_{03}^{i}E_{03}$. The "error" E is determined by the kind and the time of replacements made, and

$$E_{0t} = \prod_{\tau=1}^{\tau=t} \frac{e_{\tau}}{e_{\tau-1}}, \tag{3.2.8}$$

[117] see sec. 4.4.

[118] In this index various relatives (specified with respect to type of commodity [that is i], or to area, and type of outlet where the price quotation took place) have to be established and aggregated.

where $e_\tau = 1$ when in τ the price $p_{i\tau}$ of the original commodity i is reported, and in general e_τ is the ratio of the price $p_{k\tau}$ of the substituted commodity k and the original one $p_{i\tau}$.

In essence it amounts to an estimation of E when certain "adjustments" are made (for example for a change in the quality of a product) throughout the interval under consideration. However, this is in general *not* done, and not found necessary in the case of a chain index, which therefore is subject to an unknown E–type error.

It would be interesting to explore the overall effect of E–type errors when a number of such linkings (or chainings according to eq. 3.2.7) is made within a complicated multistage index compilation, as for example in the case of the Swiss index mentioned above.

In the same way it is maintained that a chain index might be more *flexible* or more convenient to be calculated in practice as for example $P^P_{0t} = \frac{\sum_i p_{it} q_{it}}{\sum_i p_{i0} q_{it}}$, because in chain indices no care has to be taken for comparability of prices in t with those in 0, whereas in P^P_{0t} summation in the numerator and denominator has to take place over the same commodities. One of the pretended major "advantages" of chain indices is that this does not need to be the case, hence for example $\bar{P}^{LC}_{03}$ as calculated *in practice* will not only differ from the direct index P^L_{03} due to the "drift" (see sec. 3.3), but also from $\bar{P}^{LC}_{03}$ (as defined according to the right hand side of the following equation) due to the fact that the collection of goods and services $i = 1, \ldots, n$ may be substituted by some other collection (j or k)

$$\frac{\sum_i p_{i1} q_{i0}}{\sum_i p_{i0} q_{i0}} \frac{\sum_j p_{j2} q_{j1}}{\sum_j p_{j1} q_{j1}} \frac{\sum_k p_{k3} q_{k2}}{\sum_k p_{k2} q_{k2}} \neq \frac{\sum_i p_{i1} q_{i0}}{\sum_i p_{i0} q_{i0}} \frac{\sum_i p_{i2} q_{i1}}{\sum_i p_{i1} q_{i1}} \frac{\sum_i p_{i3} q_{i2}}{\sum_i p_{i2} q_{i2}} \tag{3.2.9}$$

Hence it cannot be presupposed (as generally done, and also in what follows when axiomatic properties of chain indices are examined) that the calculated $\bar{P}^{LC}_{03}$ in practice can indeed be equated to $\bar{P}^{LC}_{0t}$. We may conclude:

> Indirect comparisons (by chaining) are *not* equivalent to direct comparisons, contrary to what is assumed in general in the chain index methodology. An indication of this is not only the lack of transitivity (eqs. 3.2.4 and 3.2.5), and the inapplicability of certain axioms (or the violation of adequately modified axioms) but also the existence of E–type errors reflecting the fact that successive links (of individual commodities or aggregated in form of link-indices), do not need to refer to the same domain of definition (selection of commodities).

The perennial updating of weights, allegedly necessary to guarantee "representativity" not only introduces an element of incomparability, but in practice also an error of unknown size as it usually goes hand in hand with a redefinition of the sample, not only for weights assigned to the same items. There is no axiomatic theory in which vectors $\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t$ in the function $P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_t)$ are allowed to differ, not only with respect to numerical values entering the vectors, but also with respect to the kind of elements represented by those vectors.

c) Chain indices (not surprisingly) may violate identity

The idea of *identity* applied to situations 0 and 2 with a situation 1 in between would require $\bar{P}_{02} = 1$ to hold whenever prices in 0 and 2 are equal. In a similar manner *monotonicity* requires $\bar{P}_{02}$ to differ from unity as soon as a single price p_{i2} differs from p_{i0} whereas all other prices remain constant. It can easily be seen that a chain Laspeyres price index violates both axioms (though the direct index P^L_{02} satisfies them of course). The following example shows that also the mean value property no longer holds (as stated in sec. 3.1).

Example 3.2.1

Given the following prices and quantities of $n = 2$ commodities ($i = 1, 2$)

	period 0		*period 1*		*period 2*	
i	*prices*	*quantities*	*prices*	*quantities*	*prices*	*quantities*
1	8	6	6	10	8	
2	12	4	15	5	12	

Quantities in period 2 are irrelevant for the calculation of $\bar{P}^{LC}_{02}$. We will also try different variants concerning prices p_{i2} and quantities q_{i1} (see ex. 3.2.2). The fields involved are set off by shadows. The results are as follows:

Direct index: *$P^L_{02} = 1$ due to identity of prices in 0 and 2, the weights (expenditure shares) in notation of sec. 3.1 are $a = 0.5$ and $b = 4/9$ for commodity 1 in period 0 and 1.*

Chain index: *$P_1{}^{LC} = P^L_{01} = 1$ and $P^{LC}_2 = 1.037$ and therefore $\bar{P}^{LC}_{02} = P^{LC}_1 P^{LC}_2 = 1.037$ indicating a rise in prices, although both price relatives $p_{i2}/p_{i0} = 1$ ($i = 1, 2$) show no change.*

With factors f_1 and f_2 as defined in sec. 3.1.b we get $g = \frac{12/15}{8/6} = 0.6$ giving $f_1 = \frac{7}{9}$ and $f_2 = \frac{35}{27}$ and eq. 3.1.2 reads as follows in this example $\bar{P}^{LC}_{02} = 1 \cdot \frac{1}{2} \cdot \frac{7}{9} + 1 \cdot \frac{1}{2} \cdot \frac{35}{27} = 1.037$ which is simply the sum of the "new weights", af_1 and $(1-a)f_2$. ◀

This example shows that a chain Laspeyres index may well violate identity defined for the interval (0,2) with *two sub*intervals (0,1) and (1,2). In the case of $n > 2$ commodities and more than two subintervals the result will only be more complicated.
Another way of expressing this is that $\bar{P}^{LC}_{02} \neq 1$ results from the Laspeyres formula not being *time reversible*[119]. In view of this shortcoming of the Laspeyres and Paasche chain index the SNA recommends using the chain version of the Fisher–index $\bar{P}^{FC}_{02}$, or the Törnquist–index $\bar{P}^{TC}_{02}$ instead (see sec. 5.1). But there are other difficulties involved in using index formulas of this kind[120]. Obviously as every link meets the time reversal

[119] SNA 93, remark no. 16.46.

[120] One of them is the difficulty to find a good interpretation for values deflated with the help of, for example a chain Törnquist–index. FORSYTH and FOWLER (1981), two prominent proponents of chain indices criticized this index for not being useful in the framework of deflation.

property, the chain will do so as well, *if* the way back corresponds exactly to the way forward, step by step. In the case of an interval (0, 2) this means that not only prices, but also quantities have to return to the original level and structure. This is highly restrictive, however, and the result of a chain may still fall short of unity as quantities in general will not react to prices identically in both ways (the reaction to a rise of prices may differ to the reaction to an equivalent decline). To require a path by which the start situation is regained precisely in the same way back as the forward path, is not only farfetched[121], but also in conflict with the general rationale of chain indices, since time reversibility implies

$$\frac{\sum p_2 q_1}{\sum p_1 q_1} = \frac{\sum p_0 q_1}{\sum p_1 q_1} = \frac{1}{\frac{\sum p_1 q_0}{\sum p_0 q_0}} = \frac{\sum p_0 q_0}{\sum p_1 q_0} \Rightarrow P^L_{12} = P_{12(0)}, \tag{3.2.10}$$

simply meaning that a change in prices $p_1 \to p_2 = p_0$ should result in the same price index irrespective of the weights (q_1 or q_0 respectively) being used. But this is clearly not in line with the idea that an updating of weights is crucial:

> If it is true that weights matter and that constantly updating weights will substantially affect the result it should *not be surprising*, on the other hand, that axioms like identity, monotonicity and time reversibility are *violated*.

It is therefore astonishing that according to the SNA the failing of the time reversal test by $\bar{P}^{LC}_{0t}$ or $\bar{P}^{PC}_{0t}$ gives rise to consider a Fisher or Törnquist index, $\bar{P}^{FC}_{0t}$ or $\bar{P}^{TC}_{0t}$ as the possibly better solution. We encounter the same conflict in international comparisons (see sec. 4.1): requiring "country reversibility" is desirable as we then have only one unique parity between any two countries, A and B respectively, because then $P_{BA} = 1/P_{AB}$. But to expect this to hold is not a matter of course if we also wish that in P_{AB} (or P_{BA} respectively) weights are chosen such that they are most "representative" (or "characteristic") for country A (or B respectively).

Thus it is not surprising that chain indices will *not* behave according to the idea of identity. The result $\bar{P}^{C}_{0t} \neq 1$ may well be compatible with a situation with identical prices in 0 and t. To see the reason for this it is useful to look at eq. 3.1.3 by using *link relatives* (sec. 0.3):

$$\bar{P}^{LC}_{02} = P^L_{02} + (l_{12} - l_{22})\left[(1-a)\cdot b \cdot l_{21} - a \cdot (1-b) \cdot l_{11}\right] \tag{3.2.11}$$

where $l_{it} = p_{it}/p_{i,t-1}$, $(i = 1, 2)$. In the case of ex. 3.2.1 we get

$$\bar{P}^{LC}_{02} = 1 + \left(\frac{8}{6} - \frac{12}{15}\right)\left(\frac{1}{2}\frac{4}{9}\frac{15}{12} - \frac{1}{2}\frac{5}{9}\frac{6}{8}\right) = 1.037.$$

[121] This shows once more – as already discussed in sec.2.2e – that it is not useful to require time reversibility. If this property is defined for a *sequence*, and not only for an interchange of just two adjacent periods, it is particularly questionable why satisfying this condition should be advantageous.

The fact that the price of the two commodities change *differently* from period 1 to period 2 (that is $l_{12} \neq l_{22}$) is responsible for the second term on the right hand side of eq. 3.2.11, and therefore for $\bar{P}^{LC}_{02} \neq 1$ despite *identical* price vectors $\mathbf{p}_0$ and $\mathbf{p}_2$.
There is also a close relationship between link relatives and g (in terms of sec. 3.1.b). We have $g = l_{22}/l_{12}$, again showing that it is the *non zero* variance in price relatives which is responsible for violating some axioms, in the case of chain indices.
Another way of looking at the non–identity problem is to describe conditions for identity. From

$$\bar{P}^{LC}_{02} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \frac{\sum p_0 q_1}{\sum p_1 q_1} = \frac{P^L_{01}}{P^P_{01}} = \frac{Q^L_{01}}{Q^P_{01}} = 1 \tag{3.2.12}$$

it follows that identity of a Laspeyres and a Paasche index should prevail. In general we expect, however, $P^L_{01}/P^P_{01} > 1$ and therefore in case of $\mathbf{p}_2 = \mathbf{p}_0$

$$P^{LC}_1 P^{LC}_2 = \frac{\sum p_1 q_0}{\sum p_0 q_0} \frac{\sum p_0 q_1}{\sum p_1 q_1} > 1 \qquad \text{and} \qquad P^{PC}_1 P^{PC}_2 = \frac{P^P_{01}}{P^L_{01}} < 1.$$

Note that what matters is not the direction of change, whether price or quantity relatives show rise or decline, but the *correlation* between price and quantity relatives (this will be discussed in more detail in sec. 3.4). Some slight variations of ex. 3.2.1 will help to examine other axioms now.

d) Monotonicity and some other axioms

In what follows we try to show that monotonicity may also be violated. For this purpose the numerical ex. 3.2.1 will be modified as follows:

Example 3.2.2
The original ex. 3.2.1 will be modified in the following way:

variant/changes in ex. 3.2.1			*results*		
variant	*quantities in period 1*	*prices in period 2*	P^{LC}_1	$P^{LC}_2 = \bar{P}^{LC}_{02}$	P^L_{02}
(a)	*5 and 10 instead 10 and 5*	*like ex. 3.2.1*	*1*	*0.888*	*1*
(b)	*like variant **a***	*7 and 12 instead of 8 and 12*	*1*	*0.8611*	*0.9375*
(c)	*like ex. 3.2.1*	*8 and 11 instead of 8 and 12*	*1*	*1*	*0.9583* < 1

Note that since $P^{LC}_1 = 1$, *we always get* $\bar{P}^{LC}_{02} = P^{LC}_2$.
Variant (a) *demonstrates that the chain index can also indicate a* decline *in prices* ($\bar{P}^{LC}_{02} = 0.888$), *not only a rise (like in ex. 3.2.1), although all prices in 2 and 0 are identical. In this case the chain index amounts to less than 1, the result of the two identical price relatives. In terms of eq. 3.1.3*

we get $b = 1/6$ *and* $g = 0.6$ *as above in ex. 3.2.1, (because variant* ***a*** *differs only with respect to quantities not to prices), and therefore* $f_1 = 2/3$ *and* $f_2 = 10/9$*. The "new weights" are* $af_1 = 1/3$ *and* $(1-a)f_2 = 5/9$ *(as compared with* $a = (1-a) = 1/2$*) summing up to 0.888.*

Variant (b) *shows that a chain index can also be smaller than the smallest individual price relative* $m_1 = p_{12}/p_{10} = 7/8 = 0.875$*, (whilst* m_2 *remains unity). Hence a chain index can violate the mean value property, not only in the case of* identical *price relatives* $m_1 = m_2 = 1$ *by exceeding this result (1.037 in ex. 3.2.1) or falling short of it (0.888 in variant* ***a****). This can also happen in the case of* $m_1 = 0.875 < m_2 = 1$.

Variant (c) *is designed to show that a chain index can also violate (*weak*)* monotonicity.[122] *We obtain* $\bar{P}_{02}^{LC} = 1$*, although the price of the second commodity is clearly declining (11 instead of 12), whilst the price of the first (8) remains constant and correctly indicates a decline.* ◀

It is interesting to note that not only identity and monotonicity, but also *proportionality* (of which identity is a special case), is not met in the case of chain indices, as indicated above already. This is derived from the fact that though $\mathbf{p}_2 = \lambda \mathbf{p}_0$ the chain $\bar{P}_{02}^{LC} = P_1^{LC} P_2^{LC}$, that is

$$\bar{P}_{02}^{LC} = \lambda \frac{\sum p_1 q_0}{\sum p_0 q_0} \frac{\sum p_0 q_1}{\sum p_1 q_1} \qquad \text{or even} \qquad \bar{P}_{02}^{FC} = \lambda \sqrt{\frac{\sum p_1 q_0}{\sum p_0 q_0} \frac{\sum \lambda p_0 q_2}{\sum p_1 q_2}}$$

will in general *not* result in λ unless some special conditions were met, (for example same quantities at different periods in time which would, however, not accommodate the idea of chaining, see sec. 1.2). Likewise proportionality when *three* links are chained and $\mathbf{p}_3 = \lambda \mathbf{p}_0$ is even more unlikely as

$$\bar{P}_{03}^{LC} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \frac{\sum p_2 q_1}{\sum p_1 q_1} \frac{\sum \lambda p_0 q_2}{\sum p_2 q_2} \qquad p_{i3} = \lambda p_{i0}, \forall i$$

will in general not amount to λ unless quantities meet certain very specific conditions.

Linear homogeneity in prices of period 2 requires $P(\mathbf{p}_0, \lambda \mathbf{p}_2) = \lambda P(\mathbf{p}_0, \mathbf{p}_2)$ and is fulfilled, however, in the case of a chain index because

$$\bar{P}_{02}^{LC}(\lambda) = \frac{\sum p_1 q_0}{\sum p_0 q_0} \frac{\sum \lambda p_2 q_1}{\sum p_1 q_1} = \lambda \frac{\sum p_1 q_0}{\sum p_0 q_0} \frac{\sum p_2 q_1}{\sum p_1 q_1} = \lambda \bar{P}_{02}^{LC}. \tag{3.2.13}$$

Other chain indices, like for example $\bar{P}_{0t}^{PC}$ or $\bar{P}_{0t}^{FC}$ are also able to satisfy this property for any t. Furthermore the following equation holds

$$\frac{\sum \lambda_1 p_1 q_0}{\sum \lambda_0 p_0 q_0} \frac{\sum \lambda_2 p_2 q_1}{\sum \lambda_1 p_1 q_1} = \frac{\lambda_2}{\lambda_0} \frac{\sum p_1 q_0}{\sum p_0 q_0} \frac{\sum p_2 q_1}{\sum p_1 q_1} \tag{3.2.13a}$$

122 Interestingly in the case of chain indices we cannot conclude : if strict monotonicity is met, then also weak monotonicity.

simply because each *link* is linear in the prices. In the same manner an isolated change of prices in the numerator of the last link, will not violate additivity in current period prices

$$\bar{P}^{LC}_{02}(\mathbf{p}_0, \mathbf{p}^*_2) = \frac{\sum p_1 q_0}{\sum p_0 q_0 \sum \lambda_1 p_1 q_1} \sum \left(p_2 + p^+_2\right) q_1$$

$$= \bar{P}^{LC}_{02}(\mathbf{p}_0, \mathbf{p}_2) + \bar{P}^{LC}_{02}(\mathbf{p}_0, \mathbf{p}^+_2). \tag{3.2.14}$$

e) Additivity and aggregative consistency reconsidered

Though eq. 3.2.14 states that "additivity" as defined in accordance with eq. 2.2.15 (sec. 2.2d) is given, we cannot infer favorable aggregation properties, like linearity or (more general) aggregative consistency (sec. 2.3b)[123]. The factor A_2 in

$$\bar{P}^{LC}_{02}(\mathbf{p}_0, \mathbf{p}^*_2) = A_2 \sum \left(p_2 + p^+_2\right) q_1 \quad \text{where } A_2 = \frac{\sum p_1 q_0}{\sum p_0 q_0 \sum p_1 q_1} \tag{3.2.14a}$$

should be examined in more detail, that is

- in the way it develops over time $A_2, A_3, \ldots$ (see sec.3.5 for more details concerning non–linearity), and
- in a breakdown into commodities or sub–indices.

As to the second point[124], take for example two commodities 1 and 2. Then A_2 is given by

$$A_2 = \frac{p_{11} q_{10} + p_{21} q_{20}}{(p_{10} q_{10} + p_{20} q_{20}) \left(\sum p_1 q_1\right)} \tag{3.2.14b}$$

and we obtain

$$\bar{P}^{LC}_{02}(\mathbf{p}_0, \mathbf{p}^*_2) = \frac{p^*_{12}}{p_{10}} (p_{10} q_{11} A_2) + \frac{p^*_{22}}{p_{20}} (p_{20} q_{21} A_2) = \frac{p^*_{12}}{p_{10}} w_1 + \frac{p^*_{22}}{p_{20}} w_2 \tag{3.2.15}$$

and by the same token

$$\bar{P}^{LC}_{02}(\mathbf{p}_0, \mathbf{p}_2) = \frac{p_{12}}{p_{10}} w_1 + \frac{p_{22}}{p_{20}} w_2 \quad \text{and} \quad \bar{P}^{LC}_{02}(\mathbf{p}_0, \mathbf{p}^+_2) = \frac{p^+_{12}}{p_{10}} w_1 + \frac{p^+_{22}}{p_{20}} w_2,$$

such that we indeed have

$$\bar{P}^{LC}_{02}(\mathbf{p}_0, \mathbf{p}^*_2) = \bar{P}^{LC}_{02}(\mathbf{p}_0, \mathbf{p}_2) + \bar{P}^{LC}_{02}(\mathbf{p}_0, \mathbf{p}^+_2).$$

[123] See definition of properties $A1^*$ and $A1$ in fig. 2.3.2 in sec. 2.3b.

[124] We turn to the first point later.

But it is noteworthy that weights w_1 and w_2, do not add up to 1, but to

$$A_2 \sum p_0 q_1 = \frac{\sum p_0 q_1 \sum p_1 q_0}{\sum p_0 q_0 \sum p_1 q_1} = \frac{Q^L_{01}}{Q^P_{01}} = \frac{P^L_{01}}{P^P_{01}}$$

instead[125], nor are they independent of prices p_0.
Expressing $\bar{P}^{LC}_{02}$ by analogy with eq. 2.2.20 gives

$$\bar{P}^{LC}_{02} = \frac{\mathbf{q}_1' \mathbf{p}_2}{\mathbf{b}_2' \mathbf{p}_0}, \text{ where } \mathbf{b}_2' = \begin{bmatrix} \frac{q_{10}}{\sum p_1 q_1 / \sum p_1 q_0} & \frac{q_{20}}{\sum p_1 q_1 / \sum p_1 q_0} \end{bmatrix} = \frac{1}{B_2} \mathbf{q}_0' \quad (3.2.16)$$

where the vectors are defined as usual

$$\mathbf{q}_1' = \begin{bmatrix} q_{11} & q_{21} \end{bmatrix}, \quad \mathbf{p}_0 = \begin{bmatrix} p_{10} \\ p_{20} \end{bmatrix}, \quad \text{and } \mathbf{p}_2 = \begin{bmatrix} p_{12} \\ p_{22} \end{bmatrix}.$$

It is interesting to note that presupposing $q_{11} = q_{10}$ and $q_{21} = q_{20}$, we get $B_2 = 1$ and hence $\bar{P}^{LC}_{02} = P^L_{02}$. This shows that $B_2 = \sum p_1 q_1 / \sum p_1 q_0 = Q^P_{01} \neq 1$ is responsible for destroying some valuable aggregation properties of P^L_{02}, but this does not mean that B_2 is necessarily also responsible for the difference $\bar{P}^{LC}_{02} \neq P^L_{02}$.

Example 3.2.3

Numerical values assumed here are broadly the same as those in ex. 3.2.1. Consider alternative quantities q_1, such that we get different values for B_2. The example tries to show how different assumptions with respect to quantities q_1 will affect the result.

general assumptions					*alternatives for q_1*			
good	p_0	p_1	p_2	q_0	A	B	C	D
1	*8*	*6*	*8*	*6*	*8*	*6*	*8.5*	*5*
2	*12*	*15*	*12*	*4*	*4*	*3*	*3*	*6*

Note that for both commodities ($i = 1, 2$) we have $p_{i2} = p_{i0}$, and thus $P^L_{02} = 1$ whatever the quantities in 1 may be, and (as in ex. 3.2.1) we have $P^{LC}_1 = P^L_{01} = 1$ and thus $P^{LC}_2 = \bar{P}^{LC}_{02}$.

alternative	$\sum p_1 q_1$	B_2 [a]	$\sum p_2 q_1$	$\bar{P}^{LC}_{02}$ [b]
	(1)	*(2)*	*(3)*	*(3)/(1)*
A	*108*	*1.125*	*112*	*1.037*
B	*81*	*0.844*	*84*	*1.037*
C	*96*	*1*	*104*	*1.083*
D	*120*	*1.250*	*112*	*0.933*

[a] $B_2 = \sum p_1 q_1 / \sum p_1 q_0 = \sum p_1 q_1 / 96$
[b] $\bar{P}^{LC}_{02} = \sum p_2 q_1 / \sum p_1 q_1$

[125] It is also noteworthy that in case of $B_2 = 1$ (considered later) the sum of weights will be 1.

Obviously if $q_{11} = q_{10}$ *and* $q_{21} = q_{20}$ *we not only have* $B_2 = 1$ *but also* $P_2^{LC} = 1 \big/ P_1^{LC}$*, which would thus yield* $\bar{P}_{02}^{LC} = 1$*. Hence equality of quantities in 1 and in 0 respectively would remove both defects simultaneously, the lack of aggregative consistency and the violation of identity.*◀

Note that in the case of $B_2 = 1$ the factor A_2 reduces to $A_2 = \left(\sum p_0 q_0\right)^{-1}$ and as $\mathbf{q}_1 = \mathbf{q}_0$ we get

$$\begin{aligned}\bar{P}_{02}^{LC}(\mathbf{p}_0, \mathbf{p}_2^*) &= A_2 \sum \left(p_2 + p_2^+\right) q_1 \\ &= \sum \left(p_2 + p_2^+\right) q_0 \Big/ \sum p_0 q_0 = P_{02}^{L}(\mathbf{p}_0, \mathbf{p}_2^*)\end{aligned} \tag{3.2.14c}$$

under such conditions.
We now turn to a *series* of chain indices, that is we consider the relationship between $\bar{P}_{03}^{LC}$ and $\bar{P}_{02}^{LC}$, and between $\bar{P}_{04}^{LC}$ and $\bar{P}_{03}^{LC}$ and so on as well as the relationship between the chain indices and (the corresponding) direct indices. By analogy with eq. 3.2.14a we get

$$\bar{P}_{03}^{LC}(\mathbf{p}_0, \mathbf{p}_3^*) = A_3 \sum \left(p_3 + p_3^+\right) q_2 \quad \text{where } A_3 = A_2 \frac{\sum p_2 q_1}{\sum p_2 q_2}, \tag{3.2.17}$$

and

$$\bar{P}_{04}^{LC}(\mathbf{p}_0, \mathbf{p}_4^*) = A_4 \sum \left(p_4 + p_4^+\right) q_3 \quad \text{where } A_4 = A_3 \frac{\sum p_3 q_2}{\sum p_3 q_3} \tag{3.2.17a}$$

and so on. Furthermore by analogy with eq. 3.2.16 we obtain

$$\bar{P}_{03}^{LC} = \frac{\mathbf{q}_2' \mathbf{p}_3}{\mathbf{b}_3' \mathbf{p}_0}, \text{ where } \mathbf{b}_3' = \frac{1}{B_3} \begin{bmatrix} q_{10} & q_{20} \end{bmatrix} = \frac{1}{B_3} \mathbf{q}_0' \tag{3.2.18}$$

and

$$B_3 = B_2 \frac{\sum p_2 q_2}{\sum p_2 q_1} = B_2 Q_2^{PC} = Q_1^{PC} Q_2^{PC} = \bar{Q}_{02}^{PC}. \tag{3.2.19}$$

It is easy to see how the series continues and therefore how $\bar{P}_{04}^{LC}$, $\mathbf{b}_4'$, B_4 etc. are defined (see tab. 3.2.1). There is a *marked* difference between this series and the series of successive direct Laspeyres indices as well as direct Paasche indices.
The two direct indices $P_{0t} = \mathbf{q}_k' \mathbf{p}_t \Big/ \mathbf{q}_k' \mathbf{p}_0$ under consideration ($k = 0 \rightarrow P_{0t}^L$, and $k = t \rightarrow P_{0t}^P$) make use of the *same* quantities in the numerator and in the denominator, and thus they can be expressed as weighted means of price relatives with weights $p_{i0} q_{ik} \Big/ \mathbf{q}_k' \mathbf{p}_0$ adding up to 1 and independent of weights in periods $1, 2, \ldots, t-1$.
However the situation changes markedly when chain indices are considered.

Table 3.2.1: Some indices expressed as ratios of expenditures

t	chain Laspeyres	direct Laspeyres	direct Paasche
2	$\bar{P}_{02}^{LC} = \frac{\mathbf{q}_1'\mathbf{p}_2}{\mathbf{b}_2'\mathbf{p}_0}, \mathbf{b}_2 = \left(\bar{Q}_{01}^{PC}\right)^{-1} \mathbf{q}_0$	$P_{02}^{L} = \frac{\mathbf{q}_0'\mathbf{p}_2}{\mathbf{q}_0'\mathbf{p}_0}$	$P_{02}^{P} = \frac{\mathbf{q}_2'\mathbf{p}_2}{\mathbf{q}_2'\mathbf{p}_0}$
3	$\bar{P}_{03}^{LC} = \frac{\mathbf{q}_2'\mathbf{p}_3}{\mathbf{b}_3'\mathbf{p}_0}, \mathbf{b}_3 = \left(\bar{Q}_{02}^{PC}\right)^{-1} \mathbf{q}_0$	$P_{03}^{L} = \frac{\mathbf{q}_0'\mathbf{p}_3}{\mathbf{q}_0'\mathbf{p}_0}$	$P_{03}^{P} = \frac{\mathbf{q}_3'\mathbf{p}_3}{\mathbf{q}_3'\mathbf{p}_0}$
4	$\bar{P}_{04}^{LC} = \frac{\mathbf{q}_3'\mathbf{p}_4}{\mathbf{b}_4'\mathbf{p}_0}, \mathbf{b}_4 = \left(\bar{Q}_{03}^{PC}\right)^{-1} \mathbf{q}_0$	$P_{04}^{L} = \frac{\mathbf{q}_0'\mathbf{p}_4}{\mathbf{q}_0'\mathbf{p}_0}$	$P_{04}^{P} = \frac{\mathbf{q}_4'\mathbf{p}_4}{\mathbf{q}_4'\mathbf{p}_0}$

As soon as a given pair of periods, (0, t) is no longer considered in isolation, the notion of "additivity" of the index function as defined in sec. 2.2d and 2.3b has to be modified: in the case of chaining, "weights" needed for aggregation of price relatives or sub–indices will change in the course of time (as t varies), and they will *not* sum up to unity. To aggregate Laspeyres chain price indices over commodities (sub–indices) we need a chain Paasche quantity index broken down appropriately by commodities or sub–indices.

To show this we examine the vector **b** determining the denominator in successive values of the chain index according to tab. 3.2.1. We have

$$\mathbf{b}_t' = \frac{1}{B_t}\left[q_{10} \quad q_{20}\right] = \frac{1}{Q_{0,t-1}^{PC}}\mathbf{q}_0' \tag{3.2.20}$$

and in the weighted–mean–of–relatives expression of the index formula we have

$$\bar{P}_{0t}^{LC} = \frac{\mathbf{a}'\mathbf{p}_t}{\mathbf{b}'\mathbf{p}_0} = \frac{\mathbf{q}_{t-1}'\mathbf{p}_t}{\mathbf{b}_t'\mathbf{p}_0} = \frac{1}{Q_{0,t-1}^{PC}}\sum \frac{p_{it}}{p_{i0}}\frac{p_{i0}q_{i,t-1}}{\mathbf{q}_0'\mathbf{p}_0} \tag{3.2.21}$$

and thus

$$\bar{P}_{0t}^{LC} = \sum \frac{\mathbf{p}_{it}}{\mathbf{p}_{i0}}\omega_{it} \tag{3.2.22}$$

with weights

$$\omega_{it} = \frac{p_{i0}q_{i0}^{*}}{\mathbf{q}_0'\mathbf{p}_0} = \frac{p_{i0}\left(q_{i,t-1}/\bar{Q}_{0,t-1}^{PC}\right)}{\mathbf{q}_0'\mathbf{p}_0},$$

as already mentioned in sec. 3.1a. The weights sum up to

$$\sum \omega_{it} = \frac{1}{Q_{0,t-1}^{PC}}\sum \frac{p_{i0}q_{i,t-1}}{\sum p_{i0}q_{i0}} = \frac{1}{Q_{0,t-1}^{PC}}\frac{\mathbf{q}_{t-1}'\mathbf{p}_0}{\mathbf{q}_0'\mathbf{p}_0} = \frac{Q_{0,t-1}^{L}}{Q_{0,t-1}^{PC}} \tag{3.2.23}$$

The results of eq. 3.2.21 and 3.2.22 not surprisingly take into account that

$$V_{0,t-1} = \bar{P}^{LC}_{0,t-1}\bar{Q}^{PC}_{0,t-1}$$

and

$$\bar{P}^{LC}_{0t} = \bar{P}^{LC}_{0,t-1}P^{LC}_{t} = \bar{P}^{LC}_{0,t-1}\frac{\sum p_t q_{t-1}}{\sum p_{t-1}q_{t-1}} = \bar{P}^{LC}_{0,t-1}\frac{\sum p_t q_{t-1}}{V_{0,t-1}\sum p_0 q_0}$$

We now try to express the result of sec. 3.1b, that is eq. 3.1.3 (or equivalently eq. 3.1.8) in terms of the equations derived in this section. From eq. 3.2.22 we get

$$\bar{P}^{LC}_{02} = \frac{1}{\bar{Q}^{PC}_{01}}\sum\frac{p_2 q_1}{\sum p_0 q_0},$$

and it is easy to verify that this equation indeed holds as

$$\left(\bar{Q}^{PC}_{01}\right)^{-1} = \frac{P^{L}_{01}}{V_{01}} = \frac{\sum p_1 q_0}{\sum p_1 q_1} = b\frac{q_{10}}{q_{11}} + (1-b)\frac{q_{20}}{q_{21}}$$

and

$$\sum\frac{p_2 q_1}{\sum p_0 q_0} = P^{LC}_{2}V_{01}.$$

Furthermore using the notations of sec. 3.1 we get

$$\bar{P}^{LC}_{02} = \left(\bar{Q}^{PC}_{01}\right)^{-1}\frac{\sum p_2 q_1}{\sum p_0 q_0} = \left(b\frac{q_{10}}{q_{11}} + (1-b)\frac{q_{20}}{q_{21}}\right)\left(m_1 a\frac{q_{11}}{q_{10}} + m_2(1-a)\frac{q_{21}}{q_{20}}\right),$$

and hence

$$\bar{P}^{LC}_{02} = m_1 ab + m_2(1-a)(1-b) + m_1 a(1-b)\frac{q_{20}q_{11}}{q_{10}q_{21}} + m_2(1-a)b\frac{q_{10}q_{21}}{q_{11}q_{20}}. \quad (3.2.24)$$

It is not self–evident that this is indeed equivalent to eq. 3.1.3. But it turns out that

$$m_1 a(1-b)\frac{q_{20}q_{11}}{q_{10}q_{21}} = m_2(1-a)b\frac{p_{12}p_{21}}{p_{11}p_{22}} = m_2(1-a)\frac{b}{g}$$

and

$$m_2(1-a)b\frac{q_{10}q_{21}}{q_{11}q_{20}} = m_1 a(1-b)\frac{p_{11}p_{22}}{p_{12}p_{21}} = m_1 a(1-b)g$$

f) Cyclical movement of prices

The purpose of the following example (ex. 3.2.4), is to show that in the case of prices changing regularly, such that for example $P^{L}_{02} = P^{L}_{04} = P^{L}_{06} = \cdots = C$, where possibly $C = 1$ the corresponding series of Laspeyres chain indices $\bar{P}^{LC}_{02}, \bar{P}^{LC}_{04}, \bar{P}^{LC}_{06}$ does not need to display a similar cycle. The movement of chain indices may rather show a cycle with increasing or decreasing amplitude and this effect is owed

- partly to the operation of chaining, and thus may also occur in the case of an unweighted chain index, as for example the chain index formula of Carli, also known as Sauerbeck chain index, and partly also
- to quantity weights[126].

A similar exercise using the Sauerbeck chain index can be found in SZULC (1983), p.535f. In the practice of price statistics, however, there is no point in considering an unweighted chain index, because chain indices are motivated by being able to adjust weights rapidly. In the absence of weights, there can be no such advantage from chaining. Our aim here is simply to demonstrate, that it is sometimes[127], the principle of chaining itself, not the existence of weights, which causes the violation of the identity axiom and other axioms.

Example 3.2.4

Again consider two commodities. In 1 the price of the first commodity will be twice the price in 0 such that $p_{11}/p_{10} = 2$*. In 2 the price* p_{12} *is half of* p_{11}*, or in other words, it returned to the original level* $p_{12}/p_{11} = 0.5$ *or* $p_{12} = p_{10}$*. The opposite process will be assumed for the other (second) commodity such that* $p_{21}/p_{20} = 0.5$ *and* $p_{22}/p_{21} = 2$*. The Carli chain index* $\bar{P}^{CC}_{02}$ *now is simply* $\bar{P}^{CC}_{02} = \frac{1}{2}(2 + 0.5) \cdot \frac{1}{2}(0.5 + 2) = 1.56 > 1$.

Assume the process will carry on and on in the same way such that we get for example the following prices at periods $0, 2, 4$ *and so on:*

commodity	*0*	*1*	*2 (=0)*	*3 (=1)*	*4 (=0)*
no. 1	*6*	*12*	*6*	*12*	*6*
no. 2	*8*	*4*	*8*	*4*	*8*

Obviously the Carli chain index will become greater and greater as such a cycle of two periods will be repeated again and again (because $\bar{P}^{CC}_{02} > 1$*):*

$$\bar{P}^{CC}_{04} = (1.56)^2 = 2.44 \quad \bar{P}^{CC}_{06} = (1.56)^3 = 3.81$$

and so on. Conversely in the case of $\bar{P}^{CC}_{02} < 1$ *the Carli chain index will become smaller and smaller. When weights are introduced the situation will not necessarily change. Assume weights* $a = (1 - a) = 0.5$ *in period 0 and* $b = 0.25$ *or* $(1 - b) = 0.75$ *in period 1, respectively for example. This yields a base weighted chain index (hence* $\bar{P}^{LC}_{02}$ *instead of* $\bar{P}^{CC}_{02}$*):*

$$\bar{P}^{LC}_{02} = \left(2 \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2}\right)\left(\frac{1}{2} \cdot \frac{1}{4} + 2 \cdot \frac{3}{4}\right) = 2.03,$$

126 The choice of weights may aggravate as well as mitigate, the change of the amplitude with each successive round of the cycle.

127 Note that the following effects are also due to Carli's index being an unweighted *arithmetic* mean. An unweighted *harmonic* mean would result in a chain over two periods amounting to 1/1.25 = 0.8 and the effect would *not* occur in case of an unweighted *geometric* mean of price relatives (Jevons' index as chain index).

considerably exceeding both individual price relatives (both being unity). It is of course easy to construct an example in which the index will be less than unity, the result we get for the relatives. For example assume weights 1/4 and 3/4 in period 0 and 3/4 and 1/4 in period 1 respectively. We then obtain $\bar{P}_{02} = (7/8)^2 = 0.766$, *and accordingly* $\bar{P}_{04} = (7/8)^4 = 0.586$ *and so on, hence an oscillation tending to zero.* ◀

This example is a simple demonstration of some well known results of chaining:

> It is an undisputed fact that chain indices in the case of cyclical movement of prices, such that prices in $k, 2k, 3k, \ldots$ resembling those in 0, tend to move cyclically with growing amplitude, rising progressively beyond limits as the chain over k periods is > 1 (or tend to zero as the chain amounts to < 1).

The reason for this intolerable result is simply derived from the fact that provided

$$\bar{P}^C_{0,k} = \bar{P}^C_{k,2k} = \bar{P}^C_{2k,3k} = \cdots = C \quad \text{then} \quad \bar{P}^C_{0,2k} = C^2, \bar{P}^C_{0,3k} = C^3, \ldots$$

and all depends on whether $C > 1$ or $C < 1$. The possibility of getting such unacceptable results has also been taken into account by the SNA, which says (para 16.47 – 49) that a chain index should

- *not* be used when prices are *cyclically* moving (rising and declining, and thereafter returning to a certain level in some regular manner) in contrast to
- a (moderate) *monotonous* rise or decline of prices, in which case a chain index is recommended.

Interestingly the SNA 93 arrives at the following rule (paragraphs 16.47 and 48)

	when the relative prices in the first (0) and last period (t)	a chain index should
1)	are very different from each other and chaining involves linking periods, like t^*, in which prices and quantities are intermediate between those of 0 and t	be used (indirect comparison via t^* is recommended)
2)	are similar to each other (and very different to an intermediate period t^*); example: seasonal variation	not be used (no indirect comparison via t^*)

This amounts to requiring a more or less smooth development in which direct indices are not so much objectionable (by contrast to abrupt and substantial changes).[128]
But in summary the SNA concludes:

[128] It will be shown in chapter 6 that this specification of situations means that a chain index is good in situations in which direct indices, especially the Laspeyres index is less objectionable, and it fails in situations in which direct indices fail.

> "On balance, situations favorable to the use of chain Laspeyres and Paasche indices over time seem more likely than those that are unfavorable" (SNA para 16.49).

Thus the SNA, despite being well aware of this shortcoming, is not calling its general preference for chain indices into question.[129]
On the other hand, we should see, however, that cyclical movement of prices – a situation the SNA is most mindful of – not necessarily entails also cycles (possibly even with a constantly increasing or declining amplitude) in *chain* indices as opposed to direct indices. It is not im-possible that in case of cycles in the prices

- a *chain* index will remain more or less constant, whereas
- the corresponding direct index reflects cycles,

that provided quantities will react (or be adjusted) appropriately too. For this to happen we may for example think of, an admittedly rather artificial reaction of quantities, in response to changes in prices as described in ex. 3.2.5.

Example 3.2.5

Consider two commodities and a reaction of households such that as the price of i ($i = 1, 2$) rises (declines) in t by $\lambda_{it} = p_{i,t+1}/p_{i,t}$ the quantity of i will be reduced (increased) according to $q_{i,t+1} = q_{i,t}/\lambda_{it}$, $t = 0, 1, \ldots$. As a consequence for each commodity i and each t the expenditure $p_{it}q_{it}$, and thus also the sum $\sum_i p_{it}q_{it}$ as well as the structure described by expenditure shares $g_i = p_{it}q_{it} \big/ \sum_i p_{it}q_{it}$ remain constant. Hence we obtain

$$P_1^{LC} = P_{01}^L = \frac{p_{10}\lambda_{10}q_{10} + p_{20}\lambda_{20}q_{20}}{p_{10}q_{10} + p_{20}q_{20}} = \lambda_{10}g_1 + \lambda_{20}g_2,$$

$$P_2^{LC} = \frac{p_{10}\lambda_{10}\lambda_{11}\frac{q_{10}}{\lambda_{10}} + p_{20}\lambda_{20}\lambda_{21}\frac{q_{20}}{\lambda_{20}}}{p_{10}\lambda_{10}\frac{q_{10}}{\lambda_{10}} + p_{20}\lambda_{20}\frac{q_{20}}{\lambda_{20}}} = \lambda_{11}g_1 + \lambda_{21}g_2,$$

$$P_3^{LC} = \lambda_{12}g_1 + \lambda_{22}g_2$$

and so on, whereas the direct index develops as follows

$$P_{02}^L = \lambda_{10}\lambda_{11}g_1 + \lambda_{20}\lambda_{21}g_2$$

and

$$P_{03}^L = \lambda_{10}\lambda_{11}\lambda_{12}g_1 + \lambda_{20}\lambda_{21}\lambda_{22}g_2$$

and so on. For the sake of simplicity we assume $g_1 = g_2 = 1/2$. Consider the following λ–values (resulting in a 4–period cycle in P_{0t}^L, and in a constant chain index $\bar{P}_{0t}^{LC} = 1$ for all t):

[129] The question is, however, why should such allegedly, most likely *favorable* situations, be exactly those situations which are *unfavorable* for direct indices of the type Laspeyres (and to a smaller degree also Paasche)? We ask this question because the enthusiasm for chain indices to a great deal, results from a criticism of the constant basket in direct Laspeyres indices.

$0 \rightarrow 1$		$1 \rightarrow 2$	
$\lambda_{10} = 1.2$	$\lambda_{20} = 0.8$	$\lambda_{11} = 1.1$	$\lambda_{21} = 0.9$

$2 \rightarrow 3$		$3 \rightarrow 4$	
$\lambda_{12} = 0.9$	$\lambda_{22} = 1.1$	$\lambda_{13} = 0.8$	$\lambda_{23} = 1.2$

We then have the following results with respect to the direct index (while the chain index remains constant):

t	P^L_{0t}	$P^L_{0t}/P^L_{0,t-1}$
1	1	1
2	1.02	1.02
3	0.99	0.9706
4	0.9504	0.96

t	P^L_{0t}	$P^L_{0t}/P^L_{0,t-1}$
5	0.9504	1
6	0.9641	1.02
7	0.9409	0.9706
8	0.9032	0.96

In conclusion it is not *compelling to expect chain indices (in contrast to direct indices) to develop in exploding or fading oscillations (or "bouncing" as Szulc called it) when prices go up and down in a cyclical fashion.* ◀

3.3 Drift–functions and growth rates

In this section we will discuss the so called "drift function" which allows us to compare a chain index with its direct counterpart. It turns out that this function is primarily determined by the differences between certain growth rates.

The next section will show how these differences can be interpreted in terms of covariances.

a) Drift–function and growth factors of chain indices

The ratio of for example the *chain* $\bar{P}^{LC}_{0t}$ with the corresponding direct index P^L_{0t}

$$D^{PL}_{0t} = \frac{\bar{P}^{LC}_{0t}}{P^L_{0t}} \tag{3.3.1}$$

is known as drift–function of Laspeyres price index.[130] The term "drift" does not imply that one of the two indices in question is regarded as being correct while the other is deemed incorrect. As the drift function is by definition a ratio it might also be reasonable to examine a difference instead. But the difference is directly related to the ratio: for example $\bar{P}^{LC}_{03} - P^L_{03}$ simply equals $P^L_{03}\left(D^{PL}_{03} - 1\right)$. Hence the difference does not convey additional information as compared with the "drift".

Analogously the drift of a Laspeyres *quantity* index is given by $D^{QL}_{0t} = \bar{Q}^{LC}_{0t}/Q^L_{0t}$. The Paasche price index drift is $D^{PP}_{0t} = \bar{P}^{PC}_{0t}/P^P_{0t}$, and the Fisher price index drift is

[130] Note, it is a comparison of a *chain* and not a link with the direct index.

$D_{0t}^{PF} = \sqrt{D_{0t}^{PL} D_{0t}^{PP}}$ respectively. For the drift of a Laspeyres *price* index we obtain

$$D_{02}^{PL} = \frac{\bar{P}_{02}^{LC}}{P_{02}^{L}} = \frac{\sum p_1 q_0 \sum p_2 q_1}{\sum p_2 q_0 \sum p_1 q_1} = \frac{P_{01}^{L} P_{12}^{L}}{P_{02}^{L}} = \frac{P_{12}^{L}}{P_{02}^{L}/P_{01}^{L}} = \frac{g_2^1}{g_2^0} \tag{3.3.2}$$

where g_t^k is the growth *factor* (and $g_t^k - 1$ the growth *rate*) measuring the change of prices from $t-1$ to t on the basis of quantities belonging to period k, that is

$$g_t^k = \frac{\sum p_t q_k}{\sum p_{t-1} q_k} = \sum \frac{p_t}{p_{t-1}} \frac{p_{t-1} q_k}{\sum p_{t-1} q_k}. \tag{3.3.3}$$

Especially the growth factor

$$g_t^{t-1} = P_t^{LC}$$

is the Laspeyres link, and the index

$$g_t^0 = P_{0t}^{L}/P_{0,t-1}^{L} = \frac{\sum p_t q_0}{\sum p_{t-1} q_0} = P_{t-1,t(0)}^{L}$$

is the rebased[131] direct Laspeyres index.

They are both arithmetic means of price relatives differing only with respect to the weights used.[132] It is easy to verify the *recursive* (or cumulative) nature of the chain Laspeyres price index drift:

Starting with $D_{01}^{PL} = 1$, we get successively

$$D_{02}^{PL} = \frac{g_2^1}{g_2^0}, D_{03}^{PL} = \frac{g_2^1 g_3^2}{g_2^0 g_3^0} = D_{02}^{PL} \frac{g_3^2}{g_3^0}, \text{and } D_{04}^{PL} = D_{03}^{PL} \frac{g_4^3}{g_4^0} \text{ etc.}$$

However, in the case of other types of drifts, the situation is not that comfortable. In terms of growth factors we can express various price indices and drifts as indicated in tab. 3.3.1.

[131] From base 0 to base $t-1$.

[132] The focus on a ratio of growth rates shows, that comparing growth rates of a Paasche– and a Laspeyres index calculation respectively (as sometimes done in official statistics) may indeed be useful to decide on the appropriate period for rebasing a Laspeyres index. To keep the "old" base (or "basket" of a Laspeyres index) may be legitimate as long as the growth rates don't differ too much.

Table 3.3.1: Price indices and drifts in terms of growth factors

	Laspeyres	Paasche
direct	$P^L_{0t} = g^0_1 g^0_2 \dots g^0_t = \prod_{\tau=1}^{\tau=t} g^0_\tau$	$P^P_{0t} = g^t_1 g^t_2 \dots g^t_t = \prod_{\tau=1}^{\tau=t} g^t_\tau$
chain	$\bar{P}^{LC}_{0t} = g^0_1 g^1_2 \dots g^{t-1}_t = \prod_{\tau=1}^{\tau=t} g^{\tau-1}_\tau$	$\bar{P}^{PC}_{0t} = g^1_1 g^2_2 \dots g^t_t = \prod_{\tau=1}^{\tau=t} g^\tau_\tau$
drift	$D^{PL}_{0t} = \frac{g^1_2 g^2_3 \dots g^{t-1}_t}{g^0_2 g^0_3 \dots g^0_t} = \prod_{\tau=2}^{\tau=t} \frac{g^{\tau-1}_\tau}{g^0_\tau}$	$D^{PP}_{0t} = \frac{g^1_1 g^2_2 \dots g^{t-1}_{t-1}}{g^t_2 g^t_3 \dots g^t_{t-1}} = \prod_{\tau=1}^{\tau=t} \frac{g^\tau_\tau}{g^t_\tau}$

It follows that there indeed is a recursive structure in the case of the Laspeyres chain index. The situation is less easier to understand in the case of the Paasche price index drift. We get

$$D^{PP}_{02} = \frac{\bar{P}^{PC}_{02}}{P^P_{02}} = \frac{g^1_1 g^2_2}{g^2_1 g^2_2} = \frac{g^1_1}{g^2_1} = \frac{Q^L_{12(0)}}{Q^L_{12(1)}}, D^{PP}_{03} = \frac{g^1_1 g^2_2}{g^3_1 g^3_2} = D^{PP}_{02} \frac{g^2_1 g^2_2}{g^3_1 g^3_2} = D^{PP}_{02} \frac{Q^L_{23(0)}}{Q^L_{23(2)}}$$

and $D^{PP}_{04} = D^{PP}_{03} \frac{g^3_1 g^3_2 g^3_3}{g^4_1 g^4_2 g^4_3}$ etc. By the same token the series of drifts of the Fisher price index develops in a rather complicated, and apparently non–recursive manner

$$D^{PF}_{02} = \sqrt{\frac{g^1_1 g^1_2}{g^2_1 g^0_2}}, D^{PF}_{03} = \sqrt{\frac{g^1_1 g^1_2 g^2_2 g^2_3}{g^3_1 g^0_2 g^3_2 g^0_3}}, D^{PF}_{04} = \sqrt{\frac{g^1_1 g^1_2 g^2_2 g^2_3 g^3_3 g^3_4}{g^4_1 g^0_2 g^4_2 g^0_3 g^4_3 g^0_4}}$$

and so on. As a consequence of eq. 3.3.1 we get the following relationship between Laspeyres direct indices and drifts

$$D^{PL}_{0t} P^L_{0t} = \left(D^{PL}_{0s} P^L_{0s}\right)\left(D^{PL}_{st} P^L_{st}\right) \tag{3.3.4}$$

such that

$$\frac{D^{PL}_{0t}}{D^{PL}_{0s} D^{PL}_{st}} = \frac{P^L_{0s} P^L_{st}}{P^L_{0t}}.$$

For other indices, like Paasche and Fisher, and their respective drifts a corresponding relation holds. Although by definition $\bar{P}_{0t} = \bar{P}_{0s}\bar{P}_{st}$ holds for any three periods 0, s, and t, drifts of a chain (Laspeyres) price index are not chains themselves. In general $D^{PL}_{0t} \neq D^{PL}_{0s} D^{PL}_{st}$. To infer how the drift may increase or decrease as time proceeds from s to t the following equation may be useful:

$$D^{PL}_{0t} = D^{PL}_{0s} \frac{D^{PL}_{st} P^L_{st}}{P^L_{0t} / P^L_{0s}} = D^{PL}_{0s} \frac{\bar{P}^{LC}_{st}}{P^L_{st(0)}} = D^{PL}_{0s} h_{st}. \tag{3.3.5}$$

Whether and if so, to which extent a Laspeyres price index drift increases ($D_{0t}^{PL} > D_{0s}^{PL}$), or decreases ($D_{0t}^{PL} < D_{0s}^{PL}$) depends on the factor h_{st}, which is the ratio of two growth factors

$$h_{st} = \frac{\bar{P}_{st}^{LC}}{P_{st(0)}^{L}} = \frac{g_{s+1}^{s} g_{s+2}^{s+1} \cdots g_{t}^{t-1}}{g_{s+1}^{0} g_{s+2}^{0} \cdots g_{t}^{0}} \tag{3.3.6}$$

As to the “behavior” of h it is difficult to draw a valid conclusion under general conditions. It is often maintained that estimates of the amount of an increase in prices using more recent weights (quantities) will yield lower results than estimates based on “old” weights. Such a pattern would for example imply $g_3^0 > g_3^1 > g_3^2$, and in the case of two adjacent drifts we have

$$D_{03}^{PL} = D_{02}^{PL} \frac{\bar{P}_{23}^{LC}}{P_{23(0)}^{L}} = D_{02}^{PL} \frac{P_{3}^{LC}}{P_{23(0)}^{L}} = D_{02}^{PL} h_{23} \qquad \text{, where } h_{23} = \frac{P_{3}^{LC}}{P_{23(0)}^{L}} = \frac{g_3^2}{g_3^0}$$

such that we would expect a decreasing drift[133] $D_{03}^{PL} < D_{02}^{PL}$ as $g_3^0 > g_3^2$.

Before examining the relative size of different growth rates in more detail, we should note in passing:

1. As a consequence of eq. 3.3.4 the drift function not only depends on the length of the interval $(0, t)$ in question, but also on how it is divided into subintervals. Thus the drift is determined by the whole pattern of the time series as well as the formula of the chain index used.

2. Again there is a reciprocal or “antithetic” relation between Laspeyres and Paasche. As a *Laspeyres* chain *price* index drifts away *up*–wards, biased from the corresponding direct price index, a chain *Paasche quantity* index will drift away *down*–wards biased, and vice versa. Both possibly with a cumulative bias as time goes on, and more and more links are linked together. Since the value index V_{0t} is always transitive (unlike most of the price and quantity indices), and the value index can be decomposed into a P^L– and a Q^P–index we for example get

$$D_{03}^{PL} = \frac{P_{01}^{L} P_{12}^{L} P_{23}^{L}}{P_{03}^{L}} = \frac{Q_{03}^{P}}{Q_{10}^{P} Q_{12}^{P} Q_{23}^{P}} = \frac{1}{D_{03}^{QP}} \tag{3.3.7}$$

   hence there is an inverse relation between price and quantity drift.

3. The product representations of eqs. 1.2.1 and 1.2.2 can also be expressed as

[133] In general we expect, however, a chain index will draw progressively away from the direct index, such that the drift grows as t becomes more and more distant from 0.

follows[134]:

$$\bar{P}^{LC}_{03} = g^0_1 g^1_2 g^2_3 \qquad (1.2.1a)$$

$$P^{L}_{03} = g^0_1 g^0_2 g^0_3. \qquad (1.2.2a)$$

The following numerical example (ex. 3.3.1) may help to understand better the development of the drift, and to assess the behavior of a drift over long intervals as determined by the drifts over shorter intervals.

To systematically summarize the information given in the growth factors g^k_t as defined in eq. 3.3.3 it is useful to consider the following matrix

$$\mathbf{G} = \begin{bmatrix} \mathbf{g}' \\ \mathbf{G}^* \end{bmatrix} \quad \text{where } \mathbf{G}^* = \begin{bmatrix} g^1_1 & g^1_2 & g^1_3 & g^1_4 \\ g^2_1 & g^2_2 & g^2_3 & g^2_4 \\ g^3_1 & g^3_2 & g^3_3 & g^3_4 \\ g^4_1 & g^4_2 & g^4_3 & g^4_4 \end{bmatrix}. \qquad (3.3.8)$$

The row vector $\mathbf{g}' = \begin{bmatrix} g^0_1 \; g^0_2 \; \ldots \end{bmatrix}$ contains the (rebased) Laspeyres indices, that is $P^L_{01(0)} = P^L_{01}$, $P^L_{12(0)}$, $P^L_{23(0)}$, and so on. The product of the main diagonal elements of $\mathbf{G}^*$ gives the value index. The matrix $\mathbf{G}$ summarizes all information which is necessary to assess the various drifts as well as the relation between all indices presented in tab. 3.3.1.

Example 3.3.1

In this example the different growth factors of a modified (quantities added) ex. 3.2.4 (chosen because of the interesting case of prices moving cyclically) will be calculated. Assume

	periods 0, 2, 4, ...		*periods 1, 3, 5, ...*	
good	*price*	*quantity*	*price*	*quantity*
no. 1	*6*	*6*	*12*	*4*
no. 2	*8*	*4*	*4*	*6*

The structure of the matrix $\mathbf{G}^*$ *is such that only three values occur: 1,* $a = 22/17 = 1.294$, *and* $a^{-1} = 17/22 = 0.773$. *In detail for the row vector we get* $\mathbf{g}' = \begin{bmatrix} a \; a^{-1} \; a \; \ldots \end{bmatrix}$, *and for the matrix*

$$\mathbf{G}^* = \begin{bmatrix} g^1_1 & g^1_2 & g^1_3 & g^1_4 \\ g^2_1 & g^2_2 & g^2_3 & g^2_4 \\ g^3_1 & g^3_2 & g^3_3 & g^3_4 \\ g^4_1 & g^4_2 & g^4_3 & g^4_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ a & a^{-1} & a & a^{-1} \\ 1 & 1 & 1 & 1 \\ a & a^{-1} & a & a^{-1} \end{bmatrix}.$$

The results for the indices listed in tab. 3.3.1 as well as for Fisher's indices are therefore as follows:

134 Interestingly in both cases, i.e. eq. 1.2.1a and eq. 1.2.2a the Paasche index analogon would be more difficult to understand. This shows once more that – contrary to a popular misunderstanding – the Laspeyres and the Paasche formula are not based on the same logic.

t	P^L_{0t}	P^P_{0t}	P^F_{0t}	$\bar{P}^{LC}_{0t}$	$\bar{P}^{PC}_{0t}$	$\bar{P}^{FC}_{0t}$
1	$a = 1.29$	1	$\sqrt{a} = 1.14$	$a = 1.29$	1	$\sqrt{a} = 1.14$
2	1	1	1	$a = 1.29$	$a^{-1} = 0.773$	1
3	$a = 1.29$	1	$\sqrt{a} = 1.14$	$a^2 = 1.68$	$a^{-1} = 0.773$	$\sqrt{a} = 1.14$
4	1	1	1	$a^2 = 1.68$	$a^{-2} = 0.597$	1
5	$a = 1.29$	1	$\sqrt{a} = 1.14$	$a^3 = 2.17$	$a^{-2} = 0.597$	$\sqrt{a} = 1.14$

Note that $\bar{P}^{FC}_{0t}$ doesn't drift away from P^F_{0t}. This is due to the fact that quantities are also changing cyclically as prices are. The example also reveals the reasons for the progressive drift of the Laspeyres and Paasche index and for the nonexistence of a Fisher drift.
According to tab. 3.3.1 we have $\bar{P}^{LC}_{0t} = g^0_1 g^1_2 \dots g^{t-1}_t$ and the factors are $a, 1, a, 1, a, \dots$. Likewise for the factors of which the chain Paasche index is composed we obtain $g^1_1, g^2_2, \dots, g^t_t$ the following series $1, a-1, 1, a-1, \dots$. Thus the drift of the Fisher index is given by

$$D^{PF}_{02} = \sqrt{\frac{g^1_1 g^1_2}{g^2_1 g^0_2}} = \sqrt{\frac{1 \cdot 1}{a \cdot a^{-1}}} = 1, D^{PF}_{03} = \sqrt{\frac{g^1_1 g^1_2 g^2_2 g^2_3}{g^3_1 g^0_2 g^3_2 g^0_3}} = \sqrt{\frac{1 \cdot 1 \cdot a^{-1} \cdot a}{1 \cdot a^{-1} \cdot 1 \cdot a}} = 1,$$

and so on. ◀

b) Drift–function and the Laspeyres–Paasche gap

It is a widespread expectation that as a rule (except for rather unusual cases like ex. 3.3.1) we will have the following relations:

a) $P^L_{0t} > P^P_{0t}$ in which case the difference γ_{0t} known as "Laspeyres–Paasche gap" (LPG) is positive $\gamma_{0t} = P^L_{0t} - P^P_{0t} > 0$

b) $\bar{P}^{LC}_{0t} < P^L_{0t} \Rightarrow D^{PL}_{0t} < 1$, and $\bar{P}^{PC}_{0t} > P^P_{0t} \Rightarrow D^{PP}_{0t} > 1$, and therefore also

c) $\gamma^C_{0t} = \bar{P}^{LC}_{0t} - \bar{P}^{PC}_{0t} < \gamma_{0t} > 0$, and finally there is also an interesting relationship between the Laspeyres–Paasche gap of links, that is $\gamma^C_t = P^{LC}_t - P^{PC}_t$ on the one hand and the overall gap $\gamma_{0t}C$ on the other hand (see eqs. 3.3.10 and 3.3.11 below).

The Australian Bureau of Statistics (ABS) has changed its deflation methodology (in response to the revised SNA and also to a similar change made in the USA[135]) by moving from "constant prices" deflation (with the direct Paasche index, that is Q^L_{0t} as the index of "volumes") to Laspeyres' *chain* volumes (that is $\bar{Q}^{LC}_{0t}$) mainly on the following grounds:

- less trouble with the appearance of new goods and services and the disappearance of old ones (a well known argument often put forward in favor of chaining)[136], and

[135] See part c) of this section.
[136] See sec. 3.7 for the argument of "greater reliability of year–to–year indexes" which for the ABS "considerably reinforces the case for chaining" (MC LENNAN (1998), p. 20).

- expectations of the kind mentioned above, that is "a reduced Laspeyres–Paasche gap" ($\gamma_{0t}^{C} < \gamma_{0t}$) because: "The changes in relative prices between consecutive pairs of years are ... obviously much smaller than the cumulative changes between the first and last years. The smaller the changes in the relative prices, the less sensitive are the volume measures to the choice of index numbers." [137].

Hence, as there is good reason to expect that γ_t^C will be small for all t, it appears reasonable to also expect $\gamma_{0t}^{C} < \gamma_{0t}$. We now try to examine the conditions in terms of growth rates, to be met in order to comply with the three expectations mentioned above.
For a) and b) to hold the following conditions have to be met:

t	gap $\gamma_{0t} = P_{0t}^{L} - P_{0t}^{P}$	$\bar{P}_{0t}^{LC} < P_{0t}^{L}$ $\Rightarrow D_{0t}^{PL} < 1$	$\bar{P}_{0t}^{PC} > P_{0t}^{P}$ $\Rightarrow D_{0t}^{PP} > 1$
2	$\gamma_{02} = g_1^0 g_2^0 - g_1^2 g_2^2 > 0$	$g_2^1 < g_2^0$	$g_1^1 > g_1^2$
3	$\gamma_{03} = g_1^0 g_2^0 g_3^0 - g_1^3 g_2^3 g_3^3 > 0$	$g_2^1 g_3^2 < g_2^0 g_3^0$	$g_1^1 g_2^2 < g_1^3 g_2^3$

It is easy to verify that condition c) (see above) follows from a) and b) because

$$g_2^1 < g_2^0 \rightarrow g_1^0 g_2^1 < g_1^0 g_2^0 \rightarrow \bar{P}_{02}^{LC} < P_{02}^{L}$$

and

$$g_1^2 < g_1^1 \rightarrow g_1^2 g_2^2 < g_1^1 g_2^2 \rightarrow \bar{P}_{02}^{PC} > P_{02}^{P}.$$

It remains, however, to be shown that $g_1^1 g_2^2 < g_1^0 g_2^1$ such that $\bar{P}_{02}^{PC} < \bar{P}_{02}^{LC}$ which can be done indirectly or may readily be inferred from eq. 3.3.9 (see below). Hence we finally have

$$g_1^2 g_2^2 < g_1^1 g_2^2 < g_1^0 g_2^1 < g_1^0 g_2^0,$$

or equivalently

$$P_{02}^{P} < \bar{P}_{02}^{PC} < \bar{P}_{02}^{LC} < P_{02}^{L}.$$

In basically the same manner we arrive at $P_{03}^{P} < \bar{P}_{03}^{PC} < \bar{P}_{03}^{LC} < P_{03}^{L}$. For the different types of Laspeyres–Paasche gaps (LPGs) we get:

t	$\gamma_{0t} = P_{0t}^{L} - P_{0t}^{P}$	$\gamma_{0t}^{C} = \bar{P}_{0t}^{LC} - \bar{P}_{0t}^{PC}$	$\gamma_{t}^{C} = P_{t}^{LC} - P_{t}^{PC}$
1	$g_1^0 - g_1^1$	$g_1^0 - g_1^1$	$g_1^0 - g_1^1$
2	$g_1^0 g_2^0 - g_1^2 g_2^2$	$g_1^0 g_2^1 - g_1^1 g_2^2$	$g_2^1 - g_2^2$
3	$g_1^0 g_2^0 g_3^0 - g_1^3 g_2^3 g_3^3$	$g_1^0 g_2^1 g_3^2 - g_1^1 g_2^2 g_3^3$	$g_3^2 - g_3^3$

[137] MC LENNAN (1998), p. 18, 20. This remark might lead to the misunderstanding that the total LPG is a kind of sum of the link LPGs which is not true however (see below eq. 3.3.11).

It is interesting to see which elements of the matrix

$$\mathbf{G} = \begin{bmatrix} g_1^0 & g_2^0 & g_3^0 \\ g_1^1 & g_2^1 & g_3^1 \\ g_1^2 & g_2^2 & g_3^2 \\ g_1^3 & g_2^3 & g_3^3 \end{bmatrix}$$

have to be taken into account in order to calculate various types of LPGs. Irrelevant elements are indicated by an asterisk such that the

direct LPGs are determined as follows:

$$\gamma_{02} \rightarrow \begin{bmatrix} g_1^0 & g_2^0 \\ * & * \\ g_1^2 & g_2^2 \end{bmatrix}, \gamma_{03} \rightarrow \begin{bmatrix} g_1^0 & g_2^0 & g_3^0 \\ * & * & * \\ * & * & * \\ g_1^3 & g_2^3 & g_3^3 \end{bmatrix}$$

and so on, and the

chain LPGs by:

$$\gamma_{02}^C \rightarrow \begin{bmatrix} g_1^0 & * \\ g_1^1 & g_2^1 \\ * & g_2^2 \end{bmatrix}, \gamma_{03}^C \rightarrow \begin{bmatrix} g_1^0 & * & * \\ g_1^1 & g_2^1 & * \\ * & g_2^2 & g_3^2 \\ * & * & g_3^3 \end{bmatrix},$$

where in each column the two elements determine the LPG of the respective link $\gamma_1^C, \gamma_2^C, \gamma_3^C$.

This shows that the time series of all sorts of indices, drifts and LPGs are uniquely determined by the structure of the matrix **G**, which in terms of ratios of expenditures, is given by

$$\mathbf{G} = \mathbf{G}_2 = \begin{bmatrix} g_1^0 & g_2^0 \\ g_1^1 & g_2^1 \\ g_1^2 & g_2^2 \end{bmatrix} = \begin{bmatrix} \frac{\sum p_1 q_0}{\sum p_0 q_0} & \frac{\sum p_2 q_0}{\sum p_1 q_0} \\ \frac{\sum p_1 q_1}{\sum p_0 q_1} & \frac{\sum p_2 q_1}{\sum p_1 q_1} \\ \frac{\sum p_1 q_2}{\sum p_0 q_2} & \frac{\sum p_2 q_2}{\sum p_1 q_2} \end{bmatrix}$$

and $\mathbf{G}_3, \mathbf{G}_4, \ldots$ correspondingly. The regular pattern of **G** in ex. 3.3.1 (and also ex. 3.4.1) is responsible for the cycles in the index functions which in turn produce some regularities concerning drifts and gaps.

We now turn to relationships between various LPGs and drifts. The LPG of chain indices (that is γ_{0t}^C) is not necessarily smaller than the LPG of the direct indices (γ_{0t}), nor is it simply a sum of the LPGs of the links (γ_t^C). For instance in ex. 3.3.1 we have:

t	γ_{0t}	γ_{0t}^C	γ_t^C
1	a	a	a
2	0	$(a^2-1)/a=0.52$	$(a-1)/a=0.23$
3	$a=1.29$	$(a^3-1)/a=0.90$	$a=1.29$
4	0	$(a^4-1)/a^2=1.08$	$(a-1)/a=0.234$
5	$a=1.29$	$(a^5-1)/a^2=1.57$	$a=1.29$
6	0	$(a^6-1)/a^2=1.71$	$(a-1)/a=0.23$

The equation $\gamma_{03}^C = P_{03}^L(D_{03}^{PL}-1)+P_{03}^P(1-D_{03}^{PP})+\gamma_{03}$, or in general

$$\begin{aligned}\gamma_{0t}^C &= P_{0t}^L(D_{0t}^{PL}-1)+P_{0t}^P(1-D_{0t}^{PP})+\gamma_{0t}\\ &= D_{0t}^{PL}\gamma_{0t}+P_{0t}^P(D_{0t}^{PL}-D_{0t}^{PP})\end{aligned} \tag{3.3.9}$$

shows how the chain LPG (γ_{0t}^C) is related to the direct LPG (γ_{0t}). It also turns out that it appears reasonable to expect $\gamma_{0t}^C < \gamma_{0t}$ when $D_{0t}^{PL} < 1$, and $D_{0t}^{PP} > 1$ hold by assumption. The term $D_{0t}^{PL}-1$ will be related to a covariance in sec. 3.4. We can also recognize a recursive system for the chain gap as follows

$$\begin{aligned}\gamma_{01}^C &= \gamma_1^C\\ \gamma_{02}^C &= g_2^1\gamma_{01}^C + g_1^1\gamma_2^C = P_2^{LC}\gamma_1^C + P_1^{PC}\gamma_2^C\\ \gamma_{03}^C &= g_3^2\gamma_{02}^C + g_1^1 g_2^2\gamma_3^C = P_3^{LC}\gamma_{02}^C + \bar{P}_{02}^{PC}\gamma_3^C\\ &= \bar{P}_{13}^{LC}\gamma_1^C + \bar{P}_{23}^{LC}P_1^{PC}\gamma_2^C + \bar{P}_{02}^{PC}\gamma_3^C\end{aligned}$$

or in general

$$\gamma_{0t}^C = P_t^{LC}\gamma_{0,t-1}^C + \bar{P}_{0,t-1}^{PC}\gamma_t^C \tag{3.3.10}$$

showing under which conditions the gap $\gamma_{0,t-1}^C$ may widen or narrow to γ_{0t}^C, and

$$\begin{aligned}\gamma_{0t}^C &= \bar{P}_{1t}^{LC}\gamma_1^C + \bar{P}_{2t}^{LC}P_1^{PC}\gamma_2^C + \bar{P}_{3t}^{LC}\bar{P}_{02}^{PC}\gamma_3^C + \dots\\ &\quad + \bar{P}_{t-2,t}^{LC}\bar{P}_{0,t-3}^{PC}\gamma_{t-2}^C + P_t^{LC}\bar{P}_{0,t-2}^{PC}\gamma_{t-1}^C + \bar{P}_{0,t-1}^{PC}\gamma_t^C.\end{aligned} \tag{3.3.11}$$

c) The "correct" and most "relevant" growth rate of real values (volumes)

It is often argued[138] that the growth factor of deflated values (i.e. "volumes") cannot be "relevant", when calculated on the basis of fixed prices related to a base period a long time ago. Furthermore some countries like the USA[139], and – as already mentioned – for example Australia, have changed their official measure of real GDP from

[138] This argument is referred to as argument **B3** in chapter 6 and it has a great deal of intuitive appeal.

[139] By January 1996 the US Bureau of Economic Analysis (BEA) adopted a chain Fisher quantity index as a measure of real growth (as recommended by SNA). Since this change coincided with a revision of the national income accounts the exact impact of the change in growth rate calculation was not easy to assess.

constant prices to a chain weighted measure in accordance with the SNA recommendations. Hence there already is some empirical evidence concerning the impact of such a change in the methodology of deflation (see also chapter 5). A decision in favor of chain–index–deflation was also made in the European System of National Accounts (ESA)[140]. Eurostat voted in favor of deflation, by using chain price indices mainly on the basis of the following two arguments[141]:

- volumes based on previous year prices give the "best", or most "accurate" estimates of changes in volume from one period to the next[142], and
- the use of chain indices would considerably ease the treatment of quality changes and new or disappearing products.

Interestingly these arguments were found compelling although very little seems to be known about how to measure the degree of "accuracy", "relevance" or "representativity". The price of greater ease in treating new or disappearing products is that volumes no longer represent values at *constant* prices of a base period, but values at *various* prices of different periods, in a sense like ordinary values. Moreover there are problems with non–additivity (see sec.5.3).[143] Several countries also criticized the possibly greatly increased costs due to the need for an annual update of weights. But on balance, in view of the vast majority of the EU Member Countries, the advantages of chaining would outstrip such disadvantages. In what follows we will try to

1. set out the *theoretical* motivation for a move to chain indices in case of deflation, (as to the idea of "better" growth rates we will mainly refer to Dutch authors) in more detail, and
2. review some of the *empirical* evidence concerning the consequences of such a move (as several countries have already made such a decision there is some

---

140 The rules laid down in ESA are in general clarified and detailed by Commission – or Council Decisions (or Regulations). They are binding for Member Countries of the European Union (EU). Moreover the introduction of the "Euro" currency entails a far-reaching monitoring of national fiscal policies of Member States. In this context the so called "Excessive Deficit Procedure", as a part of the "Stability and Growth Pact" calls for truly comparable volume figures. Therefore the deflation provisions set up so far for the EU are politically highly important. Although some decisions concerning deflation methodology were initially left undecided, until results of a special research programme were available, it was nonetheless not considered premature by Eurostat to settle the chain index issue. Germany was outvoted in this controversial issue, and was not even able to carry through its attempt to postpone a decision.

141 EU Commission Decision (as concerns the principles of measuring prices and volumes) Nr. 2223/96 approved by the Statistical Programme Committee in Stockholm, 27th May 1998.

142 This kind of comparison is considered more important than for example making consistent comparisons of growth rates between successive years. The argument outlined above is the deflation (or price weights) counterpart to the idea that in inflation measurement, chain indices are to be preferred because of their making use of the most "representative" or "relevant" quantity weights.

143 Germany emphasized in this issue not only the additivity problem, but also made some additional conceptual reservations: 1. lack of comparability of volumes over time, 2. it is not clear how to proceed in deflating quarterly national accounts and how they will be related to annual accounts, and 3. it would yield a better deflation if deflator–price indices would incorporate more prices and thus were broken down into more details rather than rebased more frequently.

knowledge about the impact on statistical results, but problems of their interpretation remain unsettled).

The main result is in short

> *Theoretical* arguments referring to the increased "relevance" of a growth rate do not necessarily call for a chain approach. Furthermore in addition to the best measure of the actual growth rate, there are other objectives of deflation of equal importance to be considered as well. *Empirically* well established is a tendency of *fixed* price volumes to overstate growth, but there is no basis for deciding on which set of growth rates is the "correct" or "most trustworthy".

The main theoretical issue is of course the relative importance attributed to the "relevance" of prices as weights, as opposed to the principle of pure comparison, which is supposed to be violated as soon as volumes are not only influenced by quantity but also by price movement.

After intense examination of the empirical consequences of a move from a fixed price (constant prices of a base period), to a chain approach in the USA YOUNG (1992) for example did not find a clear advantage of one method over the other[144].

It is often maintained that deflation with a Paasche price index referring to prices p_{i0} of a fixed base period 0 has the following shortcomings:

1. the choice of the base period affects the results of deflation,

2. each rebasing entails revising the entire "history" of GDP, and thus a new (as compared with hitherto published) series of growth rates when calculation is carried back, and

3. the traditional fixed–base year method of calculating real GDP, does not account for substitution and therefore tends to be biased.

The first point of course is trivial. It is only because of a continuously *changing* price structure that chain based measures of real growth seem to be independent of prices. The second and third point refer to what is intended to be measured in the case of "deflation". If the objective of deflation is to eliminate the impact of price movement there is clearly no reason for accounting for substitution, and the "regular rewriting of economic history" (BARR (1996)) though inconvenient is simply a consequence of choosing another set of prices.

[144] Young's result was inconclusive. After detailed comparison of various methods he ended up with some general remarks like the following: "A difference between two measures of real GDP is not evidence that one is wrong", or the worth of either method "lies in whether or not it proves useful in analysis. Viewed in this way, there can be more than one useful measure" (p. 35).

Four statisticians of the Dutch Centraal Bureau voor de Statistiek (CBS), AL, BALK, DE BOER and DEN BAKKER (1986) recommended the Vartia–I chain indices ($\bar{P}^{V1}_{0t}$) for deflation of National Accounts including Input–Output–Tables among other things[145], on the following grounds:

1. deflation requires factor reversibility, “that the value change is decomposed ... without remainder” (which however is already ensured by the weaker product test);

2. “the aim is not to isolate pure price components or pure quantity components” (p. 352) but rather to answer another (allegedly more relevant) question: “One aim might be to deflate the value of an aggregate at moment t in such a way as to afford optimum comparability with the data relating to moment 0. Alternatively, the aim might be to gain as clear a picture as possible of the changes in volume and aggregated price from one moment to the next” (p. 357);

3. because pure comparison was deemed less relevant an objective, the idea of “symmetry” gained more importance: the quantities $\mathbf{q}_0$ and $\mathbf{q}_t$ should be treated symmetrically (a fair allocation of the structural component – or substitution in response to changes in the prices – should be performed. It should neither be assigned exclusively to the price component as in the pair Q^L_{0t}, P^P_{0t} nor exclusively to the quantity component as done in P^L_{0t}, Q^P_{0t})[146];

4. direct indices have to be rebased from time to time and such “frequent changes in the index number formula used are not desirable” (p. 356), while chaining will produce less abrupt changes in the data.

Fisher indices (the SNA’s choice) though factor reversible were rejected because of their poor aggregation properties. But the Vartia formula no longer performs better in this respect as we move from P^{V1}_{0t} to $\bar{P}^{V1}_{0t}$. Lack of consistency in aggregation was even acknowledged in case of $\bar{P}^{V1}_{0t}$ by the Dutch authors.[147]
The second argument is the most important one in the present context. But aside from the fact that this argument is again an attempt to prove superiority of a chain

[145] The CBS authors for example also use the well known argument that a chain index “simplifies the treatment of disappearing or newly appearing goods or production processes” (p. 360). According to another argument, called “multiplication mystery” later (chapter 6) a direct index “fails to give a clear picture of the way in which the changes came about” (p. 358).

[146] Interestingly for Al, Balk, de Boer and den Bakker the problem of the direct Laspeyres index is not so much that it ignores the more recent (more “relevant” or “representative”) weights q_t, as chainers usually argue, but primarily its one sidedness, relying solely on q_0. Moreover average weights are supposed to be “relatively less arbitrary” and “can also be defended on the grounds that an average basket of quantities and prices is likely to be a more representative reference for the decomposition of the value change than either the initial or the final basket” (p. 356).

[147] (p. 356). In the sense of **A1** as defined in fig. 2.3.2. Violation of **A2**, that is inconsistency of volumes was well recognized too: “A practical disadvantage of this formula however, is that the indices thus obtained would not yield a table exhibiting additive consistency in real values” (p. 348), a serious defect, making a quantity index as an alternative estimate of volumes inappropriate.

index, $\bar{P}^{C}_{0t}$ by focusing on the link P^{C}_{0t} only, it is the alternative between "optimum comparability" and the "clear picture" which should be called into question. Strictly speaking the alternative is not comparability versus clarity, the problem is rather the length of the interval in time under consideration, that is $(0, t)$ on the one hand, and $(t - 1, t)$ on the other.[148]
Even if it is only the short interval that matters, and not comparability of the links among themselves, two questions remain to be answered:

- What makes comparisons over short intervals so much more important than those over longer intervals, that we should put up with some undisputed disadvantages (for example in aggregation)? And
- Why is the picture less "clear", when on the other hand primarily comparability over longer intervals is aimed at?

When a method is unable to serve two different purposes equally well it is of course always difficult to establish the greater "importance" of one purpose compared to another. Hence it is unlikely to find substantial hints to the first question. The answer to the second question is often given in terms of "accuracy" (see sec. 3.7 for this argument). It is maintained that in the Laspeyres approach the comparison of two adjacent periods, $t - 1$, and t is "contaminated"[149] by referring to a basket of a period 0 already long ago.
Moreover the alleged "clearer" picture of a comparison between t and $t - 1$, is not in itself proof that a comparison between t and 0 (a situation in which the allegedly less relevant objective "comparability" definitely gains importance) is made better indirectly not directly.
The choice between the fixed price approach and the chain approach, is often described as a choice between"simplicity" and "accuracy" (YOUNG (1992), p. 36). Simplicity, however, is nowadays far from being deemed advantageous. On the other hand "accuracy" may well be called into question, when the problem of "non–additivity" is considered, as admittedly a "main drawback with chain–weighted GDP is that GDP is no longer the sum of its components" (BARR (1996)), and that we can no longer compute "real dollar shares" or compare any two of the growth rates (YOUNG (1992), p. 35). Moreover: Is there no "contamination", or inaccuracy when a chain price index, like $\bar{P}^{V1}_{0t}$ refers to a *multitude* of quantity vectors $\mathbf{q}_0, \mathbf{q}_1, \ldots, \mathbf{q}_t$, or a resulting chain quantity index to a multitude of price vectors?
We may now turn to the empirical findings. It is known, for example that the US "significantly changed the official record of US growth" when BEA moved to a chain method. Growth since 1987 has been revised down which "caused considerable confusion", and the same applied to the then more recent periods (1993 – 1995), which

148 The controversy is not about the best growth factor (t compared with $t - 1$) taken in isolation (when links are *not* chained together), the problem is rather what is the correct representation of a development relating to a *number* of periods.

149 Op. cit.,p. 357. "In the interest of the greatest possible accuracy" a two period comparison should be based on the basket of no other period than $t - 1$ or t.

"made the current recovery look particularly weak" (BARR (1996)). The experiences made in Australia were largely the same (MC LENNAN (1998)).

The US change in methodology, was motivated mainly by the observation that some important commodities experienced a remarkable change in prices, in particular oil prices increased and prices of computers and their peripherals decreased. Under such conditions growth rates on the basis of a more recent price structure, will of course differ from growth rates based on "old" prices, and it appears reasonable to avoid an overestimation of growth by taking more recent prices.

It has often been observed that growth in more recent years measured by chain methods, falls short of growth measured for the same period using fixed prices of a year dating back long ago. The reason is that commodities, for which output grows more rapidly tends to be those for which prices increase slowly or even decrease. Thus taking more recent prices as weights will give commodities with strong growth less weight and therefore results in a smaller overall growth rate.

It should be noted, however, that correlation between prices and output varies, and that the effect described is likely to be different in the case of consumer goods as compared with investment.[150]

An interesting example for the opposite situation of *declining* oil prices, is the growth rate of "real GDP" of Norway, alternatively calculated using fixed prices of the constant base 1984 (CB) on the one hand, and calculated using previous year prices (PY) on the other hand:[151]

		1987	1988	1989
constant base period prices*	CB	4.9	3.0	5.2
previous year prices	PY	3.9	1.8	0.9

* fixed prices of 1984

The price of crude oil (one of the most important commodities for Norway's export)[152] was high in the base year (1984) and fell thereafter. Hence there is a strong plausibility, that CB overstates growth in 1989 substantially due to the high weight given to oil exports in 1984, and being kept constant thereafter, while the substantially smaller PY–growth amounting to 0.9% by 1989 will be found more realistic and more reliable. Moreover the more recent structure of prices seems to better reflect conditions actually faced.

Note that there is a substantial difference between CB and PY:

150 Conspicuously, it is precisely the task of making comparisons between growth rates of different aggregates where additivity would be useful.
151 Data taken from an internal paper distributed in a task–force meeting of Eurostat.
152 This of course makes the example of Norway rather extreme.

- What is the total and the average growth? Multiplying growth factors on the basis of PY gives $1.039 \cdot 1.018 \cdot 1.009 = 1.067$, thus a total growth of 6.7% (or an average annual growth of 2.2%) in contrast to 36% (average 4.2%) obtained by the CB method.

- Growth in 1989 using CB was higher (5.2%) than in 1988 (3% only), whereas judged on the basis of PY the opposite is true. According to PY the Norwegian economy unreletingly ran into difficulties (constantly declining growth rates), whilst growth measured by CB displays a much more comfortable or even promising picture.

Both observations, less overall growth and declining growth when measured by PY can easily be explained by the same factor: declining crude oil prices. Correspondingly the CB concept may well appear to be over–optimistic.

But comparison with the previous year is not the only relevant problem of interpreting growth rates. Estimation of "true" growth over a short interval is not an argument useful to discard the direct index approach, in favor of the chain index approach. The chain (multiplication), not the link is the problem.

> The question is not only which is the best measure of growth for the current year. Also what is important is: are successive growth rates of say five years or so (the length of a business cycle) *comparable* in the sense of truly reflecting the movement of "volume" or are the growth rates also affected by changes in prices, an influence which to eliminate is the prime concern of deflating?

We now try to show that already in a comparison of two successive PY–type growth rates (relating to period t and $t-1$ respectively) two price structures are involved, p_{t-1} and p_{t-2}.[153] The ratio of two growth factors as measured by the PY method is given by

$$\frac{Q_t^{LC}}{Q_{t-1}^{LC}} = \frac{\sum q_t p_{t-1}}{\sum q_{t-1} p_{t-2}} \cdot \frac{\sum q_{t-2} p_{t-2}}{\sum q_{t-1} p_{t-1}} = \frac{P_{t-1}^{PC} + \Delta_t}{P_{t-1}^{LC} + \Delta_{t-1}} \tag{3.3.12}$$

where the links P_{t-1}^{PC} and P_{t-1}^{LC} clearly measure the change of prices from $t-2$ to $t-1$, and where

$$\Delta_t = \frac{\sum (q_t - q_{t-1}) \cdot p_{t-1}}{\sum q_{t-1} p_{t-2}} \quad \text{and} \quad \Delta_{t-1} = \frac{\sum (q_{t-1} - q_{t-2}) \cdot p_{t-1}}{\sum q_{t-2} p_{t-2}}$$

are measures of the change in quantity. Thus when a chain approach in calculating growth rates is preferred it should be recognized that price movement seems to be

[153] Whenever real growth is measured over a longer interval using a chain quantity index there are of course many more price structures, like $p_{t-3}, \ldots$ etc. influencing the result.

incompletely eliminated. If prices would not change from $t-2$ to $t-1$ we, in eq. 3.3.12, would get $P^{PC}_{t-1} = P^{LC}_{t-1}$ and upon substituting p_{t-1} for p_{t-2}

$$\Delta_t = \frac{\sum q_t p_{t-1}}{\sum q_{t-1} p_{t-1}} - 1 \quad \text{and} \quad \Delta_{t-1} = \frac{\sum q_{t-1} p_{t-1}}{\sum q_{t-2} p_{t-1}} - 1$$

such that we obtain

$$\frac{Q^{LC}_t}{Q^{LC}_{t-1}} = \frac{\Delta_t + 1}{\Delta_{t-1} + 1} = \frac{\sum q_t p_{t-1}}{\sum q_{t-1} p_{t-1}} \cdot \frac{\sum q_{t-2} p_{t-1}}{\sum q_{t-1} p_{t-1}} \tag{3.3.13}$$

which differs from eq. 3.3.12, and clearly shows that price movement affects the comparison of two successive growth rates of volumes in the case of the chain method. By contrast such a comparison, in the case of the constant–price method is independent of such a price movement, because the expression analogous to eq. 3.3.12 is

$$\frac{Q^L_{0t}}{Q^L_{0,t-1}} \bigg/ \frac{Q^L_{0,t-1}}{Q^L_{0,t-2}} = \frac{\sum q_t p_0}{\sum q_{t-1} p_0} \cdot \frac{\sum q_{t-2} p_0}{\sum q_{t-1} p_0}. \tag{3.3.14}$$

To avoid an overrating of growth in volumes is quite different from comparing successive growth rates or from gaining a measure of the level of (deflated) production by chaining. Furthermore in assessing the "most relevant" actual rate of growth a single growth rate is taken in isolation. This does not necessarily provide an argument for chaining.

d) Comparisons of indices published monthly or quarterly

In the most part price indices are published regularly on a sub–annual basis, that is monthly or quarterly. It is most popular to compare the value, a (chain) price index in a certain month, in year t, to the value of the corresponding month of the previous year $t-1$. It is usually deemed necessary to make such a comparison on the basis of the same weighting structure, i.e. the same quantities (or "basket"). It is interesting to note, and it also gives rise to some criticism of chain indices, that such a comparison in practice is already a function of quantities relating to *two* years.
The conventional expression of a single link, like $P^{LC}_t = \sum p_t q_{t-1} / \sum p_{t-1} q_{t-1}$ or so does not make clear that t in the case of quantities refers to a year, whereas t with respect to prices (p), usually refers to months within the year in question. The problem under consideration ensues from the fact that prices are reported on a monthly basis, while a renewal of weights is not done more frequently than annually.
For example assume that a month (say May) in year t has to be compared with the same month of the preceding year $t-1$, and an updating of weights (always relating to a sum or an average [symbol ∅] of the preceding *year*) takes place only once a year, for example in December.
It might be helpful to demonstrate the problem as follows:
We have to compare:

Prices referring to	Jan. 99	...	May 99	...	Dec. 99	Jan. 00
Weights referring to	∅ 98	...	∅ 98	...	∅ 98	∅ 99

with

Prices referring to	Jan. 98	...	May 98	...	Dec. 98	Jan. 99
Weights referring to	∅ 97	...	∅ 97	...	∅ 97	∅ 98

Hence in the comparison of May 99 with May 98 there are already *two* baskets involved ∅ 97 and ∅ 98. To avoid such a situation it would be necessary to recalculate all values of the index at the end of the year (say 1998) when the new basket (to be introduced in Jan. 99) is available. Of course the result of such a recalculation (for example for May 1998 with weights ∅ 98) may well differ from the previously published data for May 1998 (based of course on the basket ∅ 97). Needless to say that such a recalculation would be arduous, costly, if not found superfluous, and thus as a rule it does not take place.

3.4 Intertemporal correlation

The drift D_{02}^{PL} can also be expressed in terms of the famous relation found by L. v. Bortkiewicz (sec. 2.1d).

a) Basic relation between drift and intertemporal covariance

Denote growth *factors* of individual prices by

$$x_{i,12} = \frac{p_{i2}}{p_{i1}}$$

(correspondingly $x_{i,23} = p_{i3} / p_{i2}$ and so on), and quantity *relatives* (*cumulative* changes) by

$$y_{i,01} = \frac{q_{i1}}{q_{i0}}$$

(correspondingly $y_{i,02} = q_{i2} / q_{i0}$ and so on), and consider weighted means, with weights being (hybrid) expenditure shares $p_{i1} q_{i0} \big/ \sum p_{i1} q_{i0}$. Given

$$\bar{x}_{12} = \sum \frac{p_2}{p_1} \frac{p_1 q_0}{\sum p_1 q_0} = \frac{\sum p_2 q_0}{\sum p_1 q_0} = P_{02}^L \big/ P_{01}^L = P_{12(0)} \tag{3.4.1}$$

or in general using the term g_t^k of sec. 3.3

$$\bar{x}_{t-1,t} = \frac{\sum p_t q_0}{\sum p_{t-1} q_0} = \frac{P_{0t}^L}{P_{0,t-1}^L} = P_{t-1,t(0)}^L = g_t^0 \tag{3.4.1a}$$

and

$$\bar{y}_{01} = \sum \frac{q_1}{q_0} \frac{p_1 q_0}{\sum p_1 q_0} = \frac{\sum p_1 q_1}{\sum p_1 q_0} = Q^P_{01} \tag{3.4.2}$$

or in general

$$\bar{y}_{0,t-1} = Q^P_{0,t-1} \tag{3.4.2a}$$

Now it can easily be seen, that the covariance between growth factors of prices, and the *cumulated* change in quantities is to be expressed as follows (omitting i for convenience)

$$\begin{aligned} \text{Cov}\,(x_{12}, y_{01}) &= \sum \left(x_{i,12} - \bar{x}_{12}\right)\left(y_{i,01} - \bar{y}_{01}\right) \frac{p_1 q_0}{\sum p_1 q_0} \\ &= \frac{\sum p_2 q_1}{\sum p_1 q_0} - \bar{x}_{12}\bar{y}_{01} = \frac{Q^P_{01}}{P^L_{01}} \left(P^L_{01} P^L_{12} - P^L_{02}\right) \\ &= Q^P_{01} \left(P^{LC}_2 - P^L_{12(0)}\right) = Q^P_{01} \left(g^1_2 - g^0_2\right) \end{aligned} \tag{3.4.3}$$

and the drift according to eq. 3.3.1 is given by

$$D^{PL}_{02} = \frac{\text{Cov}\,(x_{12}, y_{01})}{\bar{x}_{12} \cdot \bar{y}_{01}} + 1 \tag{3.4.4}$$

$$\frac{\text{Cov}\,(x_{12}, y_{01})}{\bar{x}_{12} \cdot \bar{y}_{01}} = \frac{P^{LC}_2}{P^L_{12(0)}} - 1 = \frac{g^1_2}{g^0_2} - 1 \tag{3.4.4a}$$

In a similar manner with x_{23} and y_{02} defined correspondingly and weights $p_2 q_0 / \sum p_2 q_0$ we get

$$\begin{aligned} \text{Cov}\,(x_{23}, y_{02}) &= \sum \left(x_{i,23} - \bar{x}_{23}\right)\left(y_{i,02} - \bar{y}_{02}\right) \frac{p_2 q_0}{\sum p_2 q_0} \\ &= Q^P_{02} \left(P^{LC}_3 - P^L_{23(0)}\right) = Q^P_{02} \left(g^2_3 - g^0_3\right). \end{aligned} \tag{3.4.5}$$

Again we have

$$\frac{\text{Cov}\,(x_{23}, y_{02})}{\bar{x}_{23} \cdot \bar{y}_{02}} + 1 = \frac{P^{LC}_3}{P^L_{23(0)}}$$

and the right hand side of this equation can – by virtue of eq. 3.3.1 – be expressed also as D^{PL}_{03} / D^{PL}_{02} showing the cumulative structure of the drift in terms of the covariance

$$\begin{aligned} D^{PL}_{03} &= D^{PL}_{02} \left(\frac{\text{Cov}\,(x_{23}, y_{02})}{\bar{x}_{23} \cdot \bar{y}_{02}} + 1\right) \\ &= \left(\frac{\text{Cov}\,(x_{12}, y_{01})}{\bar{x}_{12} \cdot \bar{y}_{01}} + 1\right) \left(\frac{\text{Cov}\,(x_{23}, y_{02})}{\bar{x}_{23} \cdot \bar{y}_{02}} + 1\right). \end{aligned} \tag{3.4.6}$$

Therefore we may state for the relationship between the covariance $\mathrm{Cov}(x_{t-1,t}, y_{0,t-1})$ between price growth rates (x) and quantity relatives (y) on the one hand, and the drift of the Laspeyres price index D^{PL} on the other hand

t	Cov(x, y) positive / negative if	consequence
1	$P_1^{LC} = \frac{\sum p_1 q_0}{\sum p_0 q_0} = P_{01(0)}^{L} = P_{01}^{L}$ Cov(x, y) is with necessity zero	no drift possible, first link = first direct index
2	a) $P_2^{LC} = \frac{\sum p_2 q_1}{\sum p_1 q_1} > P_{12(0)}^{L} = \frac{\sum p_2 q_0}{\sum p_1 q_0}$ then $\mathrm{Cov} > 0$ b) $P_2^{LC} < P_{12(0)}^{L}$ then $\mathrm{Cov} < 0$	a) positive drift, that is $\bar{P}_{02}^{LC} > P_{02}^{L}$ b) negative drift
3	$P_3^{LC} = \frac{\sum p_3 q_2}{\sum p_2 q_2} > P_{23(0)}^{L} = \frac{\sum p_3 q_0}{\sum p_2 q_0}$ then $\mathrm{Cov} > 0$	drift increasing (if $\mathrm{Cov} < 0$ then decreasing)

or in general

$$\begin{aligned} \text{if } P_t^{LC} > P_{t-1,t(0)}^{L} &\Rightarrow \mathrm{Cov}(x_{t-1,t}, y_{0,t-1}) > 0 \\ &\Rightarrow \text{drift is increasing } D_{0t}^{PL} > D_{0,t-1}^{PL} \\ \text{if } P_t^{LC} < P_{t-1,t(0)}^{L} &\Rightarrow \mathrm{Cov}(x_{t-1,t}, y_{0,t-1}) < 0 \\ &\Rightarrow \text{drift is decreasing } D_{0t}^{PL} < D_{0,t-1}^{PL} \end{aligned}$$

It is obviously difficult to draw general conclusions concerning the sign of the covariance, and its absolute numerical value. Therefore it appears hard to predict

- whether a chain index is drifting upwards or downwards, or
- whether or not the chain index remains in the spread between the direct Laspeyres and Paasche index [154], and
- though the covariance (Cov) determines the drift, and Cov may well change its sign, this does not imply that the drift is changing sign as well (but only that the drift is increasing or decreasing), and finally

[154] For example ALLEN (also quoted in HANSEN and LUCAS (1984), p. 28 as "the last word") conjectured "there is no reason to expect that the chain Laspeyres index drifts above the direct Laspeyres index" (ALLEN (1975), p.188). This would mean, that the drift D^{PL} may safely be assumed to be (almost necessarily) negative. This is certainly erroneous, because in case of a cyclical price movement the chain index may well exceed the direct index, and it will do so progressively as prices continue to change cyclically.

- the covariance does not give a straightforward explanation for a chain index drifting further and further away, when a cyclical movement is repeated again and again.

As a rule the amount of drift, depends on the *cumulated* change of quantities associated with the change of prices, and on the type of time series, given for prices and quantities (either monotonically increasing or decreasing, or with a cyclical movement respectively). The problem with such a statement is, however, that empirical findings concerning time series of prices and quantities are based on indices of prices and quantities. Given a cyclical movement one might suppose that a chain index will not necessarily reflect also cycles as already shown in sec. 3.3. It should also be noted that

1. the variables x and y of which the covariance is under consideration here, refer to different periods, x refers to prices of periods $t-1$ and t, but y is a measure of cumulative change of quantities over an interval $(0, t-1)$, and
2. it is the correlation between two differently defined changes, that determines the change of drift, not the drift itself,
3. the periods to which the relevant covariances refer are changing, we have a *series of covariances* $\text{Cov}(x_{12}, y_{01}), \text{Cov}(x_{23}, y_{02}), \text{Cov}(x_{34}, y_{03}), \ldots$ determining a series of drifts.

For all these reasons it is difficult to draw some general conclusions concerning the drift of a chain Laspeyres price index (relative to the direct Laspeyres price index).

Table 3.4.1: Relationships between indices, growth factors and covariances

$t-1$	$\bar{x}_{t-1,t}$	$\bar{y}_{0,t-1} = Q^P_{0,t-1}$	$\text{Cov}\left(x_{t-1,t}, y_{0,t-1}\right)/Q^P_{0,t-1}$
1	$\frac{\sum p_2 q_0}{\sum p_1 q_0} = g_2^0$	$\frac{\sum p_1 q_1}{\sum p_1 q_0} = Q^P_{01}$	$\frac{\sum p_2 q_1}{\sum p_1 q_1} - \frac{\sum p_2 q_0}{\sum p_1 q_0} = g_2^1 - g_2^0$
2	$\frac{\sum p_3 q_0}{\sum p_2 q_0} = g_3^0$	$\frac{\sum p_2 q_2}{\sum p_2 q_0} = Q^P_{02}$	$\frac{\sum p_3 q_2}{\sum p_2 q_2} - \frac{\sum p_3 q_0}{\sum p_2 q_0} = g_3^2 - g_3^0$
3	$\frac{\sum p_4 q_0}{\sum p_3 q_0} = g_4^0$	$\frac{\sum p_3 q_3}{\sum p_3 q_0} = Q^P_{03}$	$\frac{\sum p_4 q_3}{\sum p_3 q_3} - \frac{\sum p_4 q_0}{\sum p_3 q_0} = g_4^3 - g_4^0$

The results so far are summarized in tab. 3.4.1, where also a certain pattern is becoming apparent.
Obviously what determines both, the sign and the size of the covariance is the relation between two growth factors, derived by using different weights, that is

$$g_t^{t-1} = P_t^{LC} = \frac{\sum p_t q_{t-1}}{\sum p_{t-1} q_{t-1}} \quad \text{and} \quad g_t^0 = P_{t-1,t(0)}^L = \frac{P_{0t}^L}{P_{0,t-1}^L} = \frac{\sum p_t q_0}{\sum p_{t-1} q_0}$$

respectively. But it is certainly difficult to find rules governing sign and amount of this difference.

Example 3.4.1

Calculate covariance and means for ex. 3.2.1 and its variant a (see ex. 3.2.2).

1) *Calculation for ex. 3.2.1*
   Since $P_{01}^L = P_{02}^L = 1$ *we also have* $\bar{x}_{12} = P_{02}^L / P_{01}^L = 1$ *and* $\bar{y}_{01} = Q_{01}^P = 135/96 = 1.4063$, $P_{12}^L = 140/135 = 1.037$. *Hence* $Cov(x, y) = Q_{01}^P (P_1^L - 1) = 5/96 = 0.0521$, *and* $D_{02}^{PL} = P_{12}^L = 1.037 > 1$ *because* $\bar{P}_{02}^{LC} = 1.037 > P_{02}^L = 1$.
2) *Calculation for variant a, ex. 3.2.2*
   $\bar{x}_{12} = 1$, $\bar{y}_{01} = Q_{01}^P = 180/96 = 1.875$, $P_{12}^L = 160/180 = 0.888$, *and* $Cov(x, y) = -20/96 = -0.208$ *because* $\bar{P}_{02}^{LC} = 0.8888 < P_{02}^L = 1$. $D_{02}^{PL} = -20/180 + 1 = 0.888$.

◀

b) Series of drifts and intertemporal covariances, cyclical price movement

It is useful to first examine a numerical example, again with cyclical movement of prices (and quantities), a bit more general than ex. 3.3.1, in order to see the development of the Laspeyres price index drift D_{0t}^{PL} as well as the Laspeyres–Paasche gap (LPG) depending on the covariance.

Example 3.4.2

Consider the following situation with periodically recurring amounts of expenditure, called A *through* D *for short:*

	q *even*	q *uneven*
p *even*	$A =$ $\sum p_0 q_0 = \sum p_2 q_0 = \sum p_4 q_0 = \ldots$ $\sum p_0 q_2 = \sum p_2 q_2 = \sum p_4 q_2 = \ldots$	$B =$ $\sum p_0 q_1 = \sum p_2 q_1 = \sum p_4 q_1 = \ldots$ $\sum p_0 q_3 = \sum p_2 q_3 = \sum p_4 q_3 = \ldots$
p *uneven*	$C =$ $\sum p_1 q_0 = \sum p_3 q_0 = \sum p_5 q_0 = \ldots$ $\sum p_1 q_2 = \sum p_3 q_2 = \sum p_5 q_2 = \ldots$	$D =$ $\sum p_1 q_1 = \sum p_3 q_1 = \sum p_5 q_1 = \ldots$ $\sum p_1 q_3 = \sum p_3 q_3 = \sum p_5 q_3 = \ldots$

With $\alpha = C/A$ *and* $\beta = D/B$ *the structure of the matrix* **G** *(defined in eq. 3.3.8) is now*

$$\begin{bmatrix} g_1^0 & g_2^0 & g_3^0 \\ g_1^1 & g_2^1 & g_3^1 \\ g_1^2 & g_2^2 & g_3^2 \\ g_1^3 & g_2^3 & g_3^3 \end{bmatrix} = \begin{bmatrix} \alpha & \alpha^{-1} & \alpha \\ \beta & \beta^{-1} & \beta \\ \alpha & \alpha^{-1} & \alpha \\ \beta & \beta^{-1} & \beta \end{bmatrix},$$

to be compared with

$$\begin{bmatrix} a & a^{-1} & a \\ 1 & 1 & 1 \\ a & a^{-1} & a \\ 1 & 1 & 1 \end{bmatrix}$$

in ex. 3.3.1. Hence the results found in what follows can easily be translated into the more special case of $\alpha = a$ and $\beta = 1$ of ex. 3.3.1.
The covariance $Cov(x_{t-1,t}, y_{0,t-1})$ responsible for the change of the drift $D^{PL}_{0,t-1} \rightarrow D^{PL}_{0t}$ can only take two values

$$\delta = \frac{BC - AD}{C^2} = \frac{B}{C}\left(1 - \frac{\beta}{\alpha}\right)$$

and 0 indicating that either the drift is $D^{PL}_{0t} = \frac{\alpha}{\beta} D^{PL}_{0,t-1}$, or that the drift does not change such that $D^{PL}_{0t} = D^{PL}_{0,t-1}$. On the basis of this matrix **G** *we get the following results for four indices, two drifts, the covariance (Cov), and the LPG (Laspeyres–Paasche gap) of the chain indices*[155] *γ^{C}_{0t}:*

t	P^L_{0t}	P^P_{0t}	$\bar{P}^{LC}_{0t}$	D^{PL}_{0t}	*Cov*	$\bar{P}^{PC}_{0t}$	D^{PP}_{0t}	γ^C_{0t}
1	α	β	α	*1*	δ	β	*1*	$\alpha - \beta$
2	*1*	*1*	α/β	α/β	*0*	β/α	β/α	$(\alpha^2 - \beta^2)/\alpha\beta$
3	α	β	α^2/β	α/β	δ	β^2/α	β/α	$(\alpha^3 - \beta^3)/\alpha\beta$
4	*1*	*1*	α^2/β^2	α^2/β^2	*0*	β^2/α^2	β^2/α^2	$(\alpha^4 - \beta^4)/\alpha^2\beta^2$
5	α	β	α^3/β^2	α^2/β^2	δ	β^3/α^2	β^2/α^2	$(\alpha^5 - \beta^5)/\alpha^2\beta^2$
6	*1*	*1*	α^3/β^3	α^3/β^3	*0*	β^3/α^3	β^3/α^3	$(\alpha^6 - \beta^6)/\alpha^3\beta^3$
7	α	β	α^4/β^3	α^3/β^3	δ	β^4/α^3	β^3/α^3	$(\alpha^7 - \beta^7)/\alpha^3\beta^3$
8	*1*	*1*	α^4/β^4	α^4/β^4	*0*	β^4/α^4	β^4/α^4	$(\alpha^8 - \beta^8)/\alpha^4\beta^4$

Interestingly again, like in ex. 3.3.1 there is no drift of the Fisher index, that is $D^{PF}_{0t} = 1$ (or equivalently $\bar{P}^{FC}_{0t} = P^F_{0t}$), and both indices take on two alternating values only, $\sqrt{\alpha\beta}$ and 1. Though there definitely exists a LPG, the index $\bar{P}^{FC}_{0t}$ seems to enjoy an advantage over both other chain indices, $\bar{P}^{LC}_{0t}$ and $\bar{P}^{PC}_{0t}$, in that, some kind of "counteracting" of the drifts D^{PL}_{0t} and D^{PP}_{0t} (or more distinctly: an inverse relation between them, irrespective of the sign of the covariance δ, as can easily be seen by comparing the two relevant columns) prevents the drift D^{PF}_{0t} from existing. The SNA does not seem to have taken this possibility into account, because the SNA in general recommends *not applying chaining (with whichever type of links) when prices are moving cyclically.* ◀

We may now ask if it is possible for example that the covariance may change sign in such a way that $D^{PL}_{03} < D^{PL}_{02}$, or even $D^{PL}_{03} < 1$ is following a positive drift $D^{PL}_{02} > 1$, or

[155] The LPG of the *direct* indices is oscillating, taking values $\alpha - \beta$, 0 and so on.

in other words though $\bar{P}^{LC}_{02} > P^L_{02}$ we have $\bar{P}^{LC}_{03} < P^L_{03}$.[156] According to eq. 3.4.6 the *change* of the drift is related to the covariance in the following way

$$\begin{aligned} D^{PL}_{0t} &= D^{PL}_{0,t-1}\left(\frac{\mathrm{Cov}\left(x_{t-1,t}, y_{0,t-1}\right)}{\bar{x}_{t-1,t}\bar{y}_{0,t-1}} + 1\right) \\ &= D^{PL}_{0,t-1}\left(\frac{\mathrm{Cov}\left(x_{t-1,t}, y_{0,t-1}\right)}{P^L_{t-1,t(0)}Q^P_{0,t-1}} + 1\right) \end{aligned} \tag{3.4.6a}$$

where $P^L_{t-1,t(0)} = P^L_{0t} / P^L_{0,t-1}$. The term in brackets reduces to

$$\frac{P^{LC}_t}{P^L_{t-1,t(0)}} = \frac{g^{t-1}_t}{g^0_t}$$

such that

$$D^{PL}_{0t} = \frac{g^{t-1}_t}{g^0_t} D^{PL}_{0,t-1} \tag{3.4.7}$$

The relevant terms in the **G**–matrix are now

$$D^{PL}_{02} \rightarrow \begin{bmatrix} * & g^0_2 \\ * & g^1_2 \\ * & * \end{bmatrix}, \qquad D^{PL}_{03} \rightarrow \begin{bmatrix} * & * & g^0_3 \\ * & * & * \\ * & * & g^2_3 \\ * & * & * \end{bmatrix},$$

and for the situation described in the question above to be possible the conditions to be satisfied are

a) $g^0_2 > g^1_2 \rightarrow \mathrm{Cov}(x_{1,2}, y_{0,1}) = Q^P_{01}\left(g^1_2 - g^0_2\right) < 0$, or $\sum p_1 q_1 \sum p_2 q_0 > \sum p_1 q_0 \sum p_2 q_1$

b) $g^2_3 > g^0_3 \rightarrow \mathrm{Cov}(x_{2,3}, y_{0,2}) = Q^P_{02}\left(g^2_3 - g^0_3\right) < 0$, or $\sum p_3 q_2 \sum p_2 q_0 > \sum p_3 q_0 \sum p_2 q_2$

Another expression of the two inequalities is

a) $\bar{P}^{LC}_{02} < P^L_{02}$ and b) $P^L_{02} P^{LC}_3 < P^L_{03}$

In ex. 3.4.2 this would be impossible[157], because there the second (b) inequation would read as follows: $CA > CA$ (in this example we can only have $g^2_3 = g^0_3$, and hence a zero covariance). To facilitate notation we introduce the following symbols:

[156] Or the opposite inequalities as derived in ex. 3.4.3 .

[157] The restrictions in ex. 3.4.2 are such that $A_1 = A_2 = A_3$ are equal (A), and the same is true for all B, C and D.

	q even	q uneven
p even	$A_1 = \sum p_0 q_0, A_2 = \sum p_2 q_0, A_3 = \sum p_2 q_2$	$B_1 = \sum p_0 q_1, B_2 = \sum p_2 q_1$
p uneven	$C_1 = \sum p_1 q_0, C_2 = \sum p_3 q_0, C_3 = \sum p_3 q_2$	$D_1 = \sum p_1 q_1$

Now the conditions are $\frac{A_2}{C_1} > \frac{B_2}{D_1}$ and $\frac{C_2}{A_2} < \frac{C_3}{A_3}$, and the covariances $\mathrm{Cov}(x_{t-1,t}, y_{0,t-1})$ are given by

$$\frac{D_1}{C_1}\left(\frac{B_2}{D_1} - \frac{A_2}{C_1}\right) \qquad \text{for } t = 2$$

and

$$\frac{C_2}{A_2} < \frac{C_3}{A_3} \qquad \text{for } t = 3.$$

Example 3.4.3

The prices and quantities are chosen in such a way that conditions mentioned above are met:

p_0	q_0	p_1	q_1	p_2	q_2	p_3
8	6	6	3	8	8	10
12	6	15	5	18	4	14

$A_1 = 120$, $A_2 = 156$, $A_3 = 136$, $B_1 = 84$, $B_2 = 114$, $C_1 = 126$, $C_2 = 144$, $C_3 = 136$, $D_1 = 93$, $\sum p_1 q_2 = 108$, *and* $\sum p_0 q_2 = 112$. *This gives the following results (only some few entries in the cells for* $t = 3$ *because no assumptions made for* q_3*):*

	direct indices				*chain indices*			
	Laspeyres		*Paasche*		*Laspeyres*		*Paasche*	
t	P^L_{0t}	Q^L_{0t}	P^P_{0t}	Q^P_{0t}	$\bar{P}^{LC}_{0t}$	$\bar{Q}^{LC}_{0t}$	$\bar{P}^{PC}_{0t}$	$\bar{Q}^{PC}_{0t}$
1	*1.05*	*0.7*	*1.107*	*0.738*	*1.05*	*0.7*	*1.107*	*0.738*
2	*1.3*	*0.933*	*1.214*	*0.872*	*1.287*	*0.813*	*1.394*	*0.881*
3	*1.2*				*1.287*			

The drifts are $D^{PL}_{02} = 0.9901$ *and* $D^{PL}_{03} = 1.0726$*, and the covariances amount to* -0.00907 *and* $+0.06706$ *(such that they change sign as expected).*

There are two checks to make sure that the results are correct:

1) $P^L_{02} / P^P_{02} = Q^L_{02} / Q^P_{02} = 1.0706$*, and*

2) $D^{PL}_{02} = 1 \big/ D^{QP}_{02}$*, and* $D^{PP}_{02} = 1 \big/ D^{QL}_{02}$ ◀

3.5 The effect of changes in prices at different times (nonlinearity in prices)

Due to the linearity of measurement, any equal amount of change in the price of a commodity i denoted by $\Delta p_{it} = p_{it} - p_{i,t-1}$, or for short Δp_t in period t and $t^* \neq t$ (such that $\Delta p_t = \Delta p_{t^*}$) will have the same impact on the direct Laspeyres index as long as t or t^*, are periods within the range of an index with the same base 0 (that is $0 < t^*$, $0 < t$). This can easily be seen by

$$P^L_{02} = 1 + \frac{\sum q_0 \Delta p_1}{\sum q_0 p_0} + \frac{\sum q_0 \Delta p_2}{\sum q_0 p_0} = P^L_{01} + \frac{\sum q_0 \Delta p_2}{\sum q_0 p_0} \tag{3.5.1}$$

or in case of three periods

$$P^L_{03} = P^L_{02} + \frac{\sum q_0 \Delta p_3}{\sum q_0 p_0}. \tag{3.5.1a}$$

Hence a change in the prices of two commodities of say $\Delta p_{1t} = -2$ and $\Delta p_{2t} = +4$ weighted with constant quantities (6 and 4) has the same effect irrespective of the period in which the change takes place.

With weights $q_0 \big/ \sum p_0 q_0$ for the commodity 1 and 2 being 6/96 and 4/96 (the figures refer to ex. 3.2.1) we get

$$P^L_{01} - P^L_{00} = P^L_{01} - 1 = \frac{6}{96} \cdot (-2) + \frac{4}{96} \cdot (+4) = \frac{1}{24} = 0.04167$$

when the change takes place in period 1 (denoted by Δp_1 in eq. 3.5.1). When the same change also occurs in period 2 (such that $\Delta p_2 = \Delta p_1 = \Delta p$), and period 3 etc. we get

$$P^L_{02} - P^L_{00} = 0.04167 + 0.04167 = 0.08333$$

or

$$P^L_{02} = 1 + \left(2\Delta p \frac{q_0}{\sum q_0 p_0} \right) \quad \text{and} \quad P^L_{03} = 1 + \left(3\Delta p \frac{q_0}{\sum q_0 p_0} \right) \tag{3.5.2}$$

and so on. The situation is different, and much more complicated, however, in the case of a *chain* Laspeyres (price)[158] index (see tab. 3.5.1).

It is interesting to examine what happens in the case of a Laspeyres *chain* index, when prices change to the same extent in two consecutive periods (see ex. 3.5.1 for a numerical illustration). Obviously a constant (for all t) change by $\Delta p_{1t} = \Delta p_1 = -2$ and $\Delta p_{2t} = \Delta p_2 = +4$ will not have the same effect in $t = 1$, as it has in $t = 2$. Moreover

[158] The same applies mutatis mutandis in the case of a quantity index, which is perhaps more interesting due to its relation to the deflation problem.

the overall effect over two periods is (unlike eq. 3.5.2) not the same as a change by $2\Delta p_1$ and $2\Delta p_2$, but is rather given by

$$\begin{aligned}\bar{P}^{LC}_{02} &= \left(1+\frac{\sum q_0\Delta p}{\sum q_0 p_0}\right)\left(1+\frac{\sum q_1\Delta p}{\sum q_1 p_1}\right)\\ &= \bar{P}^{LC}_{01} + \left(\sum \Delta p \frac{q_1}{\sum q_1 p_1}\bar{P}^{LC}_{01}\right). \qquad (3.5.3)\end{aligned}$$

As the chain becomes longer, the relationship of course becomes more complicated. Consider possibly unequal changes of prices in periods 1, 2, and 3 denoted by $\Delta p_1, \Delta p_2$, and Δp_3. We then obtain

$$\begin{aligned}\bar{P}^{LC}_{03} &= \left(1+\frac{\sum q_0\Delta p_1}{\sum q_0 p_0}\right)\left(1+\frac{\sum q_1\Delta p_2}{\sum q_1 p_1}\right)\left(1+\frac{\sum q_2\Delta p_3}{\sum q_2 p_2}\right)\\ &= \bar{P}^{LC}_{02}\left(1+\frac{\sum q_2\Delta p_3}{\sum q_2 p_2}\right)\\ &= \bar{P}^{LC}_{02} + \left(\sum \Delta p_3 \frac{q_2}{\sum q_2 p_2}\bar{P}^{LC}_{02}\right). \qquad (3.5.4)\end{aligned}$$

The general relationship in the case of a chain index is (as shown in tab. 3.5.1) given by

$$\begin{aligned}\bar{P}^{LC}_{0t} &= \left(1+\frac{\sum q_0\Delta p_1}{\sum q_0 p_0}\right)\cdot\ldots\cdot\left(1+\frac{\sum q_{t-1}\Delta p_t}{\sum q_{t-1} p_t}\right)\\ &= \left(1+K^0_1\right)\cdot\ldots\cdot\left(1+K^{t-1}_t\right), \qquad (3.5.5)\end{aligned}$$

by contrast to the much simpler relationship, in the case of the corresponding direct index

$$P^L_{0t} = 1+\frac{\sum q_0\Delta p_1}{\sum q_0 p_0}+\cdots+\frac{\sum q_0\Delta p_t}{\sum q_0 p_0} = 1+K^0_1+\cdots+K^0_t. \qquad (3.5.6)$$

A comparison of eq. 3.5.5 with eq. 3.5.6 shows:

- In the case of P^L_{0t}, we have a linear combination of changes $\Delta p_1, \Delta p_2, \ldots$ with constant weights $q_0 / \sum p_0 q_0$, such that a constant term is added when the amount of change remains the same $\Delta p_1 = \Delta p_2 = \cdots = \Delta p_t$ in all periods. Thus the result is independent of how the interval (0, t) is subdivided into subintervals[159]
- By contrast in $\bar{P}^{LC}_{0t}$ changes of individual $\Delta p_{i1}, \Delta p_{i2}, \ldots$ are multiplied by weights which are in general not constant. Even the effect of a constant change ($\Delta p_1 = \Delta p_2 = \cdots = \Delta p_t$), will differ depending on when the same change takes place.

[159] For the direct Laspeyres index, it is for example irrelevant whether a single change by +10$ in one of the years from 1995 to 2000 takes place or a constant change of +2$ in each of the five one–year–intervals.

Table 3.5.1: Relation between changes of individual prices and the price index

t	direct Laspeyres	chain Laspeyres
1	$P^L_{01} = 1 + \sum \Delta p_1 \frac{q_0}{\sum q_0 p_0}$ $= 1 + K^0_1$	$\bar{P}^{LC}_{01} = 1 + \sum \Delta p_1 \frac{q_0}{\sum q_0 p_0}$ $= P^L_{01} = 1 + K^0_1$
2	$P^L_{02} = P^L_{01} + \sum \Delta p_2 \frac{q_0}{\sum q_0 p_0}$ $= 1 + K^0_1 + K^0_2$ $= P^L_{01} + K^0_2$	$\bar{P}^{LC}_{02} = \bar{P}^{LC}_{01} \left(1 + \sum \Delta p_2 \frac{q_1}{\sum q_1 p_1}\right)$ $= \left(1 + K^0_1\right)\left(1 + K^1_2\right)$ $= \bar{P}^{LC}_{01}\left(1 + K^1_2\right)$
3	$P^L_{03} = P^L_{02} + \sum \Delta p_3 \frac{q_0}{\sum q_0 p_0}$ $= 1 + K^0_1 + K^0_2 + K^0_3$ $= P^L_{02} + K^0_3$	$\bar{P}^{LC}_{03} = \bar{P}^{LC}_{02} \left(1 + \sum \Delta p_3 \frac{q_2}{\sum q_2 p_2}\right)$ $= \left(1 + K^0_1\right)\left(1 + K^1_2\right)\left(1 + K^2_3\right)$ $= \bar{P}^{LC}_{02}\left(1 + K^2_3\right)$

In general we may conclude that:

> The direct Laspeyres price index is linear in the prices: the absolute change of the index is uniquely determined by the absolute changes in individual prices, and the result is invariant with respect to the kind of temporal aggregation. chain indices lack these properties.

The absolute change of $\bar{P}^{LC}_{0t}$ to $\bar{P}^{LC}_{0,t+1}$ according to eqs. 3.5.3 and 3.5.4 is

$$\Delta \bar{P}^{LC}_{0,t+1} = \bar{P}^{LC}_{0,t+1} - \bar{P}^{LC}_{0t} = \sum_i \Delta p_{i,t+1} \frac{q_{it}}{\sum q_{it} p_{it}} \bar{P}^{LC}_{0t}, \tag{3.5.7}$$

and this term which is to be added to $\bar{P}^{LC}_{0t}$ in order to get $\bar{P}^{LC}_{0,t+1}$ is not only much more difficult to understand, and to interpret, than the equivalent term added to P^L_{0t} to get $P^L_{0,t+1}$, that is

$$\Delta P^L_{0,t+1} = P^L_{0,t+1} - P^L_{0,t} = \sum_i \Delta p_{i,t+1} \left(\frac{q_{i0}}{\sum q_{i0} p_{i0}}\right) = \sum_i \Delta p_{it} w_i \tag{3.5.8}$$

it also depends on $\bar{P}^{LC}_{0t}$ while in eq. 3.5.8 the change does not depend on P^L_{0t}. There are important differences in two respects.

1. There are constant weights w_i, as a consequence of P^L_{0t} being linear in the prices (by which a change in the price index ΔP^L_{0t} is related to individual changes in prices), whereas the terms to be added in the case of a chain index of Laspeyres will vary, even in case of constant changes in prices (such that $\Delta p_1 = \Delta p_2 = \cdots = \Delta p$ in all periods) unless all quantities remain constant such that $q_0 = q_1 = \cdots = q_{t-1}$, and even more important,

2. the changing weights given to individual changes in prices also depend on the level the chain index has reached so far (as the term $\bar{P}^{LC}_{0t}$ appears in eq. 3.5.7).

We may of course doubt that an index should account for a constant change of prices by simply adding constant terms $\sum \Delta p \left(\frac{q_0}{\sum p_0 q_0} \right)$, and argue that the same change of prices may not always have the same significance. The problem, however, is: there is no theory known to determine the different weights to be assigned to the change of prices depending on the time they take place, and to explain why these weights

- should be related to current and past quantities, and depend on
- the "history" of the chain index, and on how the interval (0, t) is subdivided[160]

in exactly the same way as it is done in eq. 3.5.8.

> The question is not to justify why the same increase in prices should be treated differently depending on when the increase occurs. The question is rather by which economic theory it is justified to do this precisely in the specific way of temporal aggregation (summation over time intervals) provided by chaining.

In eq. 3.5.8 successive changes of prices Δp_{it}, are not only multiplied by different factors $q_t / \sum p_t q_t$, (instead of $q_0 / \sum p_0 q_0$ in P^L_{0t}), but also by $\bar{P}^{LC}_{0t}$ reflecting the history of the chain. This variation of weights implicitly attached to price differences Δp, in the case of chaining does not seem to be well understood. It is possibly rather a by–product of the typical chain index type of temporal aggregation.

Example 3.5.1

The following numerical example is again a modification of ex. 3.2.1, and the salient feature is that the price of commodity 1 uniformly (in both periods 1 compared with 0 and 2 compared with 1) declines by -2 and the price of commodity 2 constantly rises by +4.

	period 0		*period 1*		*period 2*	
i	*prices*	*quantities*	*prices*	*quantities*	*prices*	*quantities*
1	*8*	*6*	*6*	*10*	*4*	
2	*12*	*4*	*16*	*3*	*20*	

[160] This is a property of chain indices which definitely was *not* intended in the 19th century discussion of transitivity in index theory. The problem discussed here is not lack of transitivity, that is inconsistency in aggregation over (sub–) intervals ("temporal aggregation"), but the theoretical foundation of that type of temporal aggregation which underlies the procedure of chaining.

The direct Laspeyres index yields 1.0417 and 1.0833 rising constantly by 4.17%, because the weights of the prices given by (note that $p_{10}q_{10} = p_{20}q_{20} = 48$*)*

$$\frac{q_{10}}{\sum q_{i0}p_{i0}} = \frac{6}{96} \quad \text{and} \quad \frac{q_{20}}{\sum q_{i0}p_{i0}} = \frac{4}{96}$$

are the same in both periods such that

$$\Delta p_1 \frac{q_{10}}{\sum q_{i0}p_{i0}} + \Delta p_2 \frac{q_{20}}{\sum q_{i0}p_{i0}} = (-2)\cdot\frac{6}{96} + (+4)\cdot\frac{4}{96} = 0.0417.$$

Hence the direct index treats the overall effect of a uniform absolute change of +4 and -2 respectively in both periods equally as a rise of prices by 4.17%, resulting in $P^L_{01} = 1.0417$ *and* $P^L_{02} = 1.0833$.
The chain index, on the other hand, develops in the following way:

		weights given in $\bar{P}^{LC}_{0t} - \bar{P}^{LC}_{0,t-1}$ *to**	
t	$\bar{P}^{LC}_{0t}$	*commodity 1*	*commodity 2*
1	$\frac{100}{96} = 1.0417$	$\frac{6}{96} = 0.0625$	$\frac{4}{96} = 0.0417$
2	$\frac{100}{96}\cdot\frac{100}{108} = 0.9645$	$\frac{10}{108}\cdot\frac{100}{96} = 0.09645$	$\frac{3}{108}\cdot\frac{100}{96} = 0.02894$

* *What follows are the terms* $\Delta p_{it} \cdot \left(q_{i,t-1} / \sum p_{i,t-1} q_{i,t-1} \; \bar{P}^{LC}_{0,t-1}\right)$ *for* $i = 1, 2$.

$\bar{P}^{LC}_{02} = 0.9645$ *indicates a* decline *in prices as opposed to a rise in* $P^L_{02} = 1.0833$. *This is due to the fact that the weight of commodity 1 (becoming cheaper) rose by 54.4% (from 0.0625 to 0.09645), whereas the weight of commodity 2 (becoming more expensive) declined by 30.55%.* $\bar{P}^{LC}_{02}$ *differs from* $\bar{P}^{LC}_{01}$ *by* $-0.07716 = 0.09645\cdot(-2) + 0.02894\cdot(+4)$*, derived from the fact that* $\bar{P}^{LC}_{01} > 1$ *is outweighed by the considerable increase in the quantity of commodity 1, and the reduction in the quantity of commodity 2 respectively. Assume for example the quantity of commodity 1 in period 1 is 6 (as in period 0) instead of 10, the other figures remaining unchanged. We would get:* $\bar{P}^{LC}_{02} = \bar{P}^{LC}_{01} = 1.0417$ *(no change in period 2).* ◀

It should be noted that here we argue in terms of (absolute) *differences* and not in terms of *ratios*. We might prefer P^{LC}_2 to P^L_{02}/P^L_{01} as representing a "better" growth factor, because of the more "adequate" weights q_1 (instead of q_0). Furthermore the link is independent of the history of the chain. Path dependence and the cumulative nature of implicit weights is not yet visible when only a link is considered. There is not only a marked difference between the chain index $\bar{P}^{LC}_{0t}$ and P^L_{0t}, but also between $\bar{P}^{LC}_{0t}$ and P^P_{0t}. By analogy to eq. 3.5.1 for the direct Paasche index we get

$$P^P_{02} = 1 + \frac{\sum q_2 \Delta p_1}{\sum q_2 p_0} + \frac{\sum q_2 \Delta p_2}{\sum q_2 p_0} = 1 + K^2_1 + K^2_2, \tag{3.5.9}$$

and in the case of three periods

$$P^P_{03} = 1 + K^3_1 + K^3_2 + K^3_3, \tag{3.5.9a}$$

such that the weights attached to the same differences in prices $\Delta p_1 = \Delta p_2 = \cdots = \Delta p$ are constantly redefined with the passage of time (variation of t), but nonetheless the same change will not be treated differently, depending on when it takes place. Moreover we have

$$\begin{aligned} \Delta P_{03}^{P} = P_{03}^{P} - P_{02}^{P} &= \sum \Delta p_{i1} g_i + \sum \Delta p_{i2} g_i + \sum \Delta p_{i3} \frac{q_{i3}}{\sum q_{i3} p_{i0}} \\ &= \sum \left(\Delta p_{i1} + \Delta p_{i2}\right) g_i + \frac{\sum q_3 \Delta p_3}{\sum q_3 p_0} \end{aligned} \tag{3.5.10}$$

where

$$g_i = \frac{q_3}{\sum q_3 p_0} - \frac{q_2}{\sum q_2 p_0},$$

such that we again have a function linear in the price changes.

3.6 Numerical examples, simulations and estimations with real data

Numerical *examples* are in general useful only to demonstrate certain properties of chain indices. In this sense we also make use of them here. The problem with them is usually, that they can easily be blamed for being tendentious, and that sometimes minor modifications may suffice to reach quite different conclusions.
Experience with *simulations* (for reports in literature see for example Szulc) shows that a substantial drift is to be expected when prices and quantities move cyclically. The results depend a lot on the length and temporal distribution of subintervals to be linked. The record of chain indices is, not surprisingly, also much different, depending on which type of direct index is used to make a comparison. It is likely that chain indices fare better when compared with direct indices, infrequently updated and they perform much less impressively when compared with direct indices more frequently updated.
In his textbook ALLEN (1975) quoted an *empirical study* of Fowler concerning the U.K. Retail Price Index (RPI), and he pointed out that the chain index proved its superiority, mainly because in most cases its result remained within the span between a direct Paasche and a direct Laspeyres index, as required by economic theory of index numbers[161]. Thus in Allen's eyes this finding was enough supporting evidence, for his enthusiasm about chain indices, although a direct index representing a mean of

[161] Note that a chain Laspeyres index may well fall short of a direct Paasche index as the following variant of ex. 3.5.1 demonstrates: prices change again by -2 and +4 each period (such that $\Delta p_1 = \Delta p_2 = \cdots = \Delta p_t$) and the quantity of commodity 1 (the price of which is decreasing by -2 in each period) rises by 1 in each period:

i	p_0	q_0	p_1	q_1	p_2	q_2	p_3	q_3
1	8	6	6	7	4	8	2	9
2	12	3	16	3	20	3	24	3

P^P_{0t} and P^L_{0t} like Fisher's index for example would have done the same job. Surprisingly the difference between the average annual growth rate as measured by the direct versions of Laspeyres and Paasche on the one hand, and their chain versions on the other hand was rather small. The result observed for the majority of the time series of Fowler as quoted by Allen was $P^P_{0t} < \bar{P}^{PC}_{0t} < \bar{P}^{LC}_{0t} < P^L_{0t}$. But the interval was small though sizeable compared, with studies performed some years later (for example for German data).

Statistical offices of some countries used to publish a "makeshift" (or approximated) Paasche index in addition to the official Laspeyres index (for consumer prices for example), on a more or less regular basis in order to demonstrate the size of divergence between these two direct indices, and to thereby establish the frequency at which a fixed basket should be replaced by a new one. Experience shows that divergence is smaller, and thus a more frequent updating of weights appears less urgent than commonly expected.

There are many objections that can be made against this type of analysis:

1. An "experimental" chain index or direct Paasche index calculated for the purpose of comparison with the direct Laspeyres index is often compiled on the basis of a rough disaggregation only (consumption broken down into some fifty headings). Thus such an index is unable to cover some "low level substitution" which might produce more divergence between a chain index, and its direct index counterpart, if substitution were accounted for correctly. Moreover the updating of the consumption pattern is performed by using National Account data instead of detailed Family Budget Surveys[162], such that conceptual differences might exist and a lot of the outlet or quality bias is not adjusted for.

2. Comparisons are often made for a couple of years in which inflation was moderate, such that the period in question was too short and the price increase too small to reveal enough Laspeyres–"overstatement".

In the face of some more detailed empirical investigations the first argument seems to be hardly persuasive. A German study with consumption broken down into more than 200 positions (SCHMIDT (1995)) has revealed that divergence occurs in either directions, (that is P^L_{0t} overstates *and* understates inflation with respect to certain sub–aggregates), but was mostly astoundingly small.

Another study proved that a direct Laspeyres price index P^L_{0t} with constant weights of a more recent period than the (reference) base 0, does not necessarily display less inflation as usually conjectured (HOFFMANN (1998)). It is also a common prejudice that P^L_{0t} will drift progressively upwards with the passage of time.

For the direct Laspeyres index we get $P^L_{01} = P^L_{02} = P^L_{03} = 1$ which means a *constant* price level. $\bar{P}^{LC}_{03} = 0.9353$ while $P^P_{03} = 0.9375$.

[162] Note that some countries applying officially the chain approach for many years already, like France, also use National Accounts to update the "basket" of the CPI.

It turned out, however, that the direct index did not necessarily drift progressively away from the chain index, simply because the longer the interval under consideration the more likely chain indices also suffer from cyclical price movements in some parts of the budget.
The importance of the overall inflation rate for the possible difference between a direct and a chain measurement of inflation should also not be exaggerated. Rapid inflation more often occurs in countries with a low level of living and generally bad conditions for consumption under which we do not have much room for the structure of consumption to change rapidly and fundamentally. Thus there is no reason to expect much difference between a chain and a direct index.
Under "normal" conditions we usually have surprisingly small drifts only, at least when the direct index is not updated too infrequently. This gives rise to different conclusions:

- the numerical effect of some shortcomings of chain indices, such as path dependency or non–additivity will also be less dramatic, such that the chain approach is justified, or

- given the unexpected small divergence between chain and direct indices, it is all the more questionable to incur the considerable extra burden of more frequent budget surveys (chain indices are no doubt the more expensive methodology).

Finally it is noteworthy that studies of the size and direction of the temporal correlation responsible for $P_{0t}^{P} \neq P_{0t}^{L}$, and for the drift $\bar{P}_{0t}^{LC} \neq P_{0t}^{L}$ (see sec. 3.4) were also made. NEUBAUER (1995) for example only found[163] a rather small correlation between price and quantity relatives, and he found out that this correlation differed slightly only in periods of low inflation (correlation +0.092) from periods of high inflation (-0.027). Hence we do not have much reason to expect fundamentally different results, when direct indices are replaced by chain indices, not even in times of higher inflation rates.

3.7 Accuracy and cumulation of errors

The assertion that chain indices are able to provide more accurate measures of price movement, because of comparing prices of two adjacent periods only is frequently found in literature. Problems with imputed prices in cases of disappearance of commodities and the like are supposed to occur less frequently. When direct comparison of prices over a long period in time is aimed at we might, more often face the problem that after some years comparable goods and services are possibly no longer available. Hence there *seems* to be some good reasons to expect that accurate price reports will

[163] Interestingly in this study Neubauer also found out that the difference between the official chain index for the U.K. 1988–1994 and an approximated direct Laspeyres index (estimated by himself) was small in size and varying in direction, showing that a constant basket (1987 = 100) even after five years (in 1992 and later) not necessarily yields an upward bias.

be more easily achieved for a link as for a constant basket. But this is only one aspect of accuracy to be considered when chain indices are compared with direct indices in terms of “accuracy”, and it gives rise to a more systematic examination of errors imparted to chain indices by the way they are compiled.

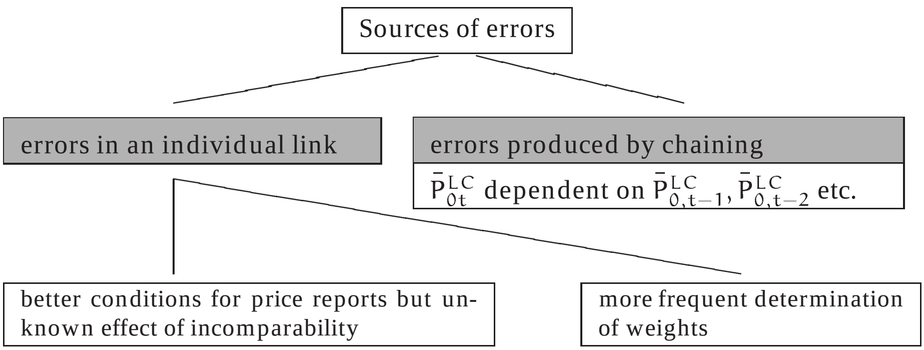
a) Types and sources of errors

To look at the conditions of making price reports is certainly not enough when direct and chain indices are compared with respect to “accuracy”. Thus in fig. 3.7.1 an attempt is made to enumerate all elements of chain index compilation possibly liable to errors. In particular it should also be taken into account that chain indices, need a more frequent empirical establishment of weights (quantities or expenditure shares), which is costly, laborious and simply due to its complexity most likely liable to errors. There are mainly two sources of bias:

1. the more frequent determination of weights, and
2. the multiplication of links to form a chain by which the result $\bar{P}^{LC}_{0t}$ is not independent of $\bar{P}^{LC}_{0,t-1}$, $\bar{P}^{LC}_{0,t-2}$ etc.,

both specific for chain indices, and hence not encountered in the direct Laspeyres index framework.

Figure 3.7.1: Sources of bias (systematic error) in compilation of chain indices



In what follows we consider the two aspects listed above separately. We will ask in particular, what the possible effect of errors in the weights is and how multiplication of links, each of them most likely subject to some kind of inaccuracy, could redound to an overall error (or bias) in the product.

Conspiciously “chainers” mostly stress the idea of “better conditions for price reports” only, but they tend to ignore completely the other two aspects displayed in fig. 3.7.1.

Furthermore it should be noted that even this argument for better conditions, prevailing in the practice of price statistics for the collection of price reports is *not* beyond doubt.
On the face of it, there is no good reason why the right hand side (RHS) of the equation

$$\frac{p_{it}}{p_{i0}} = \frac{p_{i1}}{p_{i0}} \frac{p_{i2}}{p_{i1}} \ldots \frac{p_{it}}{p_{i,t-1}} \tag{3.7.1}$$

or "translated" into the aggregate context

$$\bar{P}_{0t} = P_{01}P_{12}\ldots P_{t-1,t} = P_1^C P_2^C \ldots P_t^C \tag{3.7.1a}$$

should be any more "accurate" than the left hand side (LHS), unless we think of aspects, like

- disappearance of goods and services and the emergence of entirely new, perhaps even incomparable items, such that strictly speaking "matching" is not possible, or
- lack of "overlap", nonavailability of equivalent conditions, and in case of aggregates
- varying "weights" on the RHS of eq. 3.7.1a.

Under such conditions the calculation in *steps* (RHS–procedure) does not produce a more "accurate" result than the *single* step LHS–procedure. It is also necessary to reflect problems of changes in the domain of definition (see sec. 3.2b), and to question whether the RHS result can really be equated to the LHS result.

> The focus should not only be on conditions for price reports. We should also take into account inaccuracies, and lack of comparability resulting from changes in the selection of goods, outlets and the like, let alone possible errors due to a more frequent updating of weights.

It is unlikely that chaining will offer both advantages in reporting prices simultaneously: more accuracy and fewer problems with imputations and adjustments (for example in the case of quality improvement), made explicitly in order to ensure comparability.
Strictly speaking in eq. 3.7.1 it is assumed that the i–th commodity remains the *same* throughout the interval from 0 to t. But on the other hand chain indices are said to be superior, as

- it is no longer necessary to pay attention to comparability over a *long* time interval, *and*
- they are less susceptible to errors.

The first point is in fact a strong argument in favor of chain indices, but sometimes not in the sense of a *superior* solution deliberately chosen, but rather made in compulsion, or as a makeshift solution when comparability is difficult to establish though still deemed desirable.

It is for example interesting to see why Germany during the "Third Reich" in October 1939 moved to a chain index. According to ANDERSON (1949a), p. 479f this was primarily done because – as a consequence of war – more and more goods of the "basket" (in these days related to a five–person employee household) were no longer available, rationed or otherwise incomparable.[164] This German *chain* consumer price index was finally in March 1945 a product of no less than 66 factors (!), obviously subject to a bias (systematic error). Anderson estimated ("highly optimistically"), that the bias of each monthly (!) link of this German "Reichsindexziffer der Lebenshaltungskosten" amounting to some 0.5%, and thus he derived an accumulated overall error of 30% or so by 1946.[165] The new chain index which then was created on behest of the Allied Forces in April 1946, was again a monthly chained index and the index was due to an additional factor severely biased: prices on the black market were excluded.[166]

To start with an analysis of errors, it is necessary to introduce some conventional definitions. In

$$P_t^C = P_t^{C^*}\left(1 - \frac{a_t}{P_t^{C^*}}\right) = P_t^{C^*}\left(1 + e_t\right), \tag{3.7.2}$$

where P^* denotes the observed link as opposed to the true link P, the terms a and e are known as absolute and relative error respectively. At first glance the conventions concerning the definition, and the sign of the relative error e, as listed in fig. 3.7.2 seem to contradict intuition.

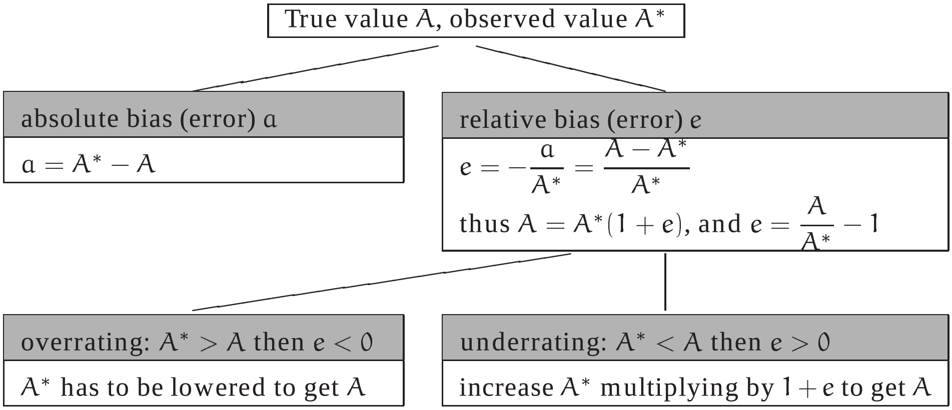
Note that relative errors are defined making use of observed, not of true values, and that for example $e = +0.3 > 0$ means the true value $A = A^*(1 + 0.3)$ is exceeding the observed one by 30%, and therefore the observed value A^* has be augmented by 30% (per cent of A^*).

164 At the beginning some attempts were made to make substitutions (in a *fixed* basket set–up) such that the nutritive value of the basket was kept constant. But since wartime conditions continued to thwart all attempts to establish a basket, in which the goods and services could consistently be priced for more than just the next few months, the fixed basket approach was finally given up in favor of a chain index. It was simply no longer *possible* to define a basket which could be kept constant.

165 Note that this estimation did not yet account for problems with cyclical price movement, in case of a sub–annual chaining (see sec. 3.2f and 3.4d). Anderson's estimation was based solely on the theory of accumulation of errors due to multiplication (as briefly recorded and summarized below).

166 It is interesting to note in passing, that a statistician of Germany's official statistics (quoted in ANDERSON (1949b), p. 177f.), tried to reconcile chain indices (in these days unfamiliar in Germany) with the principle of pure price comparison arguing that the *same* basket is used in the numerator, and in the denominator of the index. This argument which applies to the link only, not to the chain, is quite popular among adherents of chain indices.

Figure 3.7.2: Definition of bias (systematic error)



b) The exaggerated impact of errors in expenditure weights

It is conspicuous that "chainers" argue in terms of "accuracy" with respect to prices, not to weights while according to eq. 1.2.1 and 1.2.2 the chain index $\bar{P}_{0t}^{LC}$ differs from P_{0t}^{L}, not with respect to prices but rather with respect to weights (or quantities $q_1, q_2, \ldots, q_{t-1}$). Interestingly stress is also laid on the significance of new (expenditure) weights whereas errors in determining such weights seem to be neglected. The impact of such errors, typically occurring more frequently in the chain index framework was examined in detail by STRECKER (1963).[167] His main finding was:

> Errors in establishing *weights* do not tend to considerably affect the overall result of an index (also influenced by *prices*). The other side of the coin is that the impact of alternative (updated) weights is much smaller than customarily assumed.

Let e_i denote the relative "systematic" error[168] in establishing expenditure weights such that the "true" expenditure $(p_i q_t)$ for commodity i is given by

$$(p_{i0} q_{i0}) = (p_{i0} q_{i0})^* \, (1 + e_{i0}) \, ,$$

[167] Strecker referred to the "normal" direct Laspeyres index, not explicitly to the chain index variant. But of course his findings apply to chain indices as well.

[168] Textbooks in statistics commonly make a distinction between "random errors" and "systematic errors" (or "biases") respectively. Characteristic for the first type of errors, is that probability theory applies, and the errors will cancel out "in the long run" as they disperse in *both* directions, underrating as well as overrating the "true" figure. A "systematic error" by contrast diverges in one direction only, and there is no chance that such errors will diminish as the number of observations increases. The considerations set out above will apply exclusively to "systematic" errors while observations in price statistics will be subject to both types of errors of course. To focus on systematic errors is justified because they will "persist" as a factor in the chain, and they cannot, as a rule, be reduced by an increase of the number of observations.

or for short

$$g_{i0} = g^*_{i0}\,(1 + e_{i0})\,, \quad \text{thus} \quad e_{i0} = \frac{p_{i0}q_{i0}}{(p_{i0}q_{i0})^*} - 1,$$

where $(p_i q_i)^*$ is the possibly inaccurately "observed" value. Strecker examined the influence of errors in expenditure weights of the base period as they might appear in a direct Laspeyres price index (the same would apply to a Laspeyres link). The "*true*" index is given by (omitting the subscript i for convenience of presentation)

$$P^L_{0t} = \sum \frac{p_t}{q_0} \frac{p_0 q_0}{\sum p_0 q_0} = \sum a_t \frac{g_0}{\sum g_0},$$

whereas the *observed* index (with errors in the weights $g_0 / \sum g_0$) is approximately equal to

$$\sum \frac{p^*_t}{p^*_0} \frac{(p_0 q_0)^*(1 + e_0)}{\sum (p_0 q_0)^*(1 + e_0)} = \sum a^*_t \frac{g^*_0(1 + e_0)}{\sum g^*_0(1 + e_0)} = (1 - E_0)\frac{\sum a^*_t g^*_0}{\sum g^*_0}.$$

The price relatives are assumed to be observed correctly, hence $a^*_t = a_t$. The function of the overall error in establishing expenditure weights $E_0 = E$ derived by Strecker is given by

$$E = \frac{\sum g^*_0 a_t e_0}{\sum g^*_0 a_t} - \frac{\sum g^*_0 e_0}{\sum g^*_0} = E_1 - E_2. \tag{3.7.3}$$

Hence the relative systematic *overall error of weights* E depends on the observed individual expenditures (g^*_0) and their errors (e_0), as well as on the price relatives (a_t), here assumed to be without error. For each commodity i the observed base period expenditure g_{i0}, has a relative error e_{i0} which enters the formulas of two differently weighted arithmetic means, E_1 and E_2 respectively, and

> The overall error (bias) E of an index (link) resulting from errors e_{i0} in expenditure weights (or from taking "wrong" weights), is determined by the *difference* between two differently weighted means of these errors. Hence the overall bias E will most probably be small even though the individual errors e_0, might be substantial in size.

In two cases eq. 3.7.3 yields $E = 0$, that is a vanishing overall error in establishing weights of a direct Laspeyres index despite non–negative individual errors e_{i0} in expenditures (g^*_0):

1. when all errors (biases) are equal $e_{10} = e_{20} = \cdots = e_{n0}$, such that $e_{i0} = e_0$ and / or

2. all price relatives (assumed to be observed without error) happen to be equal $a_{it} = a_t$.

The first case is of course highly unrealistic, since it is unlikely that all index positions will be underrated or overrated to the same extent. It is more likely that some expenditures are overrated and others underrated. The second case is again unrealistic and irrelevant, because under such conditions different index formulas should display the same result anyway.

To demonstrate that errors in weights will make an astoundingly *small* contribution to the overall bias, Strecker considered a numerical example assuming five commodities with relative errors in expenditures ranging from -20% to +20%, such that the total expenditure observed in the base period falls short of the true value by 6.3%. Yet the total error (bias) due to inaccuracy in establishing weights turned out as -2.7%, which is rather moderate compared with individual errors amounting to -20% up to +20%.

There is even more reason to expect such a result, (small E despite sizable individual errors e), as we take into account that in practice a price index is made up of many more items, several hundreds in general, not only just five commodities and that relative errors of 20% or so, in the weights of items are highly unlikely. Ex. 3.7.1 intends to demonstrate that the impact of errors in the weights is small, even under more dramatic assumptions than those Strecker made.

Example 3.7.1

Assume two commodities with observed base period expenditures of 500 each and errors e_i as well as price relatives a_i as follows

	g^*_{i0}	e_i	a_i
1	500	+0.5	1.5
2	500	-0.5	0.5

We get

$$E_1 = \frac{500 \cdot 1.5 \cdot 0.5 + 500 \cdot 0.5 \cdot (-0.5)}{500 \cdot 1.5 + 500 \cdot 0.5} = 0.25$$

and $E_2 = 0$ since $e_1 = -e_2$. Therefore the overall error $E = 25\%$ is only half of each absolute individual error. Interestingly upon interchanging the e–terms we simply get -25% instead of +25%. ◄

To verify eq. 3.7.3 we proceed as follows: According to the errors e_0 (+0.5 and -0.5) the true expenditures are $g_{i0} = g^*_{i0}(1 + e_{i0})$, amounting to 750 and 250 respectively[169]. With the observed expenditures g^*_{i0} we get the *biased* index

$$P^* = \sum \frac{g^*_{i0}}{\sum g^*_{i0}} a_{it} = \frac{1}{2} \cdot 1.5 + \frac{1}{2} \cdot 0.5 = 1$$

[169] This means that expenditure of commodity 1 is underrated ($e > 0$) showing 500 instead of 750 and the opposite is true for commodity 2.

by contrast to

$$P = \sum \frac{g_{i0}}{\sum g_{i0}} a_{it} = \frac{750 \cdot 1.5 + 250 \cdot 0.5}{1000} = 1.25,$$

the "true" index (with "true" expenditures).
Since $P = (1 + E)P^*$ we have $E = 0.25$. Thus the observed index underrates the true increase by 25%. The situation would be even less impressive if we substituted more realistic price relatives 1.2 and 0.8, for 1.5 and 0.5 maintaining the original figures for expenditures and errors. The result then is +0.1 instead of +0.25. Thus we might safely assume that E will be small, under "normal" conditions in practice.
The result of such demonstrations is noteworthy in so far as chain indices are often preferred to direct indices, because they replace "obsolete" (we could also say "biased") weights by new (or unbiased) weights. But eq. 3.7.3 proves that under realistic conditions the impact of different weighting schemes (or errors in one scheme) is surprisingly small, in contrast to some widespread prejudice. "Statisticians know better than the public that small changes in weights usually have little effect upon an index" (TURVEY (1989), p. 38).[170]
However small errors in weights might be empirically, it should be acknowledged that

1. chain indices require a significantly *more frequent* update of weights than direct indices, and

2. biases in weights or any other errors in links are likely to *accumulate* in the chain gained by multiplying such links.

The relevance of the latter aspect will be examined now.

c) Cumulation of errors by multiplication of links

Cumulation of errors is an aspect sometimes mentioned by non–chainers[171]. With biased links according to eq. 3.7.2 we have upon multiplication

$$\bar{P}_{02}^{C^*} = P_1^{C^*} P_2^{C^*} = P_1^C P_2^C (1 + e_1)(1 + e_2) \approx P_1^C P_2^C (1 + e_1 + e_2) \qquad (3.7.4)$$

where the product is negligible $e_1 e_2 = 0$. In the same way we get

$$\begin{aligned}(1 + e_1)(1 + e_2)(1 + e_3) &= 1 + e_1 + e_2 + e_3 + e_1 e_2 + e_1 e_3 + e_2 e_3 + e_1 e_2 e_3 \\ &\approx 1 + e_1 + e_2 + e_3, \qquad (3.7.4a)\end{aligned}$$

[170] Interestingly we find a similar result (relatively small impact of weights) in Allen's textbook (ALLEN (1975)), though derived from quite a different approach (examination of the sampling distribution of prices and quantities).

[171] See ANDERSON (1949b), (1949a), but also SCHMIDT (1995), p. 61 who is not a non–chainer to the same extent as in fact Anderson is.

or in general

$$\prod_{\tau=1}^{\tau=t} P_{\tau}^{C^*} \approx \prod_{\tau=1}^{\tau=t} P_{\tau}^{C} \left(1 + \sum_{\tau=1}^{\tau=t} e_{\tau}\right). \tag{3.7.5}$$

It is not unlikely that errors made in compiling a link consistently tend into the same direction, such that the error of the chain is *cumulated* (*summated*). A systematic error of a link amounting to 0,011 (or 1.1%), for example year after year would sum up to 11% after 10 years (in absolute terms). No such summation occurs when previous year biased figures are not used to form a new estimate of a direct index P_{0t}. The situation is also entirely different, in the case of dividing one statistic by another, both statistics being liable to a bias of the same sign, as for example in determining the growth factor of a direct Laspeyres index:

$$\frac{P_{0t}^{L^*}}{P_{0,t-1}^{L^*}} = \frac{P_{0t}^{L}(1+e_t)}{P_{0,t-1}^{L}(1+e_{t-1})} \approx \frac{P_{0t}^{L}}{P_{0,t-1}^{L}}(1+e_t-e_{t-1}). \tag{3.7.6}$$

To derive the estimate

$$\frac{1+e_t}{1+e_{t-1}} \approx 1+e_t-e_{t-1}$$

the Taylor series

$$\frac{1}{1+e_{t-1}} = 1-e_{t-1}+e_{t-1}^2-e_{t-1}^3+e_{t-1}^4-\ldots$$

is used. Thus we arrive at the following statement broadly in line with our common sense expectations:

> When there are good reasons to expect that a bias of the same sign, and approximately of the same absolute size occurs in successive periods, we can safely assume that the situation will aggravate in the case of multiplication (as opposed to division where errors tend to cancel out).

Hence *multiplication* of links is unfavorable, as far as the bias problem is concerned, whereas focusing on growth rates or (equivalently) on links taken in isolation is favorable. Thus it is certainly not tenable, to assert on the one hand a greater accuracy in the chain index approach, due to the focus on two successive periods only, (to be interpreted as focusing on growth rates only where the bias in t and $t-1$ approximately equals $e_t \approx e_{t-1}$), and to ignore on the other hand, that such links are multiplied to form a chain (producing a term $\sum e_t$).
It should be borne in mind that in the chain approach an error in $\bar{P}_{0t}^{C}$, is carried forward when $\bar{P}_{0,t+1}^{C}, \bar{P}_{0,t+2}^{C}$ and so on are calculated. There is no counterpart to this type of error in the case of a direct index. The standard argument against the direct

Laspeyres index P_{0t}^{L}, is that this index is biased to an increasing extent in the course of time, because the basket is fixed, but it is not unlikely that this also applies to the chain index $\bar{P}_{0t}^{LC}$, due to the inherent cumulation of errors in the links, let alone other phenomena like path dependency and cyclical price movement that might contribute to implausible results. There is no reason why such errors should not exist, or why they should change sign.
As to the problem of bias there is also no reason why a chain index should fare better, precisely in those cases in which a direct Laspeyres index might fail, that is in *long* term comparisons.

4 Elements of a theory of chain indices

Proponents of chaining succeeded in creating the impression that chain indices are a more modern approach to index numbers (which is simply not true)[172], and more suitable for a modern dynamic economy, than traditional "fixed base" direct indices. They are often found more adequate to deal with rapid changes, because of constantly adapting weights. In addition, it is said that only chain indices can make situations comparable, that are otherwise incomparable by linking partial comparisons with a certain overlap.

The purpose of this chapter is to discuss some problems of what might be called a "theory of chaining" (of which, however, not much has been seen by now). It will be shown that

- we should make a distinction between "*chaining*", that is making long term comparisons by multiplying links to a chain and "*chainability*" (also called circularity, or transitivity etc.), a condition to make such comparisons consistent with direct comparisons[173],

- chain indices are "*path dependent*" (which is the opposite of chainability), and hence *not* in keeping with what originally was found desirable when chaining was recommended;

- chainability is most restrictive and apparently fits much more to the task of comparing countries (where no "sequence" is underlying) than to make temporal comparisons;

- furthermore the rationale of chaining is difficult to reconcile with proportionality assumptions inherent in chainability, and is necessarily in conflict with varying weights[174] and consistency requirements in aggregation; and finally

- it will be shown that latest weights, are not necessarily the most appropriate ones, and if so chaining is not the only method to account for such updated weights.

---

[172] Chain indices and direct indices both have a tradition of almost equal length: when systematic research in index numbers began, the chain approach was found no less an interesting and promising approach than the direct approach.

[173] Without such a distinction the name "chain index" is highly misleading in that it panders to the mistake that chain indices would satisfy chainability (transitivity). As will be shown in sec. 4.1 chain indices make use of "chaining" (the operation of multiplication), but they do not satisfy chainability (a property of an index function).

[174] Hence the two pretended major advantages of chain indices, that is making consistent long term comparisons by chaining and adaptability of weights are not to be reconciled.

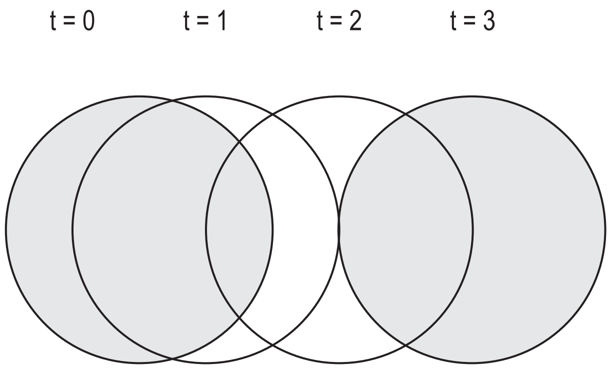
4.1 Chaining and chainability (transitivity)

This section aims at a better understanding of "circularity" or "transitivity", a notion of some importance for justifying chain indices. It will be shown that circularity is highly restrictive and introduces an element of inconsistency when combined with other criteria.

a) The theory of "comparability" of Marco Martini

According to Marco Martini chain indices are superior to direct indices, due to their ability to make situations comparable that are otherwise incomparable, by linking partial comparisons with a certain overlap, an idea made intuitively clear with the help of the picture shown in fig. 4.1.1, used by Martini to persuade Eurostat of chain indices (MARTINI (1996), (2000)).
Note that the first circle on the left hand side and the last on the right hand side have virtually no overlap at all (hence they may be called "incomparable"). But as linking proceeds from left to right they are – according to "chainers" at least – made comparable, because each link (pair of circles) has a certain overlap. Hence chaining seems to be an elegant method in situations which cannot be managed by direct comparisons.

Figure 4.1.1: Successive comparisons of partially overlapping circles



The centerpiece of Martini's theory of comparability (and also his preference of chain indices) is the idea that for two situations to be comparable they have to fulfill certain

- logical conditions, and
- what he calls "empirical conditions".

The logical requirements are firstly; identity of the domain of definition (that is the situations should be *conceptually* comparable), and secondly; the usage of the same units of measurement (which in our view not only comprises the same currency unit

as Martini states, but should also refer for example to the notion of "base" in this context) in the sense of "what is meant by 100%?", or "to which weights do we refer?"[175], and once we think of this type of logical prerequisites of comparability we will see that it is not as easy as Martini seems to think that (curiously) situations incomparable *directly* nonetheless turn out to be validly comparable *indirectly* (by virtue of multiplying).
As to the "empirical conditions" Martini postulates

1. the existence of an *overlap* of identical varieties of goods and services, and the
2. *representativity* of the selection of varieties for either or both situations to be compared.

This idea is visualized by the picture of overlapping circles (fig. 4.1.1). It is certainly a praiseworthy attempt to translate ideas like overlap (1) and "representativity"[176] (2) in operational terms. Martini uses absolute or relative numbers of items included in an index (or a link) for this purpose. But it is doubtful that *counting* items will solve the problem. Firstly, the items are of different importance (accounted for by "weights"), secondly the number of items distinguished is largely a question of the degree of detail a commodity classification is providing, and finally the same item[177], may be represented by a different number of concrete goods and services possibly also varying by type.[178] Anyway to aim at a *measure* of "representativity" is honorable, yet Martini's proposal hardly is the ultimate solution.
We now turn to Martini's idea that linking (chaining) successive direct comparisons (links) leads us to a valid indirect comparison of two situations, say 0 and t, which are not directly comparable. This position certainly needs some explanation, because by making indirect comparisons we do not remove a single impediment, which might prevent us from making a direct comparison.

b) Directly incomparable but indirectly comparable?

According to Martini multiplying links (each of which may meet the above–mentioned conditions) result in a chain which

- allegedly meets the logical and empirical conditions of comparability, while

[175] Two links of a chain differ from one another with respect to the structure of commodities and their respective weights. Hence we may not legitimately say that they "vary only in terms of time" as one of the "logical" conditions required by Martini reads.
[176] Surprisingly little if anything is known about the magnitude and effect of higher or lower "representativity" while this notion plays a most essential part in justifying chain indices.
[177] Or "elementary aggregate", or "basic heading" in the notation of Eurostat. Such an elementary aggregate is the lowest level of aggregation, the first stage where weights are attached to positions of a classification.
[178] In an attempt to assess "comparability" we are certainly in need of referring to the physical characteristics of a commodity. But as shown above an "item" is not "physically" constant, hence counting the number of equivalent items (as Martini suggests) might be difficult.

- the direct comparison for which it is at least a proxy is unable to meet these conditions.

This logic, though plausible at first glance, raises a number of questions. It is far from evident that multiplying links achieves the miracle of removing logical and empirical obstacles of comparability.[179] Whether and if so to which extent a justification for such a conclusion can be found, is rather a question of the assumptions (of proportionality for example) which we tacitly make when we conclude like Martini does. In what follows we try to show:

> The idea that linking (chaining) successive direct comparisons (links) would result in a valid indirect comparison of two situations not directly comparable is flatly wrong, because we overlook
> 1. assumptions necessary to accept an indirect comparison as a substitute for a direct comparison, and
> 2. differences in the notion of "comparability" in both cases and in the nature of the results.

In essence we tacitly assume that commodities *not* represented in the overlap or violating conditions of comparability, will undergo the same change in prices as commodities do that are included or comply with the conditions. Furthermore we have to make this assumption in case of

- each link taken in isolation, and
- each link in comparison with every other link.

But conditions are changing in the sense that there are different commodities making up the overlap in different links. This aspect is particularly important as it sheds some light on an alternative method to ensure comparability over a long period in time, conspicuously *not* mentioned by Martini: to make adjustments (as for example the well known adjustments in case of changes in quality) in order to enhance comparability.
It is difficult to understand how incomparability applies to situations 0 and t if directly compared, but not for the same pair if the comparison is made by multiplying links connecting 0 with $1, \ldots, q$ and $t-1$ with t. Once a chain is sufficiently long at least *some* links entering the formula will necessarily be no less "incomparable" among themselves as the endpoints are.
The only reason why an indirect comparison may be performed more easily than a direct one, is that we may feel entitled to take less care of comparability over the *whole* interval. There is for example no care taken for *conceptual* comparability, i.e. the above–mentioned constant domain of definition of an index function. Hence the result of indirect comparison can hardly have the same quality a direct comparison

[179] Martini justifies it by analogy with a "bridge country" in interspatial comparisons. But as shown in part c of this section, this is only an "argument" of plausibility.

would have.[180] Indeed *the result is of substantially different nature*: path dependence is an indication of this fact. Indirect comparisons are path dependent, because all intermediate situations relating to $1, \ldots, t-1$ will affect the result. Direct comparisons on the other hand are not path dependent.

Before showing this[181] we should note in passing, that there is an interesting additional inconsistency in Martini's position. The attempt to regard all index problems from the point of view of what makes comparisons valid or invalid is in principle a good idea, because in a sense the whole of statistics deals with making good comparisons. But to choose a theory of "comparability" as a starting point is insufficient, as

- it permits quite different conclusions, and
- it is too general.

It is not compelling to draw Martini–type conclusions. We would rather draw the conclusion that we should aim at making "pure" comparisons (sec. 4.4 and 8.2 for more details). Furthermore "comparability" is not a logical constant, an indisputably external visible characteristic even though we might think of measures to distinguish "more" from "less" of it.

For two things to be comparable it is necessary that there are aspects which they have in common and aspects in which they differ. It all depends on how strict the criteria are, that are applied to define the *common* aspects, and the criteria applied to find the *differences*. Moreover we should make sure that the underlying notion of comparability is the same, irrespective whether we refer to direct or to indirect comparisons. Interestingly

> Martini is very specific with respect to limitations of direct comparisons, but at the same time very unspecific as to limitations of indirect comparisons. There is not even a limit for the length of a chain. The logical and empirical conditions set out by Martini only state which *direct* comparisons are "impossible". Nothing is said about what is *indirectly* incomparable.

If we make use of Martini's conditions to direct and indirect[182] comparison most rigidly there is virtually nothing we can compare directly, but on the other hand chaining permits "comparability" of virtually everything with everything. Hence direct comparisons are readily called invalid or "impossible", but the same situations are allegedly indirectly comparable without limitations.

---

[180] Moreover the simple fact that a method is more easily applicable in face of the data situation can of course not be sufficient to justify it, or even to prove its superiority over another method, unless we are able to show that both methods will convey essentially the same information.

[181] See part d of this section.

[182] This can be derived from the fact that Martini gives no hint to how long a period over which a comparison can be made *indirectly* is allowed to be.

This indeed does not sound very logical, and we would even go a step further in that it is rather an indication that the underlying idea is erroneous. It is difficult not to suspect that there might be a contradiction in this kind of reasoning. The contradiction vanishes, however, when we see that the two types of comparisons involved in this argument have quite a different logical status. Comparing t with 0 directly means to express t in terms of 0 whereas an indirect comparison via $1, 2, \ldots, t-1$ is rather a summary description of the path connecting 0 and t, hence chaining results in a different kind of information as compared with a direct comparison.[183]

c) Limitations of the analogy with interspatial comparisons

There is a point in the idea of making situations that are incomparable, comparable by linking partial comparisons with a certain overlap when *countries*, like for example A, B, C are to be compared. Transitivity requires

$$P_{AC}P_{CB} = P_{AB}, \tag{4.1.1}$$

or

$$P_{AC}P_{CD}P_{DB} = P_{AB}, \tag{4.1.1a}$$

and is a fundamental condition in order to make a direct comparison between country A and B consistent with an indirect comparison via a third ("bridge") country C, or indirect comparisons via two "third" countries, C and D. In this sense the countries A and B are "linked" together via C or C and D or so.[184]

It appears to be no problem to translate this idea into a *temporal* framework, such that

$$P_{01}P_{12} = P_{02}, \tag{4.1.2}$$

where 0, 1, 2 denote some (not necessarily adjacent) periods in time. In connection with identity a return to the original situation gives an equation

$$P_{01}P_{12}P_{20} = P_{00} = 1, \tag{4.1.2a}$$

which explains the name "circular" test. Circularity should of course also apply to more than three periods, for example

$$P_{01}P_{12}P_{23}P_{34} = P_{04}. \tag{4.1.3}$$

[183] If *this* (i.e. the summary measure of a path) is what we want, there would be nothing wrong with chain indices, and we may call ourselves lucky for having evaded problems of direct comparison due to difficulties with data, but we should not expect the same kind of result as in the case we would aim at direct comparison, however difficult carrying out a direct comparison might be in practice.

[184] What applies to the transitivity of the parity P_{AB} between country A and B is of course also valid for direct and indirect comparisons of volumes (Q_{AB}).

The name "circular" or "chain" illustrates one of the applications: by virtue of this criterion in intertemporal comparisons the separate year to year index numbers can be joined together by successive multiplications like the links of a chain. The "circular test" is another test of Fisher's, and has always been the object of much controversy because virtually all reasonable index functions fail this test which therefore introduces an element of inconsistency.

The belief that incomparable situation could become comparable by chaining has intuitive appeal, but it is derived from a – not thoroughly thought over – equating of intertemporal conditions to comparisons between *countries*. We therefore turn to the differences between the two situations now. In some aspects intertemporal and interspatial comparisons are facing equivalent problems so that methods developed for one case can easily be translated into the other case[185]. But there are also dissimilarities, which interestingly will also shed some light to the relevance of certain axioms. The most significant differences are as follows:

1. Time has a natural order, a sequence is uniquely defined, t precedes $t+1$ whereas any of the 6 permutations (orderings) of three countries A, B, C (like $A - C - B$ or $B - C - A$) are equivalent. Hence the number of meaningful comparisons that can be made between k countries is significantly greater than in case of k periods (see below). Moreover if an additional country $k+1$ has to be taken into account results concerning the first k countries will be affected in general whilst a new period $t + 1$ will not give rise to redo calculations concerning the first t periods.

2. As a consequence axioms like time reversibility[186] and transitivity are important aspects in inter*national* (but not in inter*temporal*) comparisons, just because

   - there is no sequence, hence the direction of comparison can reasonably be inverted (and should be invariant upon such an inversion)[187]
   - in international comparisons transitivity gains importance because – as shown above – all indirect comparisons of say country A with country B (via C, or C and D and so on) should produce the same parity as the direct comparison does.[188]

---

[185] Traditional index theory was mainly concerned with comparing points or intervals in time, say 0 and t. But to make consistent (transitive) comparisons between countries with respect to prices or volumes is of growing and vital interest, especially in case of member–countries of a community like the EU. The definition of "multi–lateral" is clear in case of k countries, but far from clear in case of $t + 1$ (including 0) periods in time.

[186] Redefined as country reversal test.

[187] The idea is both, more plausible and more desirable in the interspatial context: we usually don't have any reason to prefer country A to B as a "base" of comparison and there should be only one uniquely determined parity between any two countries.

[188] Transitivity is also useful to fill gaps in an incomplete tableau of pairwise comparisons (parities). Hence there is a need for transitive parities (or parities made transitive). But the idea of "filling gaps" cannot simply be translated into "removing obstacles to comparability" as Martini does. Both axioms, time

3. Space (like countries A, B, C) is necessarily discrete only[189], whereas time can be conceived not only as discrete, but also as a continuous variate.

Other noticeable aspects of international comparisons without counterpart in intertemporal comparisons are for example: countries that have different and modifiable sizes, they can be disaggregated into smaller regions or aggregated into whole blocks, price indices (as opposed to quantity indices) are no longer pure numbers etc. Paradoxically

> There is no inter–*spatial* counterpart of chain indices and the Divisia index as both types of indices make use of a well defined order (sequence) of observations. Paradoxically time reversibility and transitivity which the idea of "chaining" is based on, play an important part primarily in the framework of interspatial comparisons. In case of temporal comparisons, on the other hand they apparently have little to recommend them.

This can be derived from the fact that attempts to demonstrate the importance of such criteria like reversibility and transitivity usually refer to the spatial, not to the temporal context.[190]

For multinational comparisons transitivity is required in the sense of compatibility of all possible direct and indirect comparisons. As mentioned already there are many more reasonable comparisons in the interspatial case. Between any two out of four countries we have 6 direct comparisons 12 indirect comparisons via *one* third country, and an additional 12 indirect comparisons via *two* third countries, that is 30 comparisons altogether. With $m = 4$ points in time involved, 0, 1, 2 , 3 there are only $m - 1$ comparisons with a *fixed* base (0–1, 0–2 and 0–3), and it makes sense to compare $m - 1$ *consecutive* time periods of the type 0–1, 1–2 and 2–3, in sum there are at best only $2(m - 1) - 1$ (avoid counting 0–1 twice) comparisons of interest.

While country reversibility may be desirable the same is not true for time reversibility. There is no point in requiring that the sequence 2 – 1 – 0 should yield a result reciprocal to 0 – 1 – 2. In the case of a Laspeyres or Paasche chain index this is in fact not guaranteed:

We see that again a kind of "antithetic" relation holds

$$\bar{P}^{LC}_{2-1-0} = \frac{1}{\bar{P}^{PC}_{0-1-2}}, \qquad \text{and} \qquad \bar{P}^{PC}_{2-1-0} = \frac{1}{\bar{P}^{LC}_{0-1-2}},$$

reversal and transitivity reflect a similar underlying independence and *invariance* idea, in the sense that an index function should be independent of a base or sequence, because they are arbitrary or nonexistent in this case.

[189] Therefore there is no counterpart to the Divisia's approach in international comparisons.

[190] Interestingly "reversal" (virtually change of *direction*) and transitivity are important just where no definite order (like time) exists. The conflict will be resolved, however, once we consider transitivity in terms of "temporal aggregation", that is when we introduce the notion of *aggregation over time*. Transitivity then means the ability to integrate consistently over an interval of time, irrespective of how this interval is subdivided into subintervals.

Table 4.1.1: "Time reversibility" in case of chain indices

	sequence 2 – 1 – 0	sequence 0 – 1 – 2
Laspeyres	$\bar{P}^{LC}_{2-1-0} = \frac{\sum p_1 q_2}{\sum p_2 q_2}\frac{\sum p_0 q_1}{\sum p_1 q_1}$	$\bar{P}^{LC}_{0-1-2} = \frac{\sum p_1 q_0}{\sum p_0 q_0}\frac{\sum p_2 q_1}{\sum p_1 q_1}$
Paasche	$\bar{P}^{PC}_{2-1-0} = \frac{\sum p_1 q_1}{\sum p_2 q_1}\frac{\sum p_0 q_0}{\sum p_1 q_0}$	$\bar{P}^{PC}_{0-1-2} = \frac{\sum p_1 q_1}{\sum p_0 q_1}\frac{\sum p_2 q_2}{\sum p_1 q_2}$

such that a chain Fisher index $\bar{P}^{FC}_{0t}$ would meet this condition of invariance upon a change of the *direction* of comparison, whatever the usefulness of such a property may be.
Even in the interspatial context the requirement of the country reversal test of a parity P_{AB} or P_{BA} between any two countries, A and B is not without problems. It would imply, that no matter, which of the two budgets q_A or q_B we refer to as a (weight) "base", we always get the same parity. It is unlikely that q_A is equally *characteristic* ("specific", or "representative") for A and B (and q_B respectively for B *and* A).[191]

> The requirements of country reversibility (and the same is true for transitivity) on the one hand and "symmetry" or "characteristicity" (and "equicharacteristicity") on the other hand follow the same idea of treating countries symmetrically. But they are clearly in conflict with one another. In the case of chain indices we encounter the same conflict: consistent temporal aggregation by chaining and representativity of weights (see sec. 4.2).

A simple method to guarantee transitivity is to choose one country, X as base land[192] in *all* comparisons between any two countries, like A, B, C etc. . Thus each parity between any two countries (P_{AB}) is defined indirectly via X. In case of the Laspeyres formula we get:

$$P^L_{AB(X)} = \frac{P^L_{XB}}{P^L_{XA}} = \frac{\sum p_B q_X}{\sum p_X q_X} \Big/ \frac{\sum p_A q_X}{\sum p_X q_X} = \frac{\sum p_B q_X}{\sum p_A q_X}, \tag{4.1.4}$$

which is the spatial analogon to the rebased index $P^L_{st(0)} = P^L_{0t}/P^L_{0s}$. But as another choice of the central country, say Y will yield different results $\left(P^L_{AB(Y)} \neq P^L_{AB(X)}\right)$ only

191 It is therefore questionable why the duality (P_{AB} and P_{BA}) should be eliminated artificially.
192 This is called "central country method". X can be one of the countries to be compared, but also an artificial country (or a kind of "average" country).

weak transitivity is given. By contrast *strict* transitivity requires consistency of *all* such indirect comparisons in relation to whichever third country.

> The temporal analogon of strict transitivity is: a unique result exists for P_{0t} irrespective of how the interval $(0, t)$ is subdivided into subintervals.

Eq. 3.2.5 shows that in general chain indices are unable to meet this criterion. The original idea of chaining over *time* periods will make the relevance of this criterion clear.

d) Temporal aggregation and chainability

Originally in the intertemporal case the underlying idea of chaining links was to make sure that a result referring to a year for example should not be affected by whether it is produced by data referring to months, or to quarters of a year, or to any other subdivision. In short: the idea is to aggregate *consistently over time*, irrespective of how an interval is subdivided. In case of relatives it is trivial as

$$\frac{p_{it}}{p_{i0}} = \frac{p_{is}}{p_{i0}} \cdot \frac{p_{it}}{p_{is}} = \frac{p_{ir}}{p_{i0}} \cdot \frac{p_{is}}{p_{ir}} \cdot \frac{p_{it}}{p_{is}} \tag{4.1.5}$$

will necessarily yield the same result, irrespective of how the interval between 0 and t is partitioned into two $(0 - s, s - t)$, or three $(0 - r, r - s, s - t)$, or whatever number of subintervals. The result is *unique*, not depending on the type and number of subintervals.

But it can easily be shown (as already indicated in eq. 3.2.5), that chain indices (unlike simple relatives) will not comply with precisely this idea[193]: for example $P_{01}P_{12}\ldots P_{56}$ will in general differ from $P_{02}P_{24}P_{46}$; that is both types of subdivisions of the same interval 0-6 will *not* yield the same result $\bar{P}_{06}$. This will be demonstrated in ex. 4.1.1.

Example 4.1.1

Given the following "data"

t = 0		t = 1		t = 2		t = 3		t = 4	
p	q	p	q	p	q	p	q	p	q
2	10	4	12	3	20	1	16	2	10
5	20	3	15	4	10	4	12	5	20

The direct index is of course $P^L_{04} = 1$ because all prices (and also quantities) in 4 equal those in 0 (indicated by shadows). The chain index not only violates identity but also yields different results:

[193] Moreover demand for chain indices ensues from a desire to make comparisons over *any* interval, in cases also where direct comparisons fails. But a chain index in action depends on which subintervals are combined. Thus chaining of indices is far from being an instrument of unambiguous temporal aggregation.

(a) $\bar{P}_{04}^{LC} = P_{02}^{LC} P_{24}^{LC} = \frac{110}{120} \cdot \frac{90}{100} = 0.825$*, with only* two *intervals (0, 2) and (2, 4), but*

(b) $\bar{P}_{04}^{LC} = P_{1}^{LC} P_{2}^{LC} P_{3}^{LC} P_{4}^{LC} = \left(\frac{100}{120} \cdot \frac{96}{93}\right) \cdot \left(\frac{60}{100} \cdot \frac{92}{64}\right) = 0.7419$ *upon dividing the* same *interval into four subintervals (0, 1), ..., (3, 4).*

The situations 0 and 4 are not uniquely compared. The result depends on the way the interval is subdivided. ◀

This is only one of several consequences of a more general feature of chain indices: they are "path–dependent" which is clearly in contradiction to what originally was intended by "chainability" and chaining of relatives.

"Circularity" in aggregating over time intervals means *ir*relevance of the path, as shown in eq. 4.1.5. By contrast chain indices depend on the "path" connecting the situations to be compared, which ironically is just the opposite of transitivity. Hence chain indices are not in line with the spirit of circularity and with what originally was intended by "chainability" and chaining. Furthermore:

> Path dependence instead of transitivity indicates that the information given by chain indices in fact differs from the information given by direct indices: chain indices provide a summary description of a process (of the "shape" of a time series) rather than a comparison of two situations only (or between the two endpoints of the chain only).

To be "path–dependent" in this context has two connotations

1. an external criterion (eq. 3.2.4): chain indices may diverge (drift) from direct indices,

2. an internal criterion (eq. 3.2.5): using different partitions of the same interval in time will in general produce different results of chain indices, as demonstrated above.

The next question of course is how "chainers" deal with this non–circularity of chain indices.

e) The role of chainability in justifying chain indices

HILL (1988) correctly states that chain indices are not transitive (chainable) in the first (external) sense as mentioned above. Moreover he also quotes a theorem to be presented in sec. 4.2 showing that transitivity is not compatible with variable weights.

To find a way out he thought it necessary to discuss the desirability of transitivity in general.

For Hill it "is tempting to discard the circularity test, as Fisher himself did", but he hesitates to go so far, mainly because transitivity has definite advantages in interspatial comparisons. He then turns the tables to arrive at the following opinion:

"...it must also be asked whether it is reasonable to judge a chain index by comparing it with its direct counterpart", and: "Advocates of chaining ought not to be in favor of circularity, because the identity between direct and indirect comparisons which satisfaction of the circularity test ensures makes the construction of a chain index superfluous. On the contrary, there must actually be a difference between the direct and the indirect measure for the latter to be superior on some criterion" (HILL (1988), p. 13).
It is interesting to study the logic underlying this kind of reasoning:

- Multiplying links (or "chaining") was originally motivated by chainability as a criterion to aggregate over periods in time. But chain indices bluntly contradict this idea. Aggregation is not unequivocal and it depends on the path connecting the endpoints.
- Next step: this being so to be *better* than direct indices a chain index should yield a *different* result. Hence discard chainability (the idea) but keep chaining (the operation).

In Hill's argument chainability is now playing the opposite role of what was originally intended. *Not* to be chainable is taken as proof of superiority[194] and the reason for multiplying links (the theory underlying the operation of chaining)[195] has been left vague.
This leads to the more fundamental problem of what kind of measurement an index aims at. Should this be the (relative) *level* of prices in period t as compared to period 0, or is an index a single numerical value reflecting the specific shape of a whole time series (see fig. 4.1.2)? In this sense it is sometimes maintained that chain indices serve *another purpose*[196] as they

- focus on *multi*lateral situations in the sense of a "multilateral index number theory" (DIEWERT and NAKAMURA (1993)) as opposed to "bilateral" (= direct indices), and as they
- provide "additional information" concerning the mechanism generating certain prices and quantities (see sec. 4.1f).

The problem with such a position is, however, that there is no longer a point in speaking of "superiority" and in finding principles (axioms) to distinguish good measures from bad ones.[197]

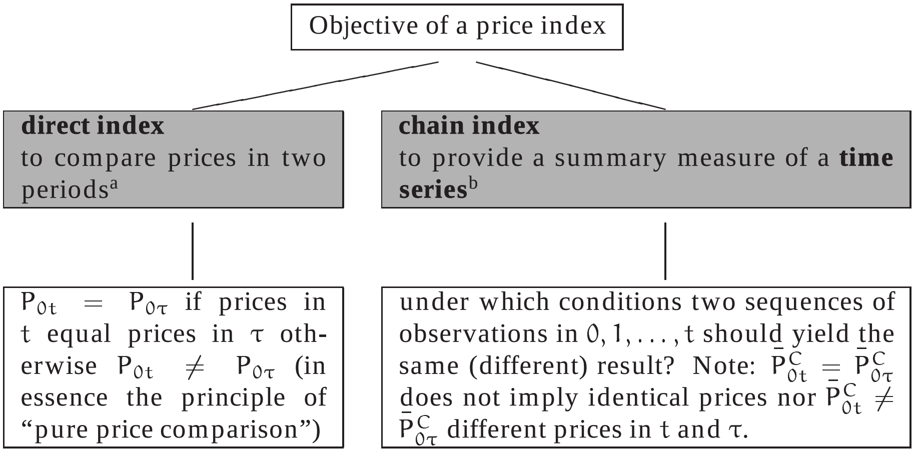
194 To be different is not a proof of being superior. Infinitely many index designs may differ from direct indices.

195 Another element of index theory, the approach of Divisia is also often viewed as a theoretical background of chaining, see chapter 7.

196 If this really were the case there would be no logic in saying that one approach (chain) is superior to the other (direct). It is nonsense to say, for example the variance is better than the correlation coefficient, as there are two measures designed for different purposes.

197 A "multilateral test approach has not been as well developed as the bilateral approach" (DIEWERT and NAKAMURA (1993), p. 10).

Figure 4.1.2: Objectives and principles of index construction



[a] the symbols t and τ denote different periods, or different data, relating to the same period.
[b] or "multilateral index number theory" as opposed to "bilateral" (= direct indices) according to DIEWERT and NAKAMURA, (1993).

In the case of direct binary comparisons a possible principle governing measurement is the principle of "pure price comparison". It permits a decision under which conditions we should obtain $P_{0t} = P_{0\tau}$, or $P_{0t} \neq P_{0\tau}$ respectively. Some fundamental axioms also follow this idea precisely. However as indicated in fig. 4.1.2 in the case of chain indices, like $\bar{P}^{LC}_{0t}$ there is no straightforward answer to the question when two index numbers P_{0t} and $P_{0\tau}$ should be equal or differ. The operation of "chaining" as such cannot play the part of a theoretical background as "pure price comparison" plays in the case of P^{L}_{0t}.

f) "Additional information" and "irrelevance of the base"

An argument in favor of chain indices already mentioned[198] is that such indices make use of time series data in a more efficient way[199], and are able to provide additional valuable (as compared with direct indices) information. This amounts to making path dependence a virtue. Unfortunately we (in general) don't hear much about the particular type of this information allegedly to be found in looking at $\bar{P}^{LC}_{02}$ for example, or at $\bar{P}^{FC}_{02}$, but not to be found in P^{L}_{02} or P^{F}_{02}.
According to ALLEN (1975), p. 177 a disadvantage of P^{L}_{02} is: "There is no reference whatever to the course of prices/quantities in between. Something better than this

[198] Also called "multiplication mystery" (**A3** and also **B1** in sec. 6.1). Authors using this argument are for example MUDGETT (1951), p. 74f., or ALLEN (1975), p. 145f and p. 177f.
[199] See sec. 4.1g showing that rather the opposite might be true.

must be sought, something more in line with economic common sense and making more efficient use of all the data". As to the meaning of common sense in this context the only hint given by Allen is, that consumer prices "in year t would be influenced by prices before year t as well as those achieved in that year" (p. 145f), which of course is trivial. But remember:

- An index should be a summary of observations, not a description of the mechanism by which they come about. In order to study such a mechanism it would be more appropriate to estimate an econometric model.
- A given result for P^L_{02} does not imply that prices of $t = 2$ were independent of prices in preceding periods. Nor does $\bar{P}^{LC}_{02} \neq P^L_{02}$ give any indication of the kind of mechanism which operates to produce the observed prices.

We may also ask: Why should an index reflect the influence of past prices in exactly the same manner as implied by a chain index[200]? And: How does it fit to "economic common sense", on the other hand, that equality of prices in 0 und 2 respectively not necessarily leads to $\bar{P}^{LC}_{02} = 1$? Obviously a better way to study the development of other factors *not* reflected in P^L_{0t} would rather be to analyze the correlation between P^L_{0t} and some other variable, as for example the change of quantities. To calculate $\bar{P}_{01}, \bar{P}_{02}, \ldots, \bar{P}_{0t}$, a series reflecting the influence of *several* variables is definitely less wise a procedure than treating different phenomena *separately*, measuring each of them as "purely" as possible and correlating them thereafter. To sum up, the argument of "additional information" is not only vague, it should be questioned for a number of reasons.
The definition of a chain, $\bar{P}_{0t} = \bar{P}_{0s}\bar{P}_{st}$ according to eq. 1.2.4 gives rise to the impression that a chain index is *independent of the base*[201] such that switching from base 0 to base s is only an insignificant formal operation. Is this correct?
The principle of chaining as a method of temporal aggregation is to make a long period comparison between 0 and t consistent with a series of short period comparisons. In the case of price *relatives* this can be done simply by multiplication (eq. 4.1.5), but for this procedure to be valid in case of *indices* as well some assumptions of proportionality are tacitly made:

- $P_{0t} = P_{0s}P_{st}$ is tantamount to assuming $\frac{P_{0t}}{P_{0s}} = \frac{P_{st}}{P_{ss}}$, $P_{ss} = 1$, by the same token
- $P_{0t} = P_{0r}P_{rs}P_{st}$ implies

$$\frac{P_{0t}}{P_{0r}} = \frac{P_{rt}}{P_{rr}} \tag{4.1.6}$$

and in addition $\frac{P_{rt}}{P_{rs}} = \frac{P_{st}}{P_{ss}}$.

200 See sec. 3.5.
201 For this impression it is important that the term "base" has two meanings in this context: weight base and reference base.

This shows:

> To proceed in the way indicated in eq. 4.1.5 in the case of indices (instead of relatives) the assumption is made implicitly that indices at different bases (and mostly with different weights) will change *in proportion*.

But ironically chain indices are motivated on the assumption that indices which have a different base (and weight structure) will in general *not* change in proportion, and switching to a new base will be better than keeping to the old one. This clearly appears contradictory.[202]

g) Restrictions imposed on a time series by transitivity

Fisher's circular test (i.e. transitivity) has always been highly controversial, and he proposed to discard this test as he already noticed that it introduced an element of inconsistency into his system of "tests". But Fisher went even further. He rigorously argued against this test and called the test "theoretically a mistaken one" and its fulfillment "should really be taken as proof that the formula which fulfills it is erroneous". As PFOUTS (1966) has shown, to construct a $(T+1) \times (T+1)$ matrix of indices ($T = 3$)

$$\mathbf{P} = \begin{bmatrix} P_{00} & P_{01} & P_{02} & P_{03} \\ P_{10} & P_{11} & P_{12} & P_{13} \\ P_{20} & P_{21} & P_{22} & P_{23} \\ P_{30} & P_{31} & P_{32} & P_{33} \end{bmatrix}$$

such that the indices P satisfy chainability and identity, is tantamount to imposing *singularity* on the matrix $\mathbf{P}$ (that is the determinant $|\mathbf{P}| = 0$) which he found is "unduly restrictive". Singularity of $\mathbf{P}$ imposes "artificial restrictions" on index numbers "that typically would not be present in the empirical data", and this would rob, in Pfouts' view, index numbers of desirable flexibility, leading also to the following intolerable consequences

- a linear relationship must exist between index numbers having the same base and even worse
- the same relationship applies to all base periods.

This can easily be verified. Taking the definitions of time reversibility, identity and circularity into account we obtain

$$\mathbf{P} = \begin{bmatrix} 1 & P_{01} & P_{01}P_{12} & P_{01}P_{12}P_{23} \\ 1/P_{01} & 1 & P_{12} & P_{12}P_{23} \\ 1/P_{01}P_{12} & 1/P_{12} & 1 & P_{23} \\ 1/P_{01}P_{12}P_{23} & 1/P_{12}P_{23} & 1/P_{23} & 1 \end{bmatrix}$$

[202] The conflict in other words: constantly adjusting *weights*, an aspect of dynamics allegedly best served by chain indices makes the assumption underlying eq. 4.1.6 invalid. Hence chaining is an operations poorly justified.

and the determinant $|\mathbf{P}|$ in fact vanishes: multiplication of row 1 by $1/P_{01}$ gives row 2, or multiplication of row 3 by P_{12} gives row 2. Hence **P** is singular ($|\mathbf{P}| = 0$). A consequence is that a single additional value, for example P_{34} is sufficient to calculate an additional fifth row and column. By definition multiplication of **P** by

$$\mathbf{c} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ P_{34} \end{bmatrix}$$

yields a column vector

$$\mathbf{Pc} = \begin{bmatrix} P_{04} \\ P_{14} \\ P_{24} \\ P_{34} \end{bmatrix}$$

summarizing all time series information of **P** expressed in terms of base 0 (first row of **P**), or base 1 (second row) etc. There are for example actually only two independent observations, P_{01} and P_{12} assembled in the 3×3 matrix

$$\mathbf{P} = \begin{bmatrix} P_{00} & P_{01} & P_{02} \\ P_{10} & P_{11} & P_{12} \\ P_{20} & P_{21} & P_{22} \end{bmatrix}.$$

It can also easily be verified that time reversibility implies all second order principal minors of **P**, to be identically singular, hence determinants like

$$\begin{vmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{vmatrix} \quad \text{or} \quad \begin{vmatrix} P_{33} & P_{35} \\ P_{53} & P_{55} \end{vmatrix}$$

and so on, all will vanish. The conditions enforced on the index numbers are indeed “unduly restrictive” and it appears plausible that they are also incompatible with varying weights or changing conditions of demand or supply. Making better use of time series data by chain indices is at least somewhat doubtful, and the justification of the operation of “chaining” does not seem to be thoroughly thought over.

Fisher already conjectured that in order to meet transitivity an index should have weights *not* depending on 0 or t, but being constant throughout the whole interval of successive periods under consideration.[203] Hence many inconsistency theorems (non–existence theorems) have to do with the circular test (BALK (1995))[204]. Furthermore chainability (the underlying idea of chaining), is not compatible with a regular adjustment of quantity weights.

[203] On the other hand to indices not depending on quantities, q_0 and q_t, the idea of factor reversibility could not apply. Thus Fisher already suspected (without proof), that adding the circular test to the factor reversal test (or vice versa) introduces an element of inconsistency. He simply could not find any index formula able to satisfy both tests, the circular test and the factor reversal test.

[204] There do not exist for example functions, P_{0t} and Q_{0t} which satisfy circularity, and identity and the product test (or the factor reversal test) simultaneously.

4.2 Chaining and constant adjustment of quantity weights

In this section we examine the relationship between weights (expenditure shares or quantities respectively), on the one hand and a) chainability, and b) chaining on the other hand. As chain indices are not *transitive* (and chainability might be viewed as a *theoretical* justification for chaining only, not as irremissable for chaining), the incompatibility of variable weights and transitivity is not in itself an argument against chaining. This gives rise to exploring in some depth the relationship between chaining and (the frequency of) updating of weights.

a) Chainability and variable weights

In the "chainers" opinion one of the main advantages of chain indices consists in solving both problems, simultaneously:

- to arrive at consistent long term inter–temporal comparisons by multiplying over subintervals, and
- to account for new situations by allowing for a constant adjustment of weights.

The latter aspect was one of the main reasons for Alfred Marshall to advocate chain indices. The dilemma, however, is that you will never get both "advantages" simultaneously. As I. Fisher already conjectured: there are chainable indices with constant weights, and there are indices with variable weights violating chainability.

But it was not until the proof of FUNKE et al. (1979) that Fisher's presumption could be rigorously proved. According to the theorem of Funke et al. the only index, satisfying the minimal requirements monotonicity, linear homogeneity, identity and commensurability and at the same time being able to pass the circular test is the so called "Cobb–Douglas index" given by

$$P_{0t}^{CD} = \prod_{i=1}^{n} \left(\frac{p_{it}}{p_{i0}} \right)^{\alpha_i} \tag{4.2.1}$$

where "weights" $\alpha_1, \alpha_2, \ldots, \alpha_n$ are real constants not depending on period 0 or t, and $\sum \alpha_i = 1$. Hence for FUNKE et al. (1979), there is no index formula which is able to combine both alleged advantages of chain indices.[205]

> Chainability (transitivity) is clearly inconsistent with a constant adjustment of weights (as it is inconsistent for example with the factor reversal test as well).

[205] "... that the main intention of the circular test, that is, the adjustment of the quantity weights to the new situation in each new dual comparison around a circle of periods or places cannot be accomplished. There simply does not exist such a formula ... " (p. 685).

Thus the history of chain indices started with a misunderstanding, an attempt to reconcile two contradictory properties.[206]

Note that P^{CD}_{0t} is a weighted geometric mean with constant weights α_i for all periods. The growth factor of P^{CD} is the geometric mean of the growth factors (links) of the individual price relatives, and it is again constant as the factors α_i are constants, whereas for example the growth factor of P^L_{0t} is an arithmetic mean with changing weights tending to the largest individual growth factor of prices.[207] P^{CD} is a "degenerate"[208] price index, however. But P^{CD} is able to fulfill the "circular test", like for example the index function of Lowe (LW)[209]

$$P^{LW}_{0s}P^{LW}_{st} = \frac{\sum p_s q}{\sum p_0 q} \cdot \frac{\sum p_t q}{\sum p_s q} = P^{LW}_{0t} = \frac{\sum p_t q}{\sum p_0 q} \tag{4.2.2}$$

with constant (fixed) quantities, that is weights again (like in the case of P^{CD}) not depending on (either or both) periods to be compared (periods 0 and t). By contrast to P^{CD} and P^{LW}, P^L is not transitive:

$$P^L_{0s}P^L_{st} = \frac{\sum p_s q_0}{\sum p_0 q_0} \cdot \frac{\sum p_t q_s}{\sum p_s q_s} \neq P^L_{0t} = \frac{\sum p_t q_0}{\sum p_0 q_0}.$$

The same is true for the logarithmic Laspeyres index

$$DP^L_{0t} = \prod \left(\frac{p_{it}}{p_{i0}}\right)^{w_{i0}} \qquad \text{where} \qquad w_{i0} = p_{i0}q_{i0} \Big/ \sum p_{i0}q_{i0} \ ,$$

the logarithmic Paasche index DP^P_{0t} with weights w_{it} defined correspondingly, and the Törnquist index

$$P^T_{0t} = \prod_{i=1}^{n} \left(\frac{p_{it}}{p_{i0}}\right)^{\bar{w}_i} = \sqrt{DP^L_{0t}DP^P_{0t}} \qquad \text{where } \bar{w}_i = \frac{1}{2}\left(w_{i0} + w_{it}\right). \tag{4.2.3}$$

Obviously

$$DP^L_{0t} = \prod \left(\frac{p_t}{p_0}\right)^{w_0} \neq \prod \left(\frac{p_s}{p_0}\right)^{w_0} \prod \left(\frac{p_t}{p_s}\right)^{w_s} = DP^L_{0s}DP^L_{st}.$$

[206] It already dawned upon Fisher for example, that the main intention of his circularity axiom, that is the rapid adjustment of budgets or "baskets" cannot be accomplished, and this was a main reason for his eventual abandoning of the circularity criterion. Interestingly unlike Alfred Marshall he was also not a proponent of chain indices.

[207] The attempt to avoid this tendency to be dominated by the fast growing prices is one of the reasons why geometric mean indices are preferred to arithmetic mean indices like P^L in case of indices for exchange rates and share prices.

[208] According to BALK (1995), P^{CD} is a "degenerate member" of the class of price indices, "because it is *only* a function of prices" and "the derived quantity indicator" V_{0t}/P^{CD} "is in general not admissible" and the pair P^{CD}, Q^{CD} is unable to pass the product test.

[209] Lowe's index, however, violates commensurability and thus P^{LW} can hardly be regarded as a reasonable "price index". There is also no quantity index which can meaningfully be related to Lowe's price index formula P^{LW}.

Likewise DP^P and P^T are not transitive either. Several attempts have also been made to formulate *economic theory* counterparts of some "tests" or "axioms" of axiomatic theory, for example the Economic Circular Test[210] (EC)

$$P_U(\mathbf{p}_0, \mathbf{p}_b, u_0, u_b) P_U(\mathbf{p}_b, \mathbf{p}_t, u_b, u_t) = P_U(\mathbf{p}_0, \mathbf{p}_t, u_0, u_t)$$

introduced by BOSSERT and PFINGSTEN (1987) along with a function F combining two utility levels $F(u_0, u_t) = F \in D_U$ for all $u_0, u_t \in D_U$ where D_U is a set of admissible (for example positive) utility levels. They then have shown that EC is met only under highly unrealistic conditions: U is a homothetic utility function or for all three periods $0, b, t$ we have $F(u_0, u_b) = F(u_b, u_t) = F(u_0, u_t) = u^*$ where u^* is a *constant* utility level or in other words "it is necessary and sufficient to use a fixed reference level irrespective of the actual utility levels"[211]. In conclusion:

> Chainability is achieved only at the expense of constant weights not related to expenditures in the periods under consideration, and only under highly unrealistic conditions from the theoretical point of view. It can therefore hardly serve as a justification for chaining.[212]

b) Chaining and frequency of updating of weights

For the indices P_{0t} (0 constant, t variable, P_{0t} expressed as a ratio of expenditures) in order to be independent of how the interval (0, t) is subdivided into subintervals (let denote P^*_{0t} the index under consideration), a sufficient condition for weights (quantities) to hold would be

$$P^*_{st} = P_{st(0)} = \frac{\sum p_t q_0}{\sum p_s q_0} = \frac{P^L_{0t}}{P^L_{0s}} = g^0_{s+1} \cdot g^0_{s+2} \cdot \dots \cdot g^0_t. \tag{4.2.4}$$

Note that this index is for base periods s other than 0 not the Laspeyres index amounting to $P^L_{st} = \sum p_t q_s / \sum p_s q_s$, but rather the *rebased* Laspeyres index gained from the condition $P_{0t} = P_{0s} P^*_{st}$, such that $P^*_{st} = P_{st(0)}$ is an index which is transitive *by construction* and in which weights are *never* $(x = 0)$ updated.
In what follows we try to examine the drift of an index $\hat{P}^x_{0t}$ which is rebased regularly (to be more distinct: x times within the interval from 0 to t) from the index P^*_{0t}, and to find a relationship between x, the frequency of updating weights, and the index P^*_{0t}, in which in fact no update is made (such that $P^*_{st} = \hat{P}^0_{st} = \sum p_t q_0 / \sum p_s q_0$, that is the 0–period–basket is kept constant). For the sake of simplicity assume $t = 8$ and x weight renewals as shown in tab. 4.2.1.

[210] Note that these tests are *not* equivalent to tests in the *statistical* (axiomatic) price index theory, where the focus is on which type of index function is satisfying the test. The question now is whether or not the test holds for a certain *utility*function.
[211] op. cit. p. 283.
[212] This does not mean that chainability is useless. It is certainly a device of achieving consistent temporal aggregations whatever the interpretation of "weights" in this context may be

Table 4.2.1: Chaining and frequency (x) of updating weights

a) Drift D^x_{0t} of $\hat{P}^x_{0t}$ against $\hat{P}^0_{0t}$

t	$\hat{P}^1_{0t}$	$\hat{P}^2_{0t}$	$\hat{P}^3_{0t}$
1	1	1	1
2	1	1	1
3	1	1	$\bar{g}^2_3$
4	1	$\bar{g}^3_4$	$\bar{g}^2_3\bar{g}^2_4$
5	$\bar{g}^4_5$	$\bar{g}^3_4\bar{g}^3_5$	$\bar{g}^2_3\bar{g}^2_4\bar{g}^4_5$
6	$\bar{g}^4_5\bar{g}^4_6$	$\bar{g}^3_4\bar{g}^3_5\bar{g}^3_6$	$\bar{g}^2_3\bar{g}^2_4\bar{g}^4_5\bar{g}^4_6$
7	$\bar{g}^4_5\bar{g}^4_6\bar{g}^4_7$	$\bar{g}^3_4\bar{g}^3_5\bar{g}^3_6\bar{g}^6_7$	$\bar{g}^2_3\bar{g}^2_4\bar{g}^4_5\bar{g}^4_6\bar{g}^6_7$
8	$\bar{g}^4_5\bar{g}^4_6\bar{g}^4_7\bar{g}^4_8$	$\bar{g}^3_4\bar{g}^3_5\bar{g}^3_6\bar{g}^6_7\bar{g}^6_8$	$\bar{g}^2_3\bar{g}^2_4\bar{g}^4_5\bar{g}^4_6\bar{g}^6_7\bar{g}^6_8$

b) Frequency of renewal of weights and number of price and quantity vectors

$\hat{P}^x_{0t}$	quantity weights	$\hat{P}^x_{08}$	kind / number of (additional) determinants
$\hat{P}^0_{0t}$	q_0 only	$\sum p_8q_0 / \sum p_0q_0$	vectors $\mathbf{p}_0, \mathbf{p}_8, \mathbf{q}_0$
$\hat{P}^1_{0t}$	q_0, q_4 (1 renewal at period 4)	$\frac{\sum p_4q_0}{\sum p_0q_0}\frac{\sum p_8q_4}{\sum p_4q_4}$	in addition to $\hat{P}^0_{0t}$: $\mathbf{p}_4, \mathbf{q}_4$ (two more determinants)*
$\hat{P}^2_{0t}$	q_0, q_3, q_6 (2 renewals at period 3 and 6)	$\frac{\sum p_3q_0}{\sum p_0q_0}\frac{\sum p_6q_3}{\sum p_3q_3}\frac{\sum p_8q_6}{\sum p_6q_6}$	in addition to $\hat{P}^0_{0t}$ four (2×2) more determinants: $\mathbf{p}_3, \mathbf{p}_6, \mathbf{q}_3, \mathbf{q}_6$
$\hat{P}^3_{0t}$	q_0, q_2, q_4, q_6 (3 renewals at period 2, 3 and 6)	$\frac{\sum p_2q_0}{\sum p_0q_0}\frac{\sum p_4q_2}{\sum p_2q_2}\frac{\sum p_6q_4}{\sum p_4q_4}\frac{\sum p_8q_6}{\sum p_6q_6}$	3×2 additional determinants: $\mathbf{p}_2, \mathbf{p}_4, \mathbf{p}_6,$ $\mathbf{q}_2, \mathbf{q}_4, \mathbf{q}_6$
$\bar{P}^{LC}_{0t}$	$x = t - 1 = 7$ renewals	$\frac{\sum p_1q_0}{\sum p_0q_0}\frac{\sum p_2q_1}{\sum p_1q_1}\cdots\frac{\sum p_8q_7}{\sum p_7q_7}$	7×2 additional determinants: $\mathbf{p}_1, \mathbf{p}_2, \ldots \mathbf{p}_7,$ $\mathbf{q}_1, \mathbf{q}_2, \ldots \mathbf{q}_7$

* Given $\mathbf{q}_4 = \mathbf{q}_0$ we then obtain the same result as in case of $x = 0$, irrespective of prices $\mathbf{p}_4$.

In case of $x = 1$ it is assumed that beginning with $t = 5$ use is made of the new "basket" of $t = 4$ instead of $t = 0$ and on account of $P_{0t} = P_{04}P_{4t}, t \geq 5$ we get

$$\hat{P}^1_{05} = \frac{\sum p_4 q_0}{\sum p_0 q_0} \frac{\sum p_5 q_4}{\sum p_4 q_4} = g^0_1 g^0_2 g^0_3 g^0_4 g^4_5,$$
$$\hat{P}^1_{06} = \frac{\sum p_4 q_0}{\sum p_0 q_0} \frac{\sum p_6 q_4}{\sum p_4 q_4} = g^0_1 g^0_2 g^0_3 g^0_4 g^4_5 g^4_6, \text{ etc.}$$

By the same token an index with two ($x = 2$) weight–renewals (baskets of 3 and 6 respectively) linked together is given by

$$\hat{P}^2_{04} = \frac{\sum p_3 q_0}{\sum p_0 q_0} \frac{\sum p_4 q_3}{\sum p_3 q_3} = g^0_1 g^0_2 g^0_3 g^3_4,\ \hat{P}^2_{05} = \hat{P}^2_{04} g^3_5,\ \hat{P}^2_{06} = \hat{P}^2_{05} g^3_6,$$

and after the second renewal

$$\hat{P}^2_{07} = \frac{\sum p_3 q_0}{\sum p_0 q_0} \frac{\sum p_6 q_3}{\sum p_3 q_3} \frac{\sum p_7 q_6}{\sum p_6 q_6} = g^0_1 g^0_2 g^0_3 g^3_4 g^3_5 g^3_6 g^6_7,$$

and so on.
It is interesting now to see how indices $\hat{P}^1_{0t}$ and $\hat{P}^2_{0t}$ drift away from $\hat{P}^0_{0t} = P^*_{0t(0)}$. Let

$$\bar{g}^k_t = g^k_t / g^0_t \ . \tag{4.2.5}$$

If for all s ($0 \leq s \leq t$) $P^*_{st} = \hat{P}^0_{st}$ is defined according to eq. 4.2.4 transitivity (invariance with respect to the partitioning into subintervals) as well as comparability (any two indices differ with respect to prices only) will hold for the whole interval (0, t). As tab. 4.2.1 shows the length of the interval in which the drift

$$D^x_{0t} = \hat{P}^x_{0t} / \hat{P}^0_{0t} \tag{4.2.6}$$

will result in $D^x_{0t} = 1$ will shorten (as indicated by shadows) as x increases. D^x_{0t} is a function of the number x of weight–renewals, and of the terms $\bar{g}^k_t$. Likewise the length of the interval in which pure comparisons are possible will shorten. For example the rise of prices in $5 \to 7$ is

$$\frac{\hat{P}^1_{07}}{\hat{P}^1_{05}} = g^4_6 g^4_7 = \frac{\sum p_7 q_4}{\sum p_5 q_4},$$

a "pure" comparison of prices p_7 with p_5 on the basis of *one* set of quantities only (q_4). The same is true for $\hat{P}^0_{07} / \hat{P}^0_{05} = \sum p_7 q_0 / \sum p_5 q_0$ with respect to q_0. By contrast

$$\frac{\hat{P}^3_{07}}{\hat{P}^3_{05}} = g^4_6 g^6_7 = \frac{\sum p_7 q_6}{\sum p_6 q_6} \frac{\sum p_6 q_4}{\sum p_5 q_4}$$

makes use of *two* quantity vectors and of prices relating to the intermediate period 6 as well. Interestingly the chain Laspeyres index is simply the limiting case in the following sense:

$$\bar{P}^{LC}_{08} = \hat{P}^{7}_{08}. \tag{4.2.7}$$

Tab. 4.2.1 clearly shows that to the extent to which the frequency x of weight–renewals is increased

- there will be more price and quantity vectors affecting the result of $\hat{P}^{x}_{0t}$, which is thereby ever less unequivocal and it also reflects more and more a quantity movement (derived from the fact, that $\mathbf{q}_4 \neq \mathbf{q}_0$ in $\hat{P}^{1}_{0t}$ or $\mathbf{q}_0 \neq \mathbf{q}_3 \neq \mathbf{q}_6$ in $\hat{P}^{2}_{0t}$ etc.) in addition to the price movement, and

- the earlier $\hat{P}^{x}_{0t}$ will drift away from $\hat{P}^{0}_{0t}$, and the shorter the interval over which comparisons of prices can be made on the basis of the same quantity weights or in which transitivity holds as in $\hat{P}^{0}_{st} = P_{st(0)}$ for all $0 \leq s \leq t$ will be.

This again demonstrates that an adjustment of weights and chainability (transitivity) are inconsistent, and more specific:

> With each additional renewal of weights ($x \rightarrow x+1$), say at period s the number of determinants of a chained index increases by 2 (one price vector $\mathbf{p}_s$, and one quantity vector $\mathbf{q}_s$ respectively), thus making the overall result sensitive for the situation at period s. A price index P_{0t} where weights q_0 are adjusted (renewed) x times (for example at $r, s, \ldots$ such that $P_{0t} = P_{0r}P_{rs}\ldots$) is a function of $2x$ vectors *in addition* to the vectors $\mathbf{p}_0, \mathbf{p}_t$, and $\mathbf{q}_0$.

In particular a chain index, like $\bar{P}^{LC}_{0t}$ or $\bar{P}^{FC}_{0t}$ is affected by t, (or $t+1$ respectively) quantity vectors and $t+1$ price vectors, as opposed to only

- *one* quantity vector ($\mathbf{q}_0$) in case of P^{L}_{0t} (or two, $\mathbf{q}_0$ and $\mathbf{q}_t$ in case of P^{F}_{0t}) and

- *two* price vectors, $\mathbf{p}_0$, and $\mathbf{p}_t$,

both numbers being independent of t. We have good reasons to believe that this difference between direct and chain indices is more than just a matter of mere numbers. Above all it is difficult to see why the number of determinants in $\bar{P}^{LC}_{0t}$ or $\bar{P}^{FC}_{0t}$ should constantly increase as time t goes on (i.e. the chain becomes longer). This is again contrary to the idea of transitivity, which permits a consistent integration over a time–interval, the result of which is depending on 0 and t only.

4.3 Chaining and consistency in aggregation

We now turn to the idea of making consistent integrations over time and trying to show how this fits to aggregation over commodities, or sub–indices (or “items” in general). It turns out that whenever more than one item is involved chaining, or comparing *indirectly* seems to be much more difficult an approach than doing so *directly*. Problems with consistency in aggregation (over types of commodities) will inevitably emerge, as soon as chaining and constant adjustment of weights comes into play. The original conceptualization of chaining by the Danish statistician H. WESTERGAARD (1890) was developed in terms of stocks and flows. Denote the initial stock by A and net increases by Z_1 and Z_2 referring to the first and second half of the year respectively. The stock then develops as follows:

period	stock	relative
0	A	$m_{00} = A\ /A = 1$
1	$A + Z_1$	$m_{01} = (A + Z_1)\ /A$
2	$A + Z_1 + Z_2$	$m_{02} = (A + Z_1 + Z_2)\ /A$

and for the relatives we find

$$m_{02} = \frac{A + Z_1}{A} \cdot \frac{A + Z_1 + Z_2}{A + Z_1} = m_{01} m_{12}.$$

Note that this only shows that the stock may be calculated directly (m_{02}) or indirectly ($m_{01} m_{12}$). None of the two procedures deserves to be preferred to the other. It is also irrelevant into which and into how many subintervals the interval is divided.

The situation becomes much more difficult, however, when stocks and flows are broken down into two sectors with stocks $A = A_1 + A_2$, and flows $Z_1 = Z_{11} + Z_{12}$ and $Z_2 = Z_{21} + Z_{22}$. Relatives considered separately for each sector ($i = 1, 2$) obviously remain transitive as $m_{(i)02} = m_{(i)01} m_{(i)12}$ for both i ($i = 1, 2$) taken in isolation.

	period 1	period 2
sector 1	$m_{(1)01} = (A_1 + Z_{11})\ /A_1$	$m_{(1)12} = (A_1 + Z_{11} + Z_{21}) \big/ (A_1 + Z_{11})$
sector 2	$m_{(2)01} = (A_2 + Z_{12})\ /A_2$	$m_{(2)12} = (A_2 + Z_{12} + Z_{22}) \big/ (A_2 + Z_{12})$

Obviously

$$m_{(1)02} = \frac{A_1 + Z_{11}}{A_1} \cdot \frac{A_1 + Z_{11} + Z_{21}}{A_1 + Z_{11}} = \frac{A_1 + Z_{11} + Z_{21}}{A_1}, \tag{4.3.1}$$

and $m_{(2)02}$ analogously.

But it is difficult to see why multiplication of *links* with reference to *different* structures, that is with weights $w_{1i} = A_i \big/ \sum A_i$ for $m_{(i)01}$, and $w_{2i} = (A_i +$

$Z_{1i}) / (\sum A_i + \sum Z_{1i})$ for $m_{(i)12}$ leading to

$$m_{02} = \left(w_{11} m_{(1)01} + w_{12} m_{(2)01}\right)\left(w_{21} m_{(1)12} + w_{22} m_{(2)12}\right)$$
$$= \frac{A_1 m_{(1)01} + A_2 m_{(2)01}}{A_1 + A_2} \cdot \frac{(A_1 + Z_{11}) m_{(1)12} + (A_2 + Z_{12}) m_{(2)12}}{(A_1 + Z_{11}) + (A_2 + Z_{12})} \tag{4.3.2}$$

should be any better than going the direct way of aggregating over *relatives* using *constant* weights w_{1i} and yielding

$$m_{02} = \frac{A_1 m_{(1)02} + A_2 m_{(2)02}}{A_1 + A_2} = w_{11} m_{(1)02} + w_{12} m_{(2)02}. \tag{4.3.3}$$

The weights to be attached to the breakdown of the link m_{01} into $m_{(1)01}$ and $m_{(2)01}$ and of the link m_{12} into $m_{(1)12}$ and $m_{(2)12}$ are themselves depending on the links such that

$$w_{21} = w_{11} m_{(1)01} / \left(w_{11} m_{(1)01} + w_{12} m_{(2)01}\right), \tag{4.3.4}$$

and

$$w_{22} = w_{12} m_{(2)01} / \left(w_{11} m_{(1)01} + w_{12} m_{(2)01}\right) \tag{4.3.4a}$$

Substituting these terms for w_{21} and w_{22} in eq. 4.3.2 we can easily see that we get

$$m_{02} = w_{21} m_{(1)01} m_{(1)12} + w_{22} m_{(2)01} m_{(2)12}$$

in accordance with eq. 4.3.3 since for both i $m_{(i)02} = m_{(i)01} m_{(i)12}$.

Unlike in the case of a breakdown of price indices into sub–indices the weights here are recursively defined: $w_{1i}, m_{(i)01} \rightarrow w_{2i}$, and $w_{2i}, m_{(i)12} \rightarrow w_{3i}$, etc.

In any case chaining does not seem to be preferable over direct computation, once the variable in question is made up of subaggregates by the simple reason that in general the structure of the aggregate will change with the passage of time. This can easily be seen from eq. 4.3.4 and 4.3.4a: if $m_{(1)01} = m_{(2)01}$ the structure would not change, such that for the weights we obtain $w_{21}/w_{22} = w_{11}/w_{12}$. There is no doubt that chain indices are more difficult to handle than direct indices, whenever aggregation or disaggregation is needed, and a change in the structure has to be accounted for. On the other hand in the opinion of “chainers” it is precisely the fact that this change is explicitly taken into account which is responsible for the alleged superiority of chain indices.

The generalization of eq. 4.3.2 is given by:

$$m_{0t} = \left(w_{11} m_{(1)01} + w_{12} m_{(2)01}\right) \cdot \dots \cdot \left(w_{t1} m_{(1)t-1,t} + w_{t2} m_{(2)t-1,t}\right)$$
$$= m_{01} \cdot \dots \cdot m_{t-1,t}, \tag{4.3.5}$$

where the establishment of weights becomes more and more difficult as t goes on

$$w_{1i} = \frac{A_i}{\sum A_i}, \qquad w_{2i} = \frac{A_i + Z_{1i}}{\sum (A_i + Z_{1i})},$$

$$w_{3i} = \frac{A_i + Z_{1i} + Z_{2i}}{\sum (A_i + Z_{1i} + Z_{2i})}, \ldots, w_{ti} = \frac{A_i + \sum\limits_{\tau=1}^{\tau=t-1} Z_{\tau i}}{\sum\limits_i (A_i + \sum_\tau Z_{\tau i})}$$

To demonstrate the difficulties in calculating indirectly instead of directly an example might be useful (ex. 4.3.1 carried on in ex. 4.3.2). It should be borne in mind, however, that in this case (change of a stock due to net inflows [increases]), the changing weights are uniquely determined by the sectoral (partial) links, a condition *not* given in case of price indices.

Example 4.3.1

Given $A_1 = 100$ and $A_2 = 200$ such that $w_{11} = 1/3$ and $w_{12} = 2/3$ as a starting point and the following sectoral links $m_{(i)01}$, $m_{(i)12}, \ldots$

i	$m_{(i)01}$	$m_{(i)12}$	$m_{(i)23}$	$m_{(i)34}$
1	1.2	1.3	1.2	1.1
2	1.1	1.2	1.1	1.3

It is much more comfortable to calculate relatives $m_{(i)0t}$ and aggregate them according to

$$m_{0t} = \frac{1}{3} m_{(1)0t} + \frac{2}{3} m_{(2)0t}$$

than to calculate the varying weights for the links and aggregate links using these weights. This yields the following results:

i		$m_{(i)02}$	$m_{(i)03}$	$m_{(i)04}$
1	1.2	1.56	1.872	2.059
2	1.1	1.32	1.452	1.888
m_{0t}	1.133	1.4	1.592	1.945

t	1	2	3	4
w_{t1}	0.33	0.3529	0.3731	0.3349
w_{t2}	0.67	0.6471	0.6269	0.6651
$m_{t-1,t}$	1.133	1.2352	1.1373	1.2330

It is of course clear that for example $1.133 \cdot 1.2353 = 1.4$ etc. ◀

Example 4.3.2

A modification of ex. 4.3.1 is made in the way, that it is assumed that three sub–indices are available, as for example

i	$P^{LC}_{(i)1}$	$P^{LC}_{(i)2}$	$P^{LC}_{(i)3}$	$P^{LC}_{(i)4}$
1	*1.2*	*1.3*	*1.2*	*1.1*
2	*1.1*	*1.2*	*1.1*	*1.3*
3	*0.9*	*1.4*	*1.1*	*1.0*

In order to compile an overall link P^{LC}_t *or an index (chain)* $\bar{P}^{LC}_{0t}$*, or a link or a chain comprising* only a part of the subaggregates, *for example* $i = 1$ *, and* $i = 3$ *it is not sufficient to know the changing weights* w_{ti} *for each subaggregate* i *in the course of time 0, 1, ... (for all* $t = 0, 1, \dots$*). You virtually have to redo the complete calculation (multiplication) for the aggregates with a possibly different composition. The situation would be much easier if the figures above would denote successive growth factors* $P^L_{(i)0,t}/P^L_{(i)0,t-1}$ *instead of* $P^{LC}_{(i)t}$*. Under the assumption of* $w_{11} = p_{10}q_{10}/\sum p_{i0}q_{i0} = 0.3$*,* $w_{12} = 0.3$*, and* $w_{13} = 0.4$ *we obtain*

aggregation	P^L_{01}	P^L_{02}	P^L_{03}	P^L_{04}
overall index	*1.05*	$1.19 \cdot 1.05 = 1.25$	$1.25 \cdot 1.13 = 1.41$	$1.12 \cdot 1.41 = 1.51$
*partial index**	*1.03*	$1.25 \cdot 1.03 = 1.40$	$1.14 \cdot 1.40 = 1.60$	$1.04 \cdot 1.60 = 1.66$

* *sector 2 excluded*

◄

Calculations of this kind would require knowing more details than just $w_{11} = w_{12} = 0.3$, and $w_{14} = 0.4$ in case of chain indices. To sum up

> To calculate *chain* indices at point t in time with various compositions of subaggregates not only links in a breakdown into the appropriate subaggregates (sectors, components) are required, but also data on the changing weights. Furthermore in general the full calculation for all intermediate points in time has to be done. For the majority of users these conditions are usually not given. By contrast the handling of aggregation and disaggregation of *direct* indices, esp. those with *constant* weights is much more convenient.

4.4 Are the most up–to–date weights automatically the best weights?

All adherents to the chain index idea share the deeply rooted prejudice that the most recent weights are also the most "relevant" and most "representative" weights. Whether or not valid, this argument has at any rate some measure of plausibility. The idea of chain indices owes much of its suggestiveness and attractiveness to this alleged automatic follow up of dynamics, progress and modernism.

Once such notions are endowed with positive emotional charge there is also a strong temptation to prefer chain indices to "fixed based" indices if not solely, so at least predominantly on the basis of some vague idea of "in" and "out". It simply appears to fit better to our days to adjust weights for changes in "relevance" or "importance" of goods rather than to keep a "basket" constant in order to make "pure" comparisons. Though having an intuitive appeal "representativeness" (or "representativity", we found both expressions in this context) is, however, an idea far from being clear and coherent, let alone the underlying assumption that most recent weights are ipso facto also the most representative ones. This section tries to

1. explore some of the less obvious problems with representativity we are all too ready to overlook, and

2. to show that chain indices are not logically the only solution allowing for a permanent updating of weights.

a) Assumptions needed to equate "last observed" with "most representative"

The reason for preferring a representative basket to a non–representative one is usually taken from the concept of "real income". It is certainly not "fair" to explain to a poor man that the purchasing power of "his" money has risen because prices of some distinctively luxury articles like caviar, jewels, sailing–ships or so have dropped. By the same token we may legitimately conclude that a non–smoker is not affected by rising prices of cigarettes. It should be noted that in both examples suitability of weights is decided on the basis of *needs and* underlying *utility* functions. We are permitted, or even obliged to exclude luxury articles from the basket of the poor or cigarettes from the basket of the non–smoker respectively, because such goods are not within the scope of goods such consumers find necessary, desirable or affordable.

In order to translate this idea of a fair measurement of real income, "fair" with respect to the particular type of consumer in question, into the context of measuring consumer prices on a regular basis, some assumptions have to be made which are *not* selfevident. The argument developed so far was mainly

- static (poverty and non–smoking remains at least for a while), and

- applies to different consumers and focuses on *their* real income and *their* preference order.

To assess representativity in such a case calls for calculation of a variety of indices, each of them on the basis of a specific basket, and to choose that index for a particular consumer where the basket of the index fits best to what the consumer would have consumed, if he had been *able* to *maximize* his utility function.[213] It should be noted, however, that

[213] This leads us to the type of microeconomic reasoning known as "economic theory of index numbers". To follow this argument into the details of the microeconomics of a utility maximizing household is beyond the scope of this book, however.

> The situation is different, when we consider dynamic comparisons applied to an average household at different points in time rather than static comparisons between consumers where the choice of the appropriate "basket" may be justified in terms of utility.

Even in the first case (static, inter–household comparisons) the notion of "representativity" is far from being clear. Anyway it can *not*, however, and should not simply be transferred from this framework (assigning a basket to a consumer) to quite a different one. In the case of chain indices it is usually assumed that the most recently observed basket is also the most "relevant" or "representative" one (though – at least to our knowledge – there is no agreed-upon measure of the degree of "relevance" or "representativity"), but it is rarely ever made explicit what justifies the equating of updateness to representativity.

In the dynamic situation there is no choice that can and should be made among a multitude of baskets at a time, and hence no basis of assessing representativity in terms of mean and variance calculated over a variety of consumers. There is simply one single basket observed at a time and this basket is assumed to be more representative than baskets of the past simply because it is the one more recently observed.

Under which conditions would such conclusions be tenable? We found mainly two conditions, set out in literature, but usually tacitly assumed in the chain index methodology (there may well be more than these two objections, however):

1. the actual observed consumption structure is the result of *voluntary* decisions made by consumers, enjoying a real income by and large the same in 0 and in t, and

2. the choice is *not restricted*, it can be made almost immediately, and the variety among which a choice can be made is not altered by activities on the *supply* side such that consumers at time t have chosen qualities and quantities q_t although they could also have chosen q_{t-1} (there should be at least some basis for distinguishing choice of q_t because

    - q_t was *preferred* to q_{t-1} as opposed to a choice of q_t because
    - q_{t-1} was no longer available).

It would clearly not be a fair gauge of inflation, if actual observed consumption patterns were taken in a period of declining real incomes. Attempts to evade rising prices by reduction of consumption *forced by reduced income* would produce a basket which might be "representative" in the sense that many or most of the households have moved to this kind of consumption in response to their decreased incomes. But would this basket automatically constitute the correct basis for comparing prices?

The notion of the "true cost of living index" (COLI) in "the economic theory" approach would avoid this ambiguity, because in such a situation the "new" basket would simply not represent the same utility level. But this is a theoretical argument, not necessarily empirical, and – what is more important – in general not easy to verify empirically.

Strictly speaking the idea of making use of the most representative basket implicitly assumes that consumers are able to achieve the same utility level throughout all changing conditions and that all changes in their consumption patterns are voluntary choices, not impaired by changes in their economic situation or by activities on the supply side such that changes of the actual basket reflect *prompt* changes in

- the *taste* (preferences) and
- in response to changes in the *structure* of prices *only*.

In reality, however, reactions are not always sufficiently prompt, and we therefore partially observe even $P^{P}_{0t} > P^{L}_{0t}$ (positive correlation!) as it is frequently the case with housing or expenditure on transportations (consumers cannot react to rising fuel prices for example by immediately buying a car with lower fuel consumption).

On the other hand, less money spent on the purchase of particular goods not necessarily indicates a voluntary substitution as a result of rational response to rising prices. It is possible that the good no longer exists, but is replaced by a new variant which may well be cheaper, but also no longer qualitatively equivalent. In pursuit of maintaining the previously attained utility level households might even be willing to pay a higher price if only the old commodity would still be available. This is a situation to which microeconomic reasoning is not well prepared. To compare consumption structures in terms of utility at different times in general requires that all goods are fully available in both periods compared.

There are reasons to call into question the underlying logic of chaining, that the most recently observed actual basket, even though "representative" is also automatically the most appropriate one for providing weights. Interestingly these doubts have to do with appearance of new and disappearance of old goods (or varieties). This is a situation in which

- chain indices are supposed to be superior, and at the same time
- the underlying assumptions of equating representativity with superiority are most difficult to verify.

Theoretically reference to the same utility level (as *assumed* in "economic theory" index theory) seems to be superior to simply referring to the most updated *actual* basket as the basis for an adequate weighting scheme, as done in chain index theory. Conspicuously this chain index theory pretends to offer the best solution in dealing with precisely such situations, like the disappearance of goods, appearance of completely new products to serve new "needs", incomparability of quality etc. which on the other hand also casts some doubt on the assumption that at any time the observed basket is also representing the same utility level and the most appropriate basket to compare prices.

b) Representativity over which interval in time?

Given we could prove that q_t as opposed to q_{t-1} or so is the most representative basket at time t, what then is the most representative basket for an interval covering a *number* of periods. Again the situation is such that an argument applying to an individual link is not automatically valid for the chain. It is altogether surprising that though say q_{t-1} may be representative at time t it will in general no longer be so at time $t+1, t+2$ and so on. A chain index comparison of prices p_t with prices p_0 is not based on most representative weights at either 0 or t, but rather on *all* weights *within* this interval.

We often hear the criticism that weights q_0 in P^L_{0t} become *progressively* irrelevant with the passage of time but in the same sense in t weights $q_{t-1}, q_{t-2}, q_{t-3}$ should also be (in this order) "progressively irrelevant". Why not delete them? In the chain index formulas such "old" weights enter the formula when they are new, but when they grow old they are not replaced, but rather retained in the formula.

Why not calculate a series of direct Paasche indices, where in each period t *all* prices, that is those of t as well as those of 0 are weighted by no other quantities than just the then most recent quantities q_t? Unlike chain indices P^P_{0t} will with each new t remove all "old" weights automatically. To the following question:

> When the most important aspect of index compilation is to make use of the most recently observed quantities as weights, why then should we not prefer a direct Paasche index over a chain index?

we never heard a satisfactory answer from "chainers". The answer can hardly be that in t weights q_t needed to calculate P^P_{0t} are not yet available, but rather weights q_{t-1}, which are sufficient to calculate the link P^{LC}_{0t} only. The difference is perhaps negligible. The only satisfactory answer seems to be that the summation over commodities taking place in P^{LC}_t as compared with P^{LC}_1, may well refer to a substantially different collection of goods. Hence

> As to the "representativity" in the sense of updateness of weights $\bar{P}^{LC}_{0t}$ does not enjoy advantages over P^P_{0t}, on the contrary $\bar{P}^{LC}_{0t}$ is affected by weights relating to periods prior to t, whilst P^P_{0t} is not. The "advantage" is that in $\bar{P}^{LC}_{0t}$ no care has to be taken for the comparability of prices in t with those in 0, whereas in P^P_{0t} an attempt has to be made that summation in $\sum p_t q_t$ (numerator of P^P_{0t}) and $\sum p_0 q_t$ in (denominator) takes place over the same selection of commodities.

An alternative to P^P_{0t} would be to choose those weights which are "multiple" weights:

a) in the sense of an average of q_0 and q_t (see fig. 1.3.2), or of

b) referring to some mid–period of the time interval.

Table 4.4.1: Price index using mid–point weights

t	formula	or equivalently
1	$\tilde{P}_{01} = \frac{1}{2}\left(1 + \frac{\sum p_2 q_1}{\sum p_0 q_1}\right)$	$\frac{1}{2}\left(P^L_{00} + P^L_{02}P^P_{01}\right) = \frac{1}{2}\left(\tilde{P}_{00} + \tilde{P}_{02}\right)$ *
2	$\tilde{P}_{02} = \frac{\sum p_2 q_1}{\sum p_0 q_1}$	$\frac{P^L_{02}}{Q^L_{01}} = P^L_{12}P^P_{01}$
3	$\tilde{P}_{03} = \frac{1}{2}\left(\frac{\sum p_3 q_1}{\sum p_0 q_1} + \frac{\sum p_3 q_2}{\sum p_0 q_2}\right)$	$\frac{1}{2}\left(P^L_{13}P^P_{01} + P^L_{23}P^P_{02}\right)$
4	$\tilde{P}_{04} = \frac{\sum p_4 q_2}{\sum p_0 q_2}$	$P^L_{24}P^P_{02}$
5	$\tilde{P}_{05} = \frac{1}{2}\left(\frac{\sum p_5 q_2}{\sum p_0 q_2} + \frac{\sum p_5 q_3}{\sum p_0 q_3}\right)$	$\frac{1}{2}\left(P^L_{25}P^P_{02} + P^L_{35}P^P_{03}\right)$

* $\tilde{P}_{00} = 1$

As to the second alternative we might think of a series like $\tilde{P}_{0t}$ as defined in tab. 4.4.1. Note that such an index would no longer "suffer" from "old" weights q_0, as it is based on one or two weights (i.e. on q_1, q_2, q_3 in periods 2, 4, 6 respectively or on an average of two baskets at uneven periods t). Such an index requires a constant updating of weights, but still consistently compares t with 0 without multiplying links. It is true that the distance between the mid–point and the end–point of the interval increases as the interval becomes longer, such that weights in the index design $\tilde{P}_{0t}$ will become older[214], but unlike the chain design in $\bar{P}_{0t}$ we also get rid of the very old weights (between 0 and the midpoint) which still influence the chain index.
This is not to advocate a somewhat curious index like $\tilde{P}_{0t}$ or so (see below), the point we wish to make is simply:

> Even though we would follow the idea that weights should account for changing consumption habits and should be updated more frequently than hitherto in the Laspeyres framework (to have for example a continuous updating) a chain index would not *necessarily* be the only reasonable index to serve that purpose.

Some other alternatives - again somewhat curious and with (by now) no relevance for official statistics at all - are sometimes discussed in literature and will be presented subsequently. They are inspired by another aspect of weighting in case of chain indices.

214 As we have seen, the implicit equating of "last observed" to "most representative" is not beyond doubt.

c) Accounting for the development of weights in all intermediate periods

It is sometimes maintained that not only accounting for the most recent basket, but also for the way in which this basket evolved from previous ones is one of the advantages of chain indices. There are some attempts made to design index functions that expressly make use of *all* baskets of an interval between 0 and t (and which will thereby also be path dependent as chain indices are). It can be seen that there are some relations between these indices and $\tilde{P}_{0t}$ in tab 4.4.1. In sec. 4.1 we considered the $(T+1) \times (T+1)$ matrix $\mathbf{P}$ of indices, defined for say $T = 2$ as follows:

$$\mathbf{P} = \begin{bmatrix} P_{00} & P_{01} & P_{02} \\ P_{10} & P_{11} & P_{12} \\ P_{20} & P_{21} & P_{22} \end{bmatrix} .$$

Rescaled to the new base 1 we get the matrix:

$$\mathbf{P}^* = \begin{bmatrix} 1 & P_{01}/P_{00} & P_{02}/P_{00} \\ 1 & P_{11}/P_{10} & P_{12}/P_{10} \\ 1 & P_{21}/P_{20} & P_{22}/P_{20} \end{bmatrix} ,$$

which is the starting point for some indices designed for "multilateral" comparisons. DIEWERT (1993), p. 55f mentioned for example a method Irving Fisher discussed (calling it "blend system"), by which a "multiperiod aggregate price level" to compare period 2 with 0 called $\breve{P}_{02}$ is gained by averaging the entries into the third column of matrix $\mathbf{P}^*$ yielding

$$\breve{P}_{02} = \frac{1}{3}\left(P_{02} + \frac{P_{12}}{P_{10}} + \frac{P_{22}}{P_{20}}\right) \tag{4.4.1}$$

and for example

$$\breve{P}_{03} = \frac{1}{4}\left(P_{03} + \frac{P_{13}}{P_{10}} + \frac{P_{23}}{P_{20}} + \frac{P_{33}}{P_{30}}\right) \tag{4.4.1a}$$

and so on.
Substituting P^L_{0t} for P_{0t} gives, due to identity of P^L_{0t} and $P^L_{t0} = \left(P^P_{0t}\right)^{-1}$

$$\breve{P}_{02} = \frac{1}{3}\left(P^L_{02} + P^P_{01}P^L_{12} + P^P_{02}\right) = \frac{1}{3}\left(P^L_{02} + \tilde{P}_{02} + P^P_{02}\right) , \tag{4.4.2}$$

and

$$\begin{aligned} \breve{P}_{03} &= \frac{1}{4}\left(P^L_{03} + P^P_{01}P^L_{13} + P^P_{02}P^L_{23} + P^P_{03}\right) \\ &= \frac{1}{4}\left(P^L_{03} + 2\tilde{P}_{03} + P^P_{03}\right) , \end{aligned} \tag{4.4.2a}$$

and since we have the following general principle to construct such indices

$$\breve{P}_{0t} = \frac{1}{t+1} \sum_{k=0}^{t} P^{P}_{0k} P^{L}_{kt}. \tag{4.4.3}$$

The results $\breve{P}_{02}, \breve{P}_{03}, \ldots$ (as well as $\tilde{P}_{02}, \tilde{P}_{03}, \ldots$) will of course differ in general from $\bar{P}^{LC}_{02}, \bar{P}^{LC}_{03}, \ldots$. More important: successive blend–system–indices and chain indices are gained recursively in a different way,

multiplicative system
$\bar{P}^{LC}_{03} = \bar{P}^{LC}_{02} P^{L}_{23}$
$\bar{P}^{LC}_{04} = \bar{P}^{LC}_{03} P^{L}_{34}$

additive system
$\breve{P}_{03} = \frac{1}{4}\left[3\breve{P}_{02} + \left(P^{L}_{03} - P^{L}_{02}\right) + P^{P}_{01}\left(P^{L}_{13} - P^{L}_{12}\right) + P^{P}_{02}\left(P^{L}_{23} - P^{L}_{22}\right) + P^{P}_{04}\right]$
$\breve{P}_{04} = \frac{1}{5}\left[4\breve{P}_{03} + \left(P^{L}_{04} - P^{L}_{03}\right) + P^{P}_{01}\left(P^{L}_{14} - P^{L}_{13}\right) + \cdots + P^{P}_{03}\left(P^{L}_{34} - P^{L}_{44}\right) + P^{P}_{04}\right]$

The purpose of such considerations is to show that constructing an index capable of comparing periods 0 and t not directly (by accounting for 0 and t only) but indirectly (by accounting for all intermediate periods $1, \ldots, t-1$ in addition to 0 and t) does not necessarily mean that links have to be *multiplied*. Indices like blend–system–indices or other formulas to be presented subsequently, however meaningful or pointless they may be, aim at ensuring transitivity. By averaging all elements of a column of **P** we get columns made up of identical expressions, and hence a vanishing determinant $|\mathbf{P}|$. Substitution of chain indices $\bar{P}^{C}_{0t}$ for the elements P_{0t} in **P** would serve the same purpose.

Another method discussed by Fisher is known as "broadened base system" ("BB" for short; base in the sense of *weight* base), and can also be seen as a generalization of index formulas using average weights, like indices of Marshall–Edgeworth (ME)[215] or Walsh:

$$P^{BB}_{01} = \frac{\sum p_1 \frac{1}{2}\left(q_0 + q_1\right)}{\sum p_0 \frac{1}{2}\left(q_0 + q_1\right)} = P^{ME}_{01} \tag{4.4.4}$$

$$P^{BB}_{02} = \frac{\sum p_2 \frac{1}{3}\left(q_0 + q_1 + q_2\right)}{\sum p_0 \frac{1}{3}\left(q_0 + q_1 + q_2\right)} = \frac{P^{L}_{02} + P^{L}_{12} V_{01} + V_{02}}{1 + Q^{L}_{01} + Q^{L}_{02}}. \tag{4.4.5}$$

[215] DIEWERT (1993), p. 56.

In a similar manner the well known index of Walsh

$$P_{0t}^{W} = \frac{\sum p_t \left(q_0 q_t\right)^{1/2}}{\sum p_0 \left(q_0 q_t\right)^{1/2}}$$

has been generalized to

$$P_{0t}^{W*} = \frac{\sum p_t \left(q_0 q_1 \ldots q_t\right)^{1/t+1}}{\sum p_0 \left(q_0 q_1 \ldots q_t\right)^{1/t+1}}.$$

The problem with such rather curious ideas is that they are mainly conceived for interspatial comparisons (especially for different countries) such that – by contrast to chaining – the sequence 0, 1, . . . , t is of no relevance.

d) Fixed basket as a model

We now return to Martini's position according to which things incomparable directly (and by definition) become comparable indirectly by some mysterious process (see sec. 4.1). There is nothing wrong with choosing a kind of "theory of comparability" as a starting point, but the conclusions following from this point may well contradict directly those Martini has drawn. In our view – agreeing with this starting point – the consequence should rather be that we should aim at making "pure" comparisons. By this is meant that situations to be compared should differ in *one* aspect only, otherwise we run into difficulties of interpretation.

In what follows we try to show that much of statistics is devoted to precisely this task of making pure comparisons.

Our task as statisticians is not to identify comparability or non–comparability[216] of data and then possibly to take data as they are or reject them. Our task is rather to present results that can be interpreted in terms of factors influencing the result, and in so doing we have to make transformations, adjustments, construct models underlying the data, keep disturbing sources of variation under control and present data in various breakdowns into subaggregates. The common feature of all these measures is to facilitate an interpretation of the results in terms of "pure" comparison. We always try to identify sources of variation in statistics. A statistical result which is affected by an unknown mix of a *number* of influences is in general of limited value only. Most of our analysis consists in elimination of "disturbances", in separating or isolating systematic influences from random variation, removal of contamination, making variables commensurate by "standardisation", making various sorts of "adjustments" to ensure that like is compared with like. Estimation of models in which we have a *small* number of causal relationships only which might constitute the data–generating–process also serves this purpose.

[216] Comparability is not "given" or absent. Conspicuously a method *not* mentioned by Martini is to make adjustments (as for example the well known adjustments in case of changes in quality) in order to enhance comparability.

Interestingly axioms also follow this idea. Under the assumption of isolated changes, stylized and simplified processes we are able to distinguish meaningful index formulas from meaningless ones. Axioms are powerful tools to gain a priori insights by construction of situations ideal for comparisons. They are assumptions simple enough to allow inferences from (fictitious) data to possible results. We cannot infer "results" from arbitrary data.
There is not much we can say when all variables are allowed to vary. When all prices *and* all quantities may vary *somehow* there is not much we can learn from the fact that different index formulas applied to such data will produce different results.
This isolation of determinants is precisely the idea of the "constant basket" approach, and it is in line with making *pure* comparisons. The chain index, on the other hand, is an index design which is characterized by the fact that sources of variation are all but under control.

5 Deflation with chain indices

In this chapter some unfavorable aggregative properties of Fisher indices in both versions, direct (P^{F}_{0t}) as well as chained ($\bar{P}^{FC}_{0t}$), to which reference was already made in sec. 2.3 and 4.4c will be discussed in more detail. The SNA recommended Fisher indices instead of Paasche indices for deflation. It is appropriate to deal with a critique of the *direct* Fisher index first (sec. 5.2), and then move on to the chained version as recommended by the SNA (sec. 5.3). Inability of chain indices to ensure structural consistency in deflation is a well known shortcoming. It is therefore interesting to see how we should best cope with such defects according to the SNA (see sec. 5.4).

5.1 The SNA recommends chained Fisher indices

As repeatedly mentioned already the SNA 93[217] recommends the use of Fisher type chain indices. In full the recommendations (para. 16.73) read as follows:

> (1) the preferred measure of year to year movement of real GDP is a Fisher volume index, changes over longer periods being obtained by chaining: that is, by cumulating the year to year movements;
> (2) the preferred measure of year to year inflation for GDP is therefore a Fisher price index, price changes over long periods being obtained by chaining the year to year price movements: the measurement of inflation is accorded equal priority with the volume measurements;
> (3) chain indices that use Laspeyres volume indices to measure movements in real GDP and Paasche price indices to year to year inflation provide acceptable alternatives to Fisher indices...

In short the recommendations express a general belief of

- the chain principle being superior to the direct binary (two situations) comparison,
- universal applicability of a price index formula for both, measurement of price levels as well as deflation of aggregates, and finally a firm belief in
- the superiority of Fisher's "ideal" index over traditional formulas (Laspeyres, Paasche), which are qualified as second best solutions only in the SNA.

In the opinion of the SNA to continuously update weights as soon as possible is the overwhelming advantage of chain indices which outweighs all shortcomings of chain indices, including all defects in the framework of deflation. The idea is that it may be even less reasonable to maintain the same (weight) base in the case of deflation as in the case of measuring the consumer price level. The SNA pointed out that industrial

[217] (System of National Accounts) Revised edition 1993.

production is usually very adaptable to changes in prices. A process of production which is efficient at one set of prices may not be very efficient at another set of relative prices. The aggregate $\sum p_0 q_t$ as opposed to $\sum p_0 q_0$ or to $\sum p_t q_t$ will not only be improbable, but may even be (technically / economically) *impossible*. Note, that the same argument is also used as an explanation for the sometimes erratic and obviously counter-intuitive results (e.g. negative real value added [VA]) the double deflation method to calculate real VA may yield in practice[218].
To summarize, the position of the SNA 93 is:

1. the purpose of deflation is another reason to reject the idea of maintaining a fixed (weight) base (or fixed "basket") and a strong case for adoption of the chain index principle, and

2. which version of a chain index should be taken (the SNA prefers the *Fisher* version of a chain index $\bar{P}^{FC}_{0t}$ to other forms as for example $\bar{P}^{PC}_{0t}$ or $\bar{P}^{LC}_{0t}$) should be decided mainly on formal grounds (see below and also sec. 3.2c).

5.2 Deflation with direct Fisher indices, a bad solution

a) Fisher deflation, additivity and quantity movement

In sec. 2.3 a number of aggregation properties were introduced and some criteria of "good" deflation were derived from the aim to define "volumes" as substitutes for "quantity". In what follows, an attempt is made to show that deflation with a *direct* Paasche deflator P^{P}_{0t} (Paasche deflation) is superior to deflation with a direct P^{F}_{0t} deflator (Fisher deflation) and that P^{F}_{0t} (let alone the chain version $\bar{P}^{FC}_{0t}$) has severe disadvantages with respect to

1. both aggregation problems as distinguished in sec. 2.3b, that is aggregation of the index function and consistent aggregation of volumes (**A1** and **A2** in fig. 2.3.2), and

2. some additional criteria introduced in sec. 2.3e and found by interpreting "volume" in terms of "quantity".[219]

Traditional deflation with (direct) Paasche price indices will meet all these criteria which appear reasonable in view of the need of aggregating / disaggregating and of interpreting volumes in terms of quantities, which are otherwise unobservable. But

[218] If prices p_0 would prevail, the production process that created q_t would perhaps not be chosen. Hence deflation will necessarily always imply a somewhat fictitious combination of prices and quantities.

[219] To recall some criteria, it was found useful that 1. the change of volumes should equal the change of quantities when all prices change *uniformly* at the same rate (which was called reflection of quantity movement, RQM), and 2. much more ambitious, that volumes should be constantly weighted sums of quantities such that the movement of volumes will not be affected by a *non–uniform* change of prices (this criterion was called pure quantity comparison, PQC, and it is in fact equivalent to linearity [additivity] in quantities of the resulting quantity index).

deflation using P_{0t}^{F} (and all the more $\bar{P}_{0t}^{FC}$) as deflator will violate most if not all these criteria.
In sec. 2.3e the following "cases" were distinguished, in which quantities change with different rates and prices change

- unanimously with the *same* rate λ (case 12), and
- prices (like quantities) changing with *different* rates (case 22).

Note that it is only in case 22 that Paasche and Fisher deflation will yield different results.
The point of interest then is not only how volumes will reflect the movement of values, prices and quantities, but also the structure of these volumes as compared with the structure of quantities. Assume quantities can be summated such that the *quantity* movement is given by

$$M_{0t} = \frac{\sum q_t}{\sum q_0} = \sum \frac{q_t}{q_0}\left(\frac{q_0}{\sum q_0}\right) = \sum \frac{q_t}{q_0} \cdot m_0, \tag{5.2.1}$$

by contrast to

$$Q_{0t}^{L} = \sum \frac{q_t}{q_0}\left(\frac{p_0 q_0}{\sum p_0 q_0}\right) = \sum \frac{q_t}{q_0} \cdot g_0 \tag{5.2.2}$$

measuring *volume* movement. Hence the structure of volumes will in general *not* coincide with the structure of quantities[220], and the difference between Q^L–volumes and quantities is, plausibly, a result of the structure of base period prices (and some additional factors in case of volumes other than the Q^L–type).

Example 5.2.1

Prices and quantities of four commodities, A_1, A_2, B_1 *and* B_2 *classified into two groups,* A *and* B *are given in the base period 0 and (in two alternatives) in the current period* t *as follows:*

	base period (0)		*period t: case 12*		*period t: case 22*	
	prices	*quantities*	*prices*	*quantities*	*prices*	*quantities*
A_1	*30*	*70*	*45*	*84*	*30*	*84*
A_2	*50*	*30*	*75*	*48*	*40*	*48*
B_1	*90*	*20*	*135*	*36*	*81*	*36*
B_2	*120*	*30*	*180*	*24*	*168*	*24*
		prices	*uniform change*		*non–uniform change*	
		quantities	*non–uniform change*		*uniform change*	

220 A (trivial) case of both structures being identical is of course given when all base period prices are the same. In general the weight assigned to an increase in quantity will be higher if a high priced good is involved and lower in the case of a low priced good (prices of base period).

*We first examine **case 12** where all prices rose by 50%. Due to the uniform change of prices we get $P^F_{0t} = P^P_{0t} = 1.5$ and $Q^F_{0t} = Q^L_{0t} = 1.2267$ and the resulting volumes are the same whichever deflator, direct Paasche or direct Fisher was used. The results are:*

	base period		current period		indices	
	$\sum p_0 q_0$	$\sum q_0$	$\sum p_0 q_t$	$\sum q_t$	Q^L_{0t}	M_{0t}
A	3600	100	4920	132	1.3667	1.32
B	5400	50	6120	60	1.1333	1.20
sum	9000	150	11040	192	1.2267	1.28

The group of commodities for which the rise of quantities was above average (group A by 32%) also experienced an increase in volume above average (36.7% > 22.7%). Conversely for a group, like B, in which quantities rose below average increase of volumes is also below average (13.3%). The structure of volumes (of $\sum p_0 q_0, \sum p_0 q_t$) also changed in accordance with the dynamics of quantities:

	$\sum p_t q_t = 1.5 \cdot \sum p_0 q_t$	structure	$\sum p_t q_0 = 1.5 \cdot \sum p_0 q_0$	structure
A	7380 = 1.5· 4920	0.446	5400 = 1.5 · 3600	0.4
B	9180 = 1.5· 6120	0.554	8100 = 1.5 · 5400	0.6
sum	16560 = 1.5· 11040	1	13500 = 1.5 · 9000	1

The share of group A, with respect to volumes, increased (0.446 > 0.4), and the share of group B decreased (0.554 < 0.6). Because in A the quantities rose above average while in B quantities rose below average. This is in line with what we expect deflation should produce.

*We now examine **case 22**: It should be noticed that this case does not differ from the preceding case as far as the* quantity *movement is concerned. Thus, not surprisingly, the results of Paasche deflation remains the same (indicated by shadows). But this is not true for "Fisher volumes" resulting from deflation with a (direct) Fisher price index:*

	value	P^P_{0t}	P^F_{0t}	Paasche volume	Fisher volume
A	4440	0.9024	0.9095	4920	4881.7
B	6948	1.1353	1.1833	6120	5871.7
sum / average	11388	1.0315	1.0684	11040	10658.6*

* *obtained by deflating 11388 with the overall P^F, but 4881.7 + 5871.7 = 10753.4 $\neq$ 10658.6, hence Fisher deflation violates structural consistency of volumes.*

The example shows that Fisher deflation is affected by differential price movement, while Paasche deflation is not. Moreover the example also makes clear that Fisher deflation violates structural consistency because the sum of (separately) deflated sub–aggregates does not equal the deflated total–aggregate. ◀

For both groups (A and B) Fisher deflation arrives at lower volumes as compared with Paasche deflation. The reason for this is that Fisher deflation amounts to calculating

$$\frac{\sum p_t q_t}{P^P_{0t}}\sqrt{\frac{P^P_{0t}}{P^L_{0t}}}\text{(Fisher volume) instead of }\frac{\sum p_t q_t}{P^P_{0t}}\text{(Paasche volume)}^{221}, \tag{5.2.3}$$

and for group B we have $\sqrt{R_{0t}} = \sqrt{\frac{P^P_{0t}}{P^L_{0t}}} = \sqrt{\frac{Q^P_{0t}}{Q^L_{0t}}} = 0.9594$ as opposed[222] to 0.9922 in the case of group A. The factor $\sqrt{R_{0t}}$ can also be related to the covariance C between price and quantity relatives, as defined in Bortkiewicz's theorem (see sec. 2.1d). According to eq. 2.1.15 we have

$$C = Q^L_{0t}P^L_{0t}\left(\frac{P^P_{0t}}{P^L_{0t}} - 1\right) = Q^L_{0t}P^L_{0t}\left(R_{0t} - 1\right), \tag{5.2.4}$$

such that

$$P^P_{0t} < P^L_{0t},\ R_{0t} < 1,\ C < 0\ \Rightarrow \text{vol (Fisher)} < \text{vol (Paasche)}$$

and

$$P^P_{0t} > P^L_{0t},\ R_{0t} > 1,\ C > 0\ \Rightarrow \text{vol (Fisher)} > \text{vol (Paasche)}$$

holds for the respective aggregates and sub–aggregates.
In ex. 5.2.1 we get for the covariances

	P^L_{0t}	P^P_{0t}	Q^L_{0t}	C	$\sqrt{R_{0t}}$
A	0.9167	0.9024	1.3667	-0.0195	0.9922
B	1.2333	1.1353	1.1333	-0.1111	0.9594
sum / average	1.1067	1.0315	1.2267	-0.0839	0.9654

It is important to note that the result of Paasche deflation does not differ in both cases (12 and 22) of ex. 5.2.1, because the assumptions relating to the *quantity* movement are the same. It is obviously only the quantity change ($q_{i0} \to q_{it}$) that matters, and the influence of different price structures in 0 and t is eliminated. Furthermore the result of Paasche deflation, that is Q^L_{0t} is a pure quantity comparison (PQC) where *constant*[223] prices p_{i0} act as weights for a linear combination of quantities, such that

221 The same result is implied already in eqs. 2.3.9 and 2.3.17.

222 This means that the Fisher volume (5871.7) falls short of the Paasche volume (6120) by 4.06% (as $0.9594 - 1 = -0.0406$), a relatively higher discrepancy as in the case of group A (-0.8% only).

223 For pure quantity comparison (PQC) it is not important to which period the prices of volumes are related, if only the period is the same for a series of volumes. Thus $\sum p_1 q_1, \sum p_1 q_2, \sum p_1 q_3, \ldots$ represents a PQC as well as for example $\sum p_0 q_1, \sum p_0 q_2, \sum p_0 q_3, \ldots$ or $\sum p_2 q_1, \sum p_2 q_2, \sum p_2 q_3, \ldots$. In general PQC seems to fit well to an intuitive understanding of the term "at constant prices" or of "volume" as a substitute for "quantity".

individual quantities q_{it} are unequivocally related to the total aggregate $\sum p_{i0}q_{it}$. The index Q^L_{0t} is a linear (additive) measure of (reference period) quantities because

$$Q\left(\mathbf{q}_0, \mathbf{q}^*_t\right) = Q\left(\mathbf{q}_0, \mathbf{q}_t\right) + Q\left(\mathbf{q}_0, \mathbf{q}^\Delta_t\right) \tag{5.2.5}$$

where $\mathbf{q}^*_t = \mathbf{q}_t + \mathbf{q}^\Delta_t$ holds. With elements q^Δ_{it} of the vector $\mathbf{q}^\Delta_t$ in the case of Q^L_{0t} we get

$$Q^L_{0t}(\mathbf{q}^*_t) = \frac{\sum q_{it}p_{i0}}{\sum q_{i0}p_{i0}} + \frac{\sum q^\Delta_{it}p_{i0}}{\sum q_{i0}p_{i0}} = Q^L_{0t}(\mathbf{q}_t) + Q^L_{0t}(\mathbf{q}^\Delta_t). \tag{5.2.6}$$

The corresponding equation for Q^F_{0t}, the result of Fisher deflation is, however,

$$Q^F_{0t}(\mathbf{q}^*_t) = \sqrt{\left(\frac{\sum q_{it}p_{i0}}{\sum q_{i0}p_{i0}} + \frac{\sum q^\Delta_{it}p_{i0}}{\sum q_{i0}p_{i0}}\right)\left(\frac{\sum q_{it}p_{it}}{\sum q_{i0}p_{it}} + \frac{\sum q^\Delta_{it}p_{it}}{\sum q_{i0}p_{it}}\right)}, \tag{5.2.7}$$

or

$$Q^F_{0t}(\mathbf{q}^*_t) = \sqrt{\left(Q^L_{0t}(\mathbf{q}_t) + Q^L_{0t}(\mathbf{q}^\Delta_t)\right)\left(Q^P_{0t}(\mathbf{q}_t) + Q^P_{0t}(\mathbf{q}^\Delta_t)\right)}, \tag{5.2.7a}$$

which is clearly different from

$$Q^F_{0t}(\mathbf{q}_t) + Q^F_{0t}(\mathbf{q}^\Delta_t) = \sqrt{Q^L_{0t}(\mathbf{q}_t)Q^P_{0t}(\mathbf{q}_t)} + \sqrt{Q^L_{0t}(\mathbf{q}^\Delta_t)Q^P_{0t}(\mathbf{q}^\Delta_t)}$$

and shows that a given result for Q^F_{0t} or an absolute volume gained by (direct) Fisher deflation does not permit any conclusions to the underlying quantity movement, to the process responsible for the result.
Linearity in quantities is useful because it allows the inference from the change of individual quantities (micro–level of individual commodities, right hand side) to the result of deflation (macro–level of aggregates, left hand side) and vice versa

Paasche deflation resulting in Q^L_{0t}	$\longleftrightarrow$	(differential) change in individual quantities

which greatly facilitates an interpretation of volumes in terms of quantities.
By contrast we cannot conclude from Q^F_{0t} to $q_{i0} \to q_{it}$, because Q^F_{0t} reflects both, a (differential) change in quantities ($q_{i0} \to q_{it}$), as well as in prices ($p_{i0} \to p_{it}$):
It is not only the well known fact, that *Fisher volumes are* liable to be *structurally inconsistent* which makes them doubtful, it is also no longer true that the same quantity movement is reflected in the same manner (irrespective of how prices changed)[224]. The impact of (differential) prices is influential, and volumes depend on *both* movements, change in quantities as well as in prices.

[224] Ex. 5.2.1 has shown that a *non–uniform* change of prices (thus a change in the *structure* of prices) will yield Fisher volumes that will hardly allow an interpretation as substitutes for volumes.

The interpretation will even become more tricky when no longer a single binary comparison is made (between 0 and *one* particular t), but rather a *series* of such comparisons between 0 and successive periods t. A sequence of Paasche deflation volumes is given by

$$\sum p_0 q_2, \sum p_0 q_3, \ldots \tag{5.2.8}$$

(the numerators of the Laspeyres quantity indices) whereas the corresponding sequence of Fisher volumes is much more complicated

$$\sum p_0 q_2 \left(\frac{\sum p_2 q_2}{\sum p_2 q_0} \frac{\sum p_0 q_0}{\sum p_0 q_2} \right)^{1/2}, \sum p_0 q_3 \left(\frac{\sum p_3 q_3}{\sum p_3 q_0} \frac{\sum p_0 q_0}{\sum p_0 q_3} \right)^{1/2}, \tag{5.2.9}$$

or

$$\begin{aligned} \sum p_0 q_2 \sqrt{Q_{02}^P / Q_{02}^L} &= \sum p_0 q_2 \sqrt{R_{02}}, \\ \sum p_0 q_3 \sqrt{Q_{03}^P / Q_{03}^L} &= \sum p_0 q_3 \sqrt{R_{03}}, \end{aligned} \tag{5.2.9a}$$

a sequence, the meaning of which does not appear to be easy to understand. It is also by no means evident why this should be called "at constant prices of period 0".
Before showing that things will become even more complicated (and less satisfactory with respect to certain criteria of deflation) in the case of chaining, that is when $\bar{P}_{0t}^{FC}$ is employed, it is useful to shortly make some general remarks to the undeserved nimbus of Fisher's formula.

b) Fisher's "ideal index" is far from being "ideal"

The SNA 93 recommended that (chained) Fisher price indices should be used for both, price level measurement as well as deflation[225], mainly on the following grounds

- price level measurement by P_{0t}^F (or deflation resulting in Q_{0t}^F) is based on more recent quantities as in case of P_{0t}^L (or more recent prices as in case of Q_{0t}^L gained by Paasche deflation),
- time and factor reversibility and other criteria[226] of this kind, like for example the symmetric treatment of periods 0 and t.

These arguments of course also apply (or all the more) to $\bar{P}_{0t}^{FC}$ and Q_{0t}^{FC} respectively. The admiration of the ideal index is mainly inspired by its compliance with reversal

[225] SNA also said that Fisher quantity indices should be used whenever *direct* deflation (i.e. updating of quantities) appears possible as an alternative to *indirect* deflation with Fisher price indices.

[226] As pointed out already factor reversibility is poorly motivated because there is no need in official statistics to solve both problems, deflation and price level measurement with the *same* price index formula. There are many reasons to take two differently constructed indices for these two purposes.

tests and the belief that "if one index measures above the 'true' value and another measures below it then an averaging of the two should yield a result closer to the 'true' value than either component" (PFOUTS (1966)). The foundation of these ideas is weak, however, and "If we abandon the notion of a 'true' index, as has been done by many writers, much of the theoretical appeal of the ideal index is lost" (Pfouts). It is true that P^F satisfies many axioms, if not more than any other well known formula.[227] Furthermore, most of the known "uniqueness theorems" refer to P^F as shown in sec. 2.2e above. But many axioms are controversial, and not related to aspects of interpretation and deflation, and thus the message of uniqueness theorems is often far from clear.

Moreover Fisher's "ideal" index has many serious shortcomings especially when used for deflation[228]. Disadvantages abound with respect to

- the interpretation of volumes in terms of the "quantities", isolation of the "quantity" component in case of Q^F_{0t}, and exposition of a pure "price" component in P^F_{0t} respectively (it is not only a change of quantities / prices to which Q^F_{0t} or P^F_{0t} will react)[229]; and

- there is neither a weighted mean of relatives, nor a ratio of expenditures interpretation [230].

Interestingly the following aggregation requirements are *all* met by traditional (direct) Paasche deflation whilst (SNA recommended) Fisher deflation satisfies *none* of them viz.

- aggregative consistency of the index function[231], and the much weaker equality test (**ET**)[232]

[227] This is one of the main arguments of DIEWERT (1998), p. 48. For him P^F is the best choice because of its relation to the "economic theory" approach in index theory and Diewert's preference for "symmetric" averages ("the base period basket ... is just as valid as the current period basket"), and since P^F satisfies a greater number of tests or axioms than any of its competitors. In some of his articles Diewert counted no less than 20 or so such axioms. He conspiciously gave no mention to problems of deflation, interpretation and understandability or to aspects related to the practice of official price statistics.

[228] Our position is that traditional solutions for deflation are more reasonable and justified, and furthermore they are easier to implement and to understand. The SNA recommendations should be viewed with great suspicion.

[229] As shown above Fisher volumes do not reflect PQC in the sense of elimination of the price movement (or isolation of the quantity movement).

[230] According to PFOUTS (1966), an index should answer a specific question, but "the purpose of the ideal index is at best opaque", and to interpret P^F there is not much left apart from saying, it is the geometric mean of P^L and P^P.

[231] **A1** in fig. 2.3.2, that is sub–indices should be additively consistent among themselves. No simple function exists by which sectoral P^F–type indices can be aggregated to a total P^F–index.

[232] See ex. 2.3.2. According to STUVEL (1989), p. 49 ff. Fisher's index is also unable to pass the (compared to ET) even weaker "withdrawal–and–entry test" (WT). A test becomes "weaker" as additional conditions are introduced in its description. We arrive from ET to WT assuming a subaggregate is consisting of only one single commodity. Once all subaggregates consist of one commodity only, we get proportionality (P), assuming identical price relatives λ. As $\lambda = 1$ we get the even more special condition identity (I).

- aggregation of the volumes[233], and their interpretation in terms of "quantities" (sec. 2.3e).

In short: P^F is not additive (in all connotations distinguished in sec. 2.3) nor are volumes gained by Fisher–deflation structurally consistent (that is in contrast to Q^L the resulting quantity index Q^F is not linear [additive] in quantities q_t).
In view of the defects with respect to criteria like interpretation[234], consistency and comparability, the "ideal" index P^F_{0t} (or Q^F_{0t}) looses much of its attraction. In our view such aspects should come first but they were put last if not ignored in the SNA:

> With respect to criteria established here, and in sec. 2.3 (reflection and isolation of quantity movement, and structural consistency) traditional deflation with direct Paasche price indices P^P seems to be much more justified and intuitively appealing (i.e. in line with "volumes" as substitutes for "quantities") than deflation with direct Fisher indices P^F. All that is said here about the *direct* Fisher price index applies to the *chain* Fisher price index as recommended in the SNA with even greater force.

There is a fundamentally different philosophy underlying this SNA–recommendation as compared with ours:

- the *SNA* emphasized criteria like reversibility, symmetry, "crossing", updating of weights, and avoidance of prices p_{i0} (or quantities q_{i0}) related to a base period long ago in the definition of "volumes" (or in price level measurement respectively), in contrast to

- *our* preferred criteria like "pure" comparisons, structural consistency of volumes, interpretation in terms of quantities, and aggregative properties of index formulas.

The SNA unfortunately decided to recommend the pair P^F_{0t} (deflator) and Q^F_{0t} (index of volumes), or, even worse, $\bar{P}^{FC}_{0t}$ and $\bar{Q}^{FC}_{0t}$ instead of the pair P^P_{0t}, and Q^L_{0t} hitherto preferred. It is difficult to understand this choice and to disentangle the complex of influences that makes a Fisher index rise or decline, and thus to interpret such indices. In the case of *chain* Fisher indices ($\bar{P}^{FC}_{0t}$, $\bar{Q}^{FC}_{0t}$) the interpretation is even more "opaque" (Pfouts). We agree with Pfouts' conclusion that Fisher's index is far from "ideal", and that "this index should be abandoned".[235]

Hence from left to right a test will become weaker: ET → WT → P → I. It is possible to conclude in this order: for example an index to which ET applies will also meet P and I. The converse order of conclusion is not permitted. Thus Fisher's index P^F satisfies only the last (most right) two criteria, P and I, Paasche and Laspeyres indices (P^P and P^L) satisfy all four criteria, and chain indices none of them.

233 **A2** in fig. 2.3.2, that is structural consistency of volumes at various levels of aggregation, or aggregation of *volumes* follows the same rules as aggregation of values.

234 There is a need for empirical figures that can be viewed as effects of some instruments or targets (as for instance the money stock, as measured by M_1, M_2, M_3 and so on). It is also desirable that an index is understandable and will be accepted by the general public.

235 In a more recent survey of the axiomatic approach BALK (1995) reached a similar conclusion: "the evidence for choosing the Fisher price index as the ultimate index is not conclusive".

5.3 Deflation with chain Fisher indices, an even worse solution

An axiom violated in case of deflation with chain indices, in *addition* to all those shortcomings, we have already met in the case of (direct) Fisher deflation is "proportionality in prices and quantities" **PPQ** (see sec.2.3e)[236]. According to PPQ in a situation in which all quantities remain constant $q_{it} = q_{i0}$, $\forall i$, and all prices change identically by λ such that $p_{it} = \lambda p_{i0}$ the quantity index should indicate no change $Q(\mathbf{q}_0, \mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t) = 1$ and the value index should amount to[237]

$$V_{0t} = \frac{\mathbf{p}_t'\mathbf{q}_t}{\mathbf{p}_0'\mathbf{q}_0} = P(\mathbf{p}_0, \mathbf{q}_0, \mathbf{p}_t, \mathbf{q}_0) = \frac{\mathbf{p}_t'\mathbf{q}_0}{\mathbf{p}_0'\mathbf{q}_0} = P_{0t}^L = \lambda, \tag{5.3.1}$$

as $Q = 1$, or

$$Q_{0t} = V_{0t} / \lambda \; . \tag{2.3.13a}$$

More general: if all quantities change identically by ω such that $q_{it} = \omega q_{i0}$, $\forall i$ we shall have $V_{0t} = \lambda\omega$ and $Q_{0t} = \omega$.[238] It is interesting to note that under conditions described above (eq. 5.3.1), that is $q_{it} = q_{i0}$, and $q_{it} = \lambda p_{i0}$ with *necessity*

- a value index $V_{0t} = \mathbf{p}_t'\mathbf{q}_t \big/ \mathbf{p}_0'\mathbf{q}_0 = \lambda$ since V_{0t} is a mere definition[239], and
- *direct* price and quantity indices $P_{0t}^P = P_{0t}^F = \lambda$ and $Q_{0t}^L = Q_{0t}^P = Q_{0t}^F = 1$.

Thus deflation using direct indices, P_{0t}^P, or P_{0t}^F will necessarily comply with PPQ, a criterion which seems to be only plain logic. But a *chain* index deflation (using for example $\bar{P}_{0t}^{FC}$, or also $\bar{P}_{0t}^{PC}$ as a deflator) must *not* pass this (value index preserving) test: hence given $\mathbf{q}_t = \mathbf{q}_0$

$$\bar{Q}_{0t}^{FC} \neq 1 \qquad \text{and} \qquad \bar{P}_{0t}^{FC} \neq V_{0t} \tag{5.3.2}$$

is well possible (provided $t \geq 3$). An example will demonstrate this (ex. 5.3.1). Thus

> Violation of PPQ is (at least) one more deficiency we face when deflation is made with a chained Fisher index, in addition to all those shortcomings we have already met when a direct Fisher index is used (e.g. structurally inconsistent volumes). To fail PPQ is a common feature of all types of *chain* indices not only of the chain *Fisher* price index.

[236] This criterion is also known as value index preserving test (VOGT (1978a), BALK (1995)).
[237] It is assumed that the pair of index functions P, Q will pass the product test.
[238] As usual no change (identity) $\omega = 1$, and thus eq. 5.3.1, is a special case of proportionality.
[239] A value index V_{0t} is, by contrast to a price or quantity index, observation only in the sense that there cannot be any doubt which formula to take to calculate V_{0t}.

Example 5.3.1

Assume prices of both goods, A and B rise uniformly by 50% from 0 to 3 (such that $\lambda = 1.5$), and quantities are the same in 0 and 3 (thus we get $V_{03} = P^L_{03} = P^P_{03} = P^F_{03} = \lambda = 1.5$). Direct *Fisher– and Paasche–deflation, therefore will produce the same result. By contrast, however, deflating using a* chain *Fisher index, $\bar{P}^{FC}_{03}$, results in volumes indicating a decline, despite zero change in quantities. The "logic" or "economic common sense" of this result is difficult to discover. Assume for example the following numerical values:*

	period 0		period 1		period 2		period 3	
good	p	q	p	q	p	q	p	q
A	30	5	40	3	50	2	45	5
B	10	15	5	20	10	13	15	15

As $\sum p_3q_3 = \sum p_3q_0 = 1.5 \cdot \sum p_0q_0$, and $\sum p_3q_2 = 1.5 \cdot \sum p_0q_2$ we have

$$\bar{P}^{FC}_{03} = 1.5 \cdot \sqrt{\frac{\sum p_1q_0 \sum p_2q_1 \sum p_0q_2}{\sum p_2q_0 \sum p_0q_1 \sum p_1q_2}},$$

and

$$\bar{P}^{PC}_{03} = 1.5 \cdot \frac{\sum p_0q_0 \sum p_1q_1 \sum p_2q_2}{\sum p_2q_0 \sum p_0q_1 \sum p_1q_2} = 1.5\frac{\bar{P}^{PC}_{02}}{P^L_{02}},$$

both indices not necessarily amounting to 1.5. The results for links and chain indices are as follows

	P^C_1	P^C_2	P^C_3	chain $(\bar{P}^C_{03})$
Laspeyres	275/300	350/220	285/230	665/368 = 1.807 > 1.5
Paasche	220/290	230/145	450/400	2277/1682 = 1.354 < 1.5
Fischer	0.834	1.589	1.181	$\bar{P}^{FC}_{03} = \sqrt{2.4463} = 1.564$

Dividing $\sum p_3q_3 = 450$ by $\bar{P}^{FC}_{03}$ gives a chain–index Fisher volume *of 287.71, hence a* reduction *by 4.1% compared with $\sum p_0q_0 = 300$ (as $\bar{Q}^{FC}_{03} = 0.959$). Curiously we observe a reduction although quantities remain unchanged, and prices as well as values were rising by 50%. Likewise we get $\sum p_3q_3 / \bar{P}^{PC}_{03} = 332.41$, that is an* increase *by 10.8% (since $\bar{Q}^{LC}_{03} = V_{03}/\bar{P}^{PC}_{03} = 1.108$) from deflation with a Paasche* chain *index.* ◀

To understand this result from the previous example it is useful to express the corresponding quantity index in the case of multiplying three or four links, and to compare this with the result obtained by other deflation methods (see tab. 5.3.1).[240] The table shows that direct Paasche volumes (the traditional deflation method) simply form the sequence $\sum p_0q_3, \sum p_0q_4, \ldots$. It appears legitimate to speak of values "at constant

[240] It is interesting that there is no such table in SNA in which "volumes" or quantity indices, like $\bar{Q}^{FC}_{0t} = V_{0t}/\bar{P}^{FC}_{0t}$ or so, as a result of deflation are spelled out in detail, e.g. presented in terms of all aggregates involved, as done in this table.

prices of period 0” in this case. Things are much more difficult to understand and interpret, however, in the case of the alternative “volumes”, for example when a *chain* Paasche index (instead of a direct one) is taken as a deflator. We then have the sequence:

$$\sum p_0 q_3 \frac{P^{P}_{03}}{P^{PC}_{03}} = \sum p_0 q_3 \left(D^{PP}_{03}\right)^{-1}, \sum p_0 q_4 \left(D^{PP}_{04}\right)^{-1}, \ldots$$

and it can also easily be seen from the fact that in general the inverse drifts $D^{PP}_{03}, D^{PP}_{04}, \ldots$ or the equivalent factors in tab. 5.3.1 will not be identical for all sub–aggregates, that:

No deflation method other than traditional *direct* Paasche deflation, especially no method using *chain* index deflators is capable of yielding structurally consistent volumes.

This is so, simply because the only index function able to fulfill this criterion is the *direct* Paasche price index (see sec. 2.3d). It is interesting to see that also the *chain* version of Paasche $\bar{P}^{PC}_{0t}$, let alone the chain Fisher (CF) deflation with $\bar{P}^{FC}_{0t}$, is *not* structurally consistent[241].

Table 5.3.1 shows that a change of CF–volumes may well be a result of a *variety* of influences. Therefore it is difficult to state what exactly has changed and to what extent when a volume of this type indicates a change of say x%. It is also clear from this table that there is a cumulative pattern in all chain index volumes, such that we get more and more complicated terms

$$\bar{Q}^{FC}_{03} = \sqrt{\frac{\sum p_3 q_3}{\sum p_1 q_0}} \sqrt{\frac{\sum p_0 q_1 \sum p_1 q_2 \sum p_2 q_3}{\sum p_1 q_0 \sum p_2 q_1 \sum p_3 q_2}}, \tag{5.3.3}$$

$$\bar{Q}^{FC}_{04} = \sqrt{\frac{\sum p_4 q_4}{\sum p_0 q_0}} \sqrt{\frac{\sum p_0 q_1 \sum p_1 q_2 \sum p_2 q_3 \sum p_3 q_4}{\sum p_1 q_0 \sum p_2 q_1 \sum p_3 q_2 \sum p_4 q_3}} \tag{5.3.3a}$$

and so on, and in general

$$\bar{Q}^{FC}_{0t} = \prod_{k=1}^{k=t} \left(\frac{\sum p_{k-1} q_k}{\sum p_k q_{k-1}} V_{0t} \right)^{1/2}. \tag{5.3.4}$$

Also note that it is sometimes (e.g. in using production indices to update volumes), necessary to make sure that deflation is consistent with an estimate of the physical quantity. But it is not easy to discern a “quantity” in an expression like $\bar{Q}^{FC}_{04}$ or in the “volume” of eq. **??** instead of $\sum p_0 q_4$. Nor is it easy to understand why such an

[241] Hence it is not ensured that the components of *real* values (deflated by use of *chain* indices) will add up to the total of that aggregate, *whatever* type of chain index is used.

expression is said to be a measure of an aggregate "at constant prices of $t = 0$" since $\bar{Q}_{04}^{FC}$ is affected by prices of all other periods (1, ..., 4) as well.[242]

Table 5.3.1: Comparison of different deflation methods (deflators P_{0t})

DP = direct Paasche; DF = direct Fischer; CP = chain Paasche

P_{0t}	$t = 3\ (\sum p_3 q_3 / P_{03})$	$t = 4\ (\sum p_4 q_4 / P_{04})$
DP	$\sum p_0 q_3$	$\sum p_0 q_4$
DF	$\sum p_0 q_3 \left(\frac{\sum p_3 q_3}{\sum p_3 q_0} \frac{\sum p_0 q_0}{\sum p_0 q_3} \right)^{1/2} = \sum p_0 q_3 \sqrt{R_{03}} = \sum p_0 q_3 \frac{P_{03}^P}{P_{03}^F}$ *	$\sum p_0 q_4 \left(\frac{\sum p_4 q_4}{\sum p_4 q_0} \frac{\sum p_0 q_0}{\sum p_0 q_4} \right)^{1/2} = \sum p_0 q_4 \sqrt{R_{04}} = \sum p_0 q_4 \frac{P_{04}^P}{P_{04}^F}$
CP	$\sum p_2 q_3 \frac{\sum p_0 q_1}{\sum p_1 q_1} \frac{\sum p_1 q_2}{\sum p_2 q_2} = \frac{\sum p_2 q_3}{P_{02}^{PC}} = \sum p_0 q_3 \frac{P_{03}^P}{P_{03}^{PC}}$	$\sum p_3 q_4 \frac{\sum p_0 q_1}{\sum p_1 q_1} \frac{\sum p_1 q_2}{\sum p_2 q_2} \frac{\sum p_2 q_3}{\sum p_3 q_3} = \frac{\sum p_3 q_4}{P_{03}^{PC}} = \sum p_0 q_4 \frac{P_{04}^P}{P_{04}^{PC}}$

* $R_{03} = P_{03}^P / P_{03}^L$ (see eq. 5.2.3 and 5.2.4, R_{04} defined analogously)

In the case of CF (chain Fisher) we get:

$t = 3$	$\left(\sum p_3 q_3 \sum p_2 q_3 \frac{\sum p_0 q_0}{\sum p_1 q_0} \frac{\sum p_0 q_1}{\sum p_2 q_1} \frac{\sum p_1 q_2}{\sum p_3 q_2} \right)^{1/2} = \sum p_0 q_3 \left(P_{03}^P / \bar{P}_{03}^{FC} \right)$ (5.3.5a)
$t = 4$	$\left(\sum p_3 q_3 \sum p_2 q_3 \frac{\sum p_0 q_0}{\sum p_1 q_0} \frac{\sum p_0 q_1}{\sum p_2 q_1} \frac{\sum p_1 q_2}{\sum p_3 q_2} \right)^{1/2} = \sum p_0 q_3 \left(P_{03}^P / \bar{P}_{03}^{FC} \right)$ (5.3.5b)

[242] The SNA rightly said in the light of this result: "Only in the special case in which time series of fixed base Laspeyres volume indices are used ... it is legitimate to equate ... real GDP with ... GDP 'at constant prices'. When chain indices are used, it is not appropriate to describe real GDP as GDP at constant prices"(para 16.71).

Another interesting observation is (see sec. 3.3c) that volume changes in the case of DP–deflation from period 2 to period 3 by a factor

$$\frac{Q^L_{03}}{Q^L_{02}} = \frac{\sum q_3 p_0}{\sum q_2 p_0} = Q^L_{23(0)},$$

a Laspeyres–quantity index rebased on base 2, depending on *base* period prices only. As all those growth factors of volumes Q^L_{03}/Q^L_{02}, Q^L_{04}/Q^L_{03}, ... depend on prices p_0 only, they are fully comparable.
The corresponding factor in the case of chain Fisher index deflation (CF deflation) on the other hand, is $\sqrt{Q^L_{23} Q^P_{23}} = Q^{FC}_3$ an index depending on the two price structures of period 2 and 3 (p_2 and p_3, or in general the prices of $t-1$ and t), both *other* than p_0. This is of course entirely in line with the chain index logic according to which (price–) weights of volumes have got to vary in the course of time and they should *not* depend on prices of the past. But such growth factors Q^{FC}_t are *not* comparable over time.

5.4 The SNA on how to deal with structural inconsistency

The SNA though most enthusiastic about $\bar{P}^{FC}_{0t}$–deflation was well aware of the structural inconsistency of volumes (in SNA called "non–additivity" [243]) resulting from this kind of deflation. Non-additivity was apparently supposed to be the most serious problem in this context and it is thus interesting to see what kind of remedy the SNA recommended. The SNA as well as the European System of National Accounts (ESA)[244]

1. noticed that non–additivity "can produce counter intuitive and unacceptable results in the long run" (SNA, para 16.65), but the SNA also made an attempt to play down this problem and nonetheless arrived at the conclusion that a Fisher chain index still "may provide the best volume measure of value added from a theoretical point of view" (para 16.67), a point of view described in some detail already in sec. 3.3c, and

2. the SNA discussed several ways out of (or at least concealing) the non–additivity problem.

Firstly:
The SNA correctly realized the restrictive nature of additivity (in the sense of structural consistency of volumes) because this "virtually defines" the pair of indices Q^L_{0t} and P^P_{0t}. But for the SNA being restrictive[245] was obviously reason enough *not* to follow this way. Interestingly instead of calling the "theoretical" point of view in question, a theory liable to counter–intuitive results, the SNA emphasized the idea that

[243] This is of course not the only problem with lack of additivity (another one is called A1 in fig. 2.3.2).
[244] See sec. 3.3c above.
[245] "Although desirable from an accounting viewpoint, additivity is actually a very restrictive property" (para 16.55).

the traditional Q^L_{0t}, P^P_{0t}–approach has to put up with lack of additivity as well once the base year is changed. It is only when a series is considered in which the base is kept constant when such problems do *not* emerge. Assuming tacitly that a series of volumes is considered, in which the base is renewed repeatedly the SNA turned the table and spoke of *advantages* of chaining (in *this* case). In oblivion of non–additivity (disadvantage) the SNA even discovered an "advantage ... that chaining avoids introducing apparent changes in growth or inflation as a result of changing the base year" (p. 16.74). The "advantage" consists in avoiding abrupt changes that inevitably destroy structural consistency at the expense of *always* having non–additivity, not only when direct indices are linked together, but also in *every* period. The logic seems to be: a defect is less damaging when it occurs more frequently.

Secondly:
The SNA discussed several ideas how to deal with non–additivity, but rejected most of them, like for example the following remedies (para 16.58)

- publish non–additive volumes "as they stand without adjustment"[246];
- to "distribute the discrepancies over the components at each level of aggregation";
- to "eliminate the discrepancies by building up the values of the aggregates as the sum of the values of the components at each level of aggregation" a possibility the SNA itself rejected, because this would not only contradict the results obtained by deflating the aggregate as a whole, but would also make the results "depend quite arbitrarily on the level of disaggregation".

Finally:
The SNA ended up with a recommendation giving a really poor account for the alleged advantages and theoretic superiority of chain indices:

> "Although the preferred measure of real growth and inflation for GDP is a chain Fisher index, ... it must be recognized that the lack of additive consistency can be a serious disadvantage for many types of analysis .... It is therefore recommended that disaggregated constant price data should be compiled and published in addition to the chain indices for the main aggregates. The need to publish two sets of data ... should be readily appreciated by analysts engaged in macro–econometric modeling and forecasting" (para 16.75).[247]

[246] The SNA also discussed the possibility to "publish data only in the form of index numbers" and to avoid absolute figures, but SNA rejected this idea "because it may merely conceal the problem from users" (para 16.56).

[247] We find exactly the same statements in ESA and in the Council Regulation (No. 2223/96) quoted in sec. 3.3c (there para 10.66). Hence these are rules of deflation binding for all Member States. ESA as well as SNA states that the "preferred measure" is a chained Fisher index of volume (direct deflation)

Another recommendation in view of "the well–known 'non–additivity' problem" (in the words of a European Decision quoted in sec. 3.3c) reads as follows:

> "...it will have to be explained to users why there is no additivity in the tables. The non additive 'constant price' data is published without any adjustment. This method is transparent and indicates to users the extent of the problem" (para 10.67)[248].

The solution does not appear to be the ultimate wisdom:

> Results are to be published showing discrepancies. These discrepancies have to be explained and there are also additional series to publish without such discrepancies. Furthermore the need to "explain" should be viewed against the background that the result of a chain Fisher deflation as shown above is in itself not easy to understand, as there are a number of influences in addition to prices and quantities in period 0 and t that may affect the result.

What makes the new method of deflation better than the old one when for a certain group of users results gained by the old method should be published in addition? What makes users other than econometricians and forecasters less interested in *consistent* data? The SNA has some users in mind "whose interest are confined to a few global measures of real growth and inflation". Why should they be more happy with non–consistent data resulting from a new method than they used to be with the traditional deflation data? Do prices which are more recent (but also varying with the passage of time) really increase satisfaction and make us forget about problems with discrepancies in aggregates and "explanations"[249] thereof?
It should also be noted that the great majority of Member States of the European Union were well aware of having

- to face the problem of non–additivity in volumes, and
- increased costs (expressly to the regret of some Member States) due to the need of an annual update of weights,[250]

or prices (indirect deflation) (para 10.62). *Chain* indices of Laspeyres and Paasche are only qualified as "acceptable alternatives", and there is no mention at all given to the direct version of those indices, that is to P^L and P^P (para 10.65). Note that "additive consistency" in the quotation above refers to what we would call structural consistency.

[248] In the same spirit we find in the decision just quoted (in the preceding footnote) as an example the following remarks (p. 84): "Clearly, the sum of A and B is no longer equal to the total. This is the famous 'non–additivity problem' .... The discrepancies between A, B and their total should ... be ... explained to the users. These discrepancies cannot be interpreted as indication of the reliability of the results".

[249] The "non additivity problem" which needs explanations and supplementary tables is not the only problem which may give rise to discussions. Another problem could be for example the *double deflation* method, like chaining a method prone to possibly producing counter–intuitive results such as negative value added at constant prices. The effect of a combination of double deflation with chain index methods (applied to both, output as well as input series) seems to be a problem not yet thoroughly investigated (see chapter 8).

[250] Moreover a decision was made without much knowledge about the impact on statistical results, let alone problems of their interpretation.

to name only some disadvantages the States knowingly and admittedly have to put up with adopting the chain index approach.
In the meantime some experiences (for example made in Australia where the response to methods is studied by MC LENNAN (1998)) show, however, that users seem to have less problems with “the well–known ’non–additivity’ problem” and the interpretation of $\bar{Q}_{0t}^{FC}$ than expected (for example in the texts of SNA and ESA just quoted). Nonetheless it is at least a conceptual problem.

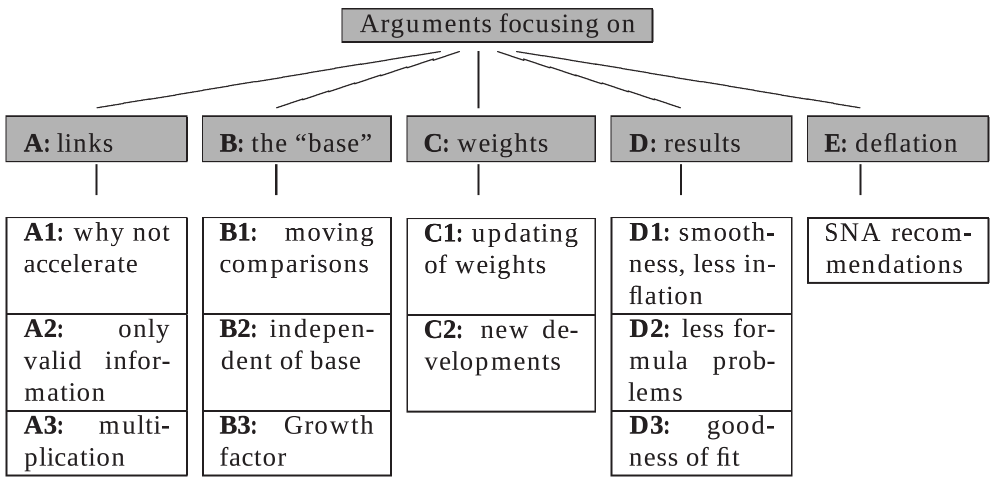
6 Arguments in favor of the chain index approach

In this chapter an attempt is made to review arguments put forward in favor of chain indices. As already mentioned at the outset the present author became more and more a non-chainer after having examined the arguments advanced by chainers. The presentation of the reasons given for the pretended superiority of chain indices over the direct approach in sec. 6.1 is followed by a critical assessment of the logical status and the validity of these ideas in sec. 6.2. It turns out – in our view at least – that the praise of chain indices is largely built on vaguenesses and inconsistencies. The point we want to make[251] is that none of the arguments (amounting to twelve altogether) is really compelling and convincing. It is also pertinent to comment on some more fundamental methodological issues concerning the rhetoric of chainers as done in the final sec. 6.3 of this chapter.

6.1 A parade of arguments for chain indices

It is difficult to find a system for the arguments of chaining. Nonetheless a try is made in what follows. In principle there are five types of arguments to support the idea of chaining (see fig. 6.1.1, and also tab. 6.2.1). The system already gives some hints to the critical notions referred to in the arguments in question.

Figure 6.1.1: Twelve arguments in favor of chain indices, an overview



A careful examination shows that most, if not all of the arguments known to us are simply built on some confusion with notions like "base", "chainability" etc., and on

251 See sec. 6.2 for more details.

some mysticism concerning the multiplication of links. The fifth group of arguments (not subdivided) is contributed by the SNA recommendation, and it adds the aspect of deflation to the alleged superiorities of chain indices (ironically it is precisely in the case of deflation where disadvantages of chaining abound as shown in chapter 5).

a) Focus on links rather than on the chain

The common trait of the following arguments (and some related arguments as well, as for example **B3**) consists in disregarding the existence of *two* elements in defining a chain index, the link and the chain (see sec. 1.1). Many statements pretending advantages of chain indices over direct indices in fact focus on the links rather than the chain, and they compare P_t^C instead of $\bar{P}_{0t}^C$ with the direct index P_{0t}, or they consist in giving too wide an interpretation of the fact that $\bar{P}_{0t}^C$ is derived by multiplying links, if not in a mystification of this multiplication.

A1: The "why not" or "limiting case" argument:

One of the most widespread and intuitively appealing arguments goes as follows: The chain approach is simply a "fixed base" index rebased more frequently. Instead of rebasing at five year intervals or so, updating is simply done at shorter intervals, viz. every year. Thus the chain index is the "limiting value" or "borderline case" to which a direct index tends. Or as ALLEN 1975, p. 177 has put it: "why not accelerate and go for annual chaining? There is no reason why not." We might call this most insinuating and enticing argument simply for short *"why not"*–argument.[252] On the basis of this argument we sometimes also find the statement that a chain index is (compared with the direct index) simply the more general approach[253]. In this sense the SNA (para 16.77) states for example:

> "In effect, the underlying issue is not whether to chain or not but how often to rebase. Sooner or later the base year for fixed weight Laspeyres ... indices ... has to be updated because the prices of the base year become increasingly irrelevant .... Long runs of data therefore almost inevitably involve some form of chain indices. Annual chaining is simply the limiting case in which rebasing is carried out each year instead of every five or ten years."

So why not accelerate?[254]

[252] It is the intention of this book to show that there are indeed *several* good reasons "why not", and to work out in detail what these reasons are.

[253] It is tacitly assumed that chaining of links were the general rule in index compilation, while keeping a basket constant were only the exception. See also sec. 1.2 and the following footnote.

[254] Argument **A1**, esp. as expressed in the SNA quotation above, is a misinterpretation of the guiding principle of the fixed–basket–approach: what is primarily wanted in this approach is comparability *within* an interval, not so much the long term comparison across a great number of such intervals.

A2: The "only valid information" argument (accuracy issue):
According to this argument (put forward by Mudgett in particular) the chain index gives the only valid obtainable information, the direction of *change* from year to year, whereas the fixed based index concentrates on the more doubtful information, like the actual *level* attained in terms of the base period level (doubtful especially when the base period is years ago).[255] If this "only valid information" – argument were correct we should rather refrain from multiplying links to a chain (an operation according to **A1** characteristic for index calculations of all sorts)[256]. Given that links are the only meaningful measures whereas P_{0t} is not valid simply because of 0 being too distant from t, why then should $\bar{P}_{0t}^{LC}$ be able to provide a valid information? In some sense or other with this argument the same problem arises as with the multiplication mystery and Martini's theory of indirect comparison (see sec. 4.1) on the strength of which we are allegedly able to compare validly indirectly (by chaining) only what is incomparable directly. There is also some resemblance with the following arguments, which again suggests that the mere subdivision of the interval is able to do wonders.

A3: The multiplication mystery
In some arguments the trivial operation of multiplying links is regarded as an advantage as such. BANERJEE (1975), p. 48 for example wrote: "The chain is thus the resultant of a series of comparisons, each between two consecutive periods. The chain method thus eliminates the limitations involved in the comparison between two distant periods".
It is immediately apparent that much of the persuasive power of this argument is owed to the vagueness of the term "resultant". On a level with this idea is also the (unproved) assumption that chaining results in a smaller total *error* in computations using data each liable to errors, or in a smaller deviation in terms of the factor reversal test or time reversal test respectively.
The last mentioned criterion is sometimes referred to as "formula error". To quote BANERJEE (1975), p. 59 once more: "if the comparison is made between two consecutive periods, the formula error should be the least. This incidentially, is a justification for the use of the chain index".[257] Interestingly this argument is again (like **A2**) very much stressed by Bruce Mudgett one of the most convinced proponents of chaining. In his view "the rule that comparisons of price and quantity change should be made between consecutive periods or areas" (MUDGETT (1951), p.57)[258] also helps to im-

[255] We find a similar idea in a somewhat arcane argument introduced by U. P. Reich (see sec. 7.3a) according to which it is only the change (or "dynamics", "growth rate" etc.) and not the level that counts in economics.

[256] Hence in a sense from **A2** ensues that long term comparisons should better *not* be made while **A1** states that just this is the essence of all sorts of index calculations.

[257] Note that Banerjee, like MUDGETT (1951) defines the extent to which an index deviates from the time reversal, factor reversal, or circular test as various types of "formula error". It is clear that in general the "error" will be smaller the shorter the interval over which a comparison is made. See also the discussion of argument **B1** below.

[258] Mudgett not only expressly treats intertemporal and interspatial comparisons equally (in contrast to "consecutive periods" it is not clear what is meant by "consecutive areas"), it is also the emphasis he

prove what he calls "formula consistency", i.e. a small difference between P^L_{0t} and P^P_{0t} (see also **D2**), and "homogeneity", i.e. more common commodities when two *adjacent* periods are compared (MUDGETT (1951), p. 43, 48, see also under **C2**). Thus according to Mudgett there are two sources for "advantages" of chain indices, both in terms of "accuracy" and "validity": chain indices are superior as they

- make comparisons exclusively between two *consecutive* periods (concentrating on links), that is by focusing on *short* period comparisons, or on the "only valid information", and *at the same time* as they
- provide "a valid procedure" to make comparisons over *long* series or distant areas by multiplication of those links (MUDGETT (1951), p. 58).[259]

To sum up: it appears inconsistent to take the link for the whole chain, and to ignore the aspect of multiplying links, as done in argument **A2** (and in a way also in **A1**), and at the same time (as argued especially under **A3**) to derive anew some "advantages" of the chain approach from the simple fact that links are multiplied to form a chain. Therefore

> We call argument **A3**, the "multiplication *mystery*" argument, because it consists in deducing some far–reaching consequences from the fact that the interval (0,t) is subdivided, links are considered for the subintervals, and they are multiplied. The alleged advantages consist in: more accuracy (**A3**), and – closely related to this argument – in making long distance comparisons possible which are otherwise (i.e. directly) impossible, in utilizing the information represented in the time series data in a more efficient way, and thereby also in providing valuable additional information (argument **B1**).

The idea of **A3** (and implicitly in related arguments, like **B1**, **C1** and **C2** as well) is often found highly convincing and also the key to Martini's thinking: chain indices permit long distance comparisons by comparing any two consecutive periods only, and aggregating over such short distance comparisons. Thus in some mysterious manner, what is incomparable directly becomes comparable indirectly.

b) Ambiguities concerning the notion of "base"

While the common feature of arguments of group **A** is to emphasize the existence of "links"and their multiplication, a second group of arguments in favor of chain indices (group **B**) rests on some confusion related to the concept of "base".
The impression that there is no need to refer to a base in the chain type of comparisons, or that there is some automatic choice of the base period, is again a result of taking the link for the chain, and of some vagueness and embroilment created by notions like "base", "chainability" etc. The fundamental idea of this group of arguments is

---

placed on Fisher's "reversal" ideas with which we most decidedly disagree.

[259] More than that, Mudgett even claims that it is sometimes the only possible procedure (as also Martini argues), or at least a more accurate procedure. See sec. 3.7c for our position in this question.

> that chain indices provide a different type of comparison by making use of a "moving" base, and that thus they have "solved", or rather made irrelevant a problem of sizeable difficulty to all sorts of direct indices, that is the choice of an appropriate base period.

In our view this is bluntly a misunderstanding of the term "base". chain indices in fact represent another way of analysing time series data on prices and quantities, by no means a more elegant way, however. The interesting feature, however, is not the "moving" base of the links but rather the path dependency of the resulting chain.[260]

B1: Moving comparison, additional information

chain indices are said to reflect whole time series or "*runs*" of index numbers instead of simple (direct) binary comparisons of two points only, and to provide some kind of "*moving*" or "rolling" comparison (ALLEN (1975), p. 177), apparently independent of some arbitrary point in time, called "base". It is also pretended that users of index calculations are even freed from choosing a base year which should be a "normal year" (see also argument **B2**).

Here it is noteworthy already, without entering the discussion of this argument, that there is some relationship between **B1** on the one hand, and **A1** and **A2** on the other, in that the idea of a "moving" or "rolling" base again applies to the links, and not to the chain. It is only the "base" of the link, $t-1$ in P_t^{LC}, which is "moving", but of course *not* the base 0 in $\bar{P}_{0t}^{LC}$. It is also difficult to see, why the sequence $\bar{P}_{01}^{LC}, \bar{P}_{02}^{LC}, \ldots$ is called a "run", while $P_{01}^{L}, P_{02}^{L}, \ldots$ is not.

By argument **B1** is possibly also meant, that an index should account for not only periods 0 and t, but somehow (how exactly?)[261] also for all intermediate periods 1, $\ldots, t-1$. But note that this does not necessarily mean that *links* have to be *multiplied* to form a chain. In sec. 4.4c we introduced some, curious though, index formulas considered in literature[262] which also may allow the taking into account of the whole time series of prices and quantities (weights).

Argument **B1** also comprises the somewhat arcane idea that chain indices provide *valuable additional information* due to making better use of all time series data. Again a short remark may suffice without anticipating the discussion in sec. 6.2: to our knowledge none of the authors making use of this quite popular and prevalent argument gave any indication of the type of information he or she has in mind, let alone wherein its value consists. By second thoughts it turns out that the "additional information" is a mere figment which indeed fits well to the mystery of **A3**.

[260] Or the specific way in which the value of the chain in t depends on the value in $t-1, t-2, \ldots$ due to cumulative effects and the nonlinearity (see sec. 3.5)
[261] See sec. 3.5 for the lack of a theoretical justification for the nonlinear pattern of aggregation over time.
[262] Like the blend–system or the broadened base system.

B2: Independence of a base

It is argued that the chain index is *independent of the base*[263], in the sense of being shiftable without error because (as shown in eq. 4.1.6)

$$P_{34} = \frac{P_{04}}{P_{03}} = \frac{P_{14}}{P_{13}} = \frac{P_{24}}{P_{23}} \tag{6.1.1}$$

holds for chains $\bar{P}_{0t}$ by definition, since $\bar{P}_{0t}$ is *derived from* multiplication. The interpretation given to eq. 6.1.1 is usually this: A chain index is no more tied to one base than to another, and the result P_{34} in eq. 6.1.1 is the same, irrespective of whether 0, 1, or 2 is the base. This alleged "solution" or "irrelevance" of the problem of choosing a base period[264] is often viewed as a great advantage of the chain principle.

Note that this is simply implied by the *definition* of a "chain" (or the operation of "*chaining*" as opposed to the property of "*chainability*", see also sec. 4.1f for this distinction).[265] Interestingly this is sometimes misunderstood as some kind of non–existence of a base. CRAIG (1969) for example claimed, that the chain index' violation of identity (as demonstrated in sec. 3.2a) is no problem on the following grounds: to get $\bar{P}^{LC}_{0t} = 1$ in case of identical prices in 0 and t is not necessarily desirable, because 0 can well be badly chosen as base period, and it is better *not* to return to such a base. This "argument" for a chain index completely disregards the fact that $\bar{P}^{LC}_{0t}$ has of course a base, which is 0, and obviously the same as the base of P^{L}_{0t}. From such confusions with the notion of "base" ensue some other inconsistencies:

> In chain indices the *reference* base (RB) is deemed irrelevant, and chain indices are praised not least for being independent of the RB (argument **B2**). On the other hand increased attention is given to the weight base (WB), the up–to–dateness of which is said to be of the utmost importance (as claimed in argument **B3**, but also in **C1** where the chain index is praised above all for making use of the most recent, and hence most "relevant" WB).

Such inconclusive statements give rise to the question[266] whether now, in the case of chain indices the term "base" still has the same meaning as it has in the direct index framework.

B3: A more relevant growth factor

A chain index (in fact, however, the link) is said to represent directly the growth factor[267] (compared with the previous year) of an index. This argument is applied

[263] We also found in literature the most revealing misunderstanding that a chain index has *no* base (!), or that its base is automatically always $t-1$.

[264] We have already discussed this argument along with the following one (**B3**) in sec.3.3c.

[265] Moreover there is no reason given in **B2**, why P_{04}/P_{03} should equal P_{14}/P_{13}, and P_{24}/P_{23}. Taken in isolation the relationship given by eq. 6.1.1 is not desirable, though in a sense it guarantees irrelevance of the reference base.

[266] To be discussed in more detail in sec. 6.2.

[267] This is of course trivial because according to eq. 6.1.1 the chain is *defined* by multiplication of such growth factors (links).

- to price indices (that is to an increase of a price level), depending on quantities, and (more often) to
- growth rates of volume (real value), that is to quantity indices depending on prices (see sec. 3.3c).

What (allegedly) makes

$$\frac{\bar{P}^{LC}_{0t}}{P^{LC}_{0,t-1}} = P^{LC}_{t} = \frac{\sum p_t q_{t-1}}{\sum p_{t-1} q_{t-1}} \tag{6.1.2}$$

a "better" growth factor than

$$\frac{P^{L}_{0t}}{P^{L}_{0,t-1}} = \frac{\sum p_t q_0}{\sum p_{t-1} q_0} \tag{6.1.3}$$

is of course the weight q_{t-1} in contrast to q_0. The same applies to weight p_{t-1} in contrast to p_0 in case of quantity indices

$$Q^{LC}_{t} = \frac{\sum q_t p_{t-1}}{\sum q_{t-1} p_{t-1}} \quad \text{as opposed to} \quad \frac{Q^{L}_{0t}}{Q^{L}_{0,t-1}} = \frac{\sum q_t p_0}{\sum q_{t-1} p_0}.$$

c) Adjustment of weights, new products and changes in quality

The common feature of this group of arguments is the contention that chain indices are able to solve almost unsurmountable problems involved in fixed basket indices and at the same time these arguments give chain indices an image of flexibility, adaptability and thus also of a much greater suitability to modern needs[268]. It is claimed that

> chaining is an elegant device to elude the trouble with keeping a basket constant, despite changes in types and qualities of goods, in the relevance of certain outlets and the like. Chaining is not only of great promise as it avoids such problems, but also as it provides "indices whose weighting structures are as up–to–date and relevant as possible" (SNA).

Interestingly the decision for a new method of deflation in Europe in 1998 (see sec. 3.4c) was primarily made on the basis of exactly this twofold argument of type **C**:

- a more modern and flexible index design with weights constantly updated (**C1**) *and* at the same time

[268] Aspects like these are often regarded as most convincing in support of the idea of chaining. At first sight one is sorely tempted to advocate for chain indices precisely on those grounds: chain indices seem to match far better with a rapidly changing world like ours, and *at the same time* they also avoid problems inherent in a fixed basket approach which is (in contrast to the chain approach) unfit for keeping pace with the times.

- an elegant solution of a burdensome problem we always had to cope with in case of the old design of a fixed basket (**C2**).

C1: Most up–to–date weights

Especially the rev. SNA brought this argument into play, praising emphatically the chain indices as (see above) "indices whose weighting structures are as up–to–date and relevant as possible". Note that this argument compares a direct index with a chain index, as if they both had the same kind of weighting, as for example a single weighting scheme only. To say we always make use of the most up–to–date previous year weights is correct only for P_t^{LC}, not for $\bar{P}_{0t}^{LC}$, where we have *a number* of the weights which in addition also are cumulative that is they produce in conjunction with one another the final result $\bar{P}_{0t}^{LC}$.

C2: Less problems with new developments, quality adjustment etc.

The SNA also found it worth mentioning that chain indices have *practical* advantages. They make it "possible to obtain a much better match between products in consecutive time periods ..., given that products are continually disappearing from markets to be replaced by new products, or new qualities". It is often said, that the chain principle facilitates (or accelerates, and thus improves) the adoption to new developments by making use of the most recent weights, and to handle the withdrawal of old and the entry of new commodities better and more easily. It is well known that the fixed basket approach of the Laspeyres direct index inevitably runs into difficulties as new products emerge, old ones are no longer available, and the characteristics of goods change substantially with an impact on consumers' utility level and on the classification of commodities used to select the specific basket. Hence every index design that appears able to prevent such type of problems from existing will be welcomed.

On the face of it the chain principle thus not only allows for a constant adjustment of quantity weights, and for the comparability of non–overlapping selections of items (commodities) in 0 and t by comparing 0 with 1, 1 with 2, ..., $t-1$ with t. It also pretends to "solve" one of the most dodgy problems in practical price statistics, that is to make adjustments for changes in the selection of goods (and specification of their quality), outlets, etc. It is said that "A chain index naturally incorporates changes in the qualities and types of goods as it follows changed consumption patterns through the years" (FORSYTH (1978), p. 349), or it "automatically incorporates" changes in consumption (FORSYTH and FOWLER (1981), p. 229).

A second thought reveals, however, that the problem with such statements is that it is not clear what is meant by "incorporates" in this context[269]. It is also left open in which sense the practical problems in question (i.e. quality adjustments and the like) have been "solved".[270]

[269] Vagueness is not only typical for this argument to advocate chain indices, we also find it in considerations like chain indices which use more "relevant" weights (how to measure relevance?), they utilize time series data "better" (in which sense?), or they provide some additional information (which one?).

[270] It will be shown in sec. 6.2 that argument **C2** is particularly unfair, because difficulties and problems

d) Arguments referring to expected and desired results of index calculations

The following group of arguments focus on the numerical results to be expected when chain indices are used. The arguments do not refer to conceptual aspects whatsoever. Nor do they apply to *all* sorts of empirical data, like axiomatic considerations do for example. Thus arguments of group **D** are not independent of the kind of data entering an index formula, and they are in a way ambiguous, in that they imply advantages as well as disadvantages.

D1: Smoother development, smaller inflation rates
It is sometimes maintained that chain indices are likely to yield these type of results, unless some cyclical movement or regular fluctuation occurs. Chain indices are found unfavorable in case of fluctuations[271], but useful (in terms of desirable numerical results, like low inflation), especially "if individual prices and quantities tend to increase or decrease monotonically over time" (SNA 93, para. 16.44).

D2: Less formula problems, small LPG[272]
It is also often conjectured that the chain index version of various index formulas will yield less divergent results than the corresponding direct index version. For example Laspeyres and Paasche chain indices, $\bar{P}^{LC}_{0t}$ and $\bar{P}^{PC}_{0t}$ will have less different results than P^{L}_{0t} and P^{P}_{0t} if applied to the same data. Under certain conditions the spread of index formulas (or the "gap" between them) will be smaller for chain indices as opposed to direct indices[273] , with the result that "the choice of index number formula assumes less significance" (SNA 93, no. 16.51). This argument may be called "less formula problems" or somewhat ironically "reduction of conceptual expense". As shown in sec. 3.3c the argument **D2** apparently was also decisive for the Australian Bureau of Statistics (ABS, see Mc. Lennan (1998)).

D3: Goodness of fit in an econometric model
According to Selvanathan and Prasada Rao (1994), p.125, satisfactorily high values of the coefficient of determination r^2 found in econometric models, "reinforce the general claim of superiority of the chain base index numbers over fixed base indices, especially when the current period is further away from the base period". Note that a higher goodness of fit in a model is taken as a proof of *conceptual* superiority, and there is no consideration of concepts and axioms. As a justification of methods this is clearly insufficient.[274]

from which the direct Laspeyres index suffers will by no means be "solved" but rather "dissolved" or made irrelevant.

271 As is well known this is why SNA does not recommend linking more frequently than annually. In case of cycles chain indices are *not* useful. On the other hand it is implicitly assumed when no cycles occur the converse is true, such that chain indices do have advantages.

272 Laspeyres–Paasche gap (see sec. 3.3b).

273 They may even converge on a single chain index, such that the gap (difference) between them vanishes.

274 Nobody would claim A being a better "domestic product" than a differently defined aggregate B, only because the correlation of A with, for example money stock, M3 is higher than correlation between B and M3.

e) Chain indices recommended for deflation (SNA)

The SNA recommendation adds another argument to the alleged advantages of chain indices: superiority of chain indices for purposes of deflation. The main reason for the preference of chain indices is again the desire to guarantee a continuous update of weights as soon as possible. The reasons given by the SNA 93 have been demonstrated and commented at length in chapter 5. To sum up we may say:

> In the opinion of the SNA for index formulas there isn't anything quite as important as the up–to–dateness of weights. To use the most recent quantity weights in order to measure inflation and the most recent prices in the framework of deflation is capable of outweighing all possible shortcomings involved in adopting the chain index principle.

The idea is to avoid terms like $\sum p_0 q_t$ in which prices and quantities refer to most different periods[275], thus making the products somewhat artificial, as opposed to $\sum p_0 q_0$ or to $\sum p_t q_t$, where prices and quantities refer to the *same* period. To make this idea (prices and quantities should not belong to periods differing too much) a criterion or even a guiding principle in index construction eventually amounts to discarding all attempts of calculating "volumes" (being *necessarily* different from "values" $\sum p_t q_t$ as in the case of volumes prices and quantities of *different* periods are combined), and hence of deflation, if not of indices altogether[276].

It can easily be seen, and will be demonstrated in the following section, that it would be fundamentally mistaken to ban all expressions combining prices and quantities belonging to different periods (like $\sum p_0 q_t$ or so), on some pseudo–theoretic grounds as this would deprive us of most of what index theory has accomplished.

6.2 Rebuff of the arguments for chain indices

In this section we will first try to find some general aspects applicable to all arguments brought into play in the chainers' praise for chain indices, and validate the rebuttal of *all* those arguments, and thereafter we will deal with specific considerations related to *some* selected arguments of chainers.

---

275 As pointed out in chapter 5 the SNA said: a process of production which is efficient at one set of prices may not be very efficient at another set of relative prices; an artificial aggregate like $\sum p_0 q_t$ may not only be improbable, it may even be (technically / economically) *impossible* (if prices p_0 really would prevail, the production process that created q_t would not be used). As shown above this kind of reasoning, exemplified in the SNA, would eventually result in an alleged "impossibility" of deflation. All sorts of "volumes" are with necessity some *artificial* combinations of prices and quantities other than the *observed* combination $\sum p_t q_t$.

276 We come back to this argument in sec. 8.2c .

a) General remarks valid for all arguments of chainers

We start with a summary table in which the arguments are listed and succinctly described once more. Moreover some hints are given to the parts of the book where we already dealt with the argument in question (see tab. 6.2.1). In reviewing the arguments it is eye–catching that

1. Justification of chain indices is not theory–driven. There is not only no "new" conceptual foundation designed esp. for chain indices, but we also encounter inconsistency, onesidedness, and decidedly result–oriented measurement (esp. in type **D** arguments).
2. "Advantages" of chain indices are mainly derived from a criticism of the fixed basket (direct Laspeyres) approach. However in some arguments the original issue with regard to the length of the time series seems to have been lost out of sight. The direct index is objectionable in the case of the long series, but not in the short series. So why recommend chaining in the case of short series? There is no need to cure a nonexistent disease, i.e. for chaining in situations, where a fixed basket approach is not defective.
3. It should be recognized that the pretension, that a fixed–basket–index problem is "solved" by moving to chain indices is impressive only as long as the problem still exists when the fixed–basket approach is abandoned and the chain approach is adopted.

The last type of considerations will be called "solution vs. dissolution" for short.

1. Lack of theory, inconsistencies, onesidedness, and result–oriented measurement
It is generally accepted that we have two types of conceptual foundations of index formulas:

- the idea of a (fixed) "basket", for which the expenditure has to be observed on a regular basis (since expenditure is constantly changing due to changing prices), or more distinct: the idea of a pure price comparison, and
- the so called "economic theory approach", which compares expenditures necessary for maintaining the same *utility* level (rather than buying the same basket–*quantities*).

Interestingly chain indices do not add a third theoretic underpinning to index theory[277]. They are based on the basket–concept, modifying it only by the simple idea that it is primarily the updating (instead of the constancy) of the basket that matters.

[277] It is sometimes maintained that the Divisia index is capable of providing such a theoretic foundation for chaining. We will examine this assertion in chapter 7. Note also that all of chapter 4 was expressly devoted to the search for a "theory" – if there is any – underlying the chain index approach.

We also find a number of *inconsistent* and *inconclusive* statements. The most striking, as already mentioned above[278], being the standpoint the SNA has taken with regard to unit value indices as opposed to chain indices: Unit value indices were rightly rejected by the SNA as being

> "affected by changes in the mix of items as well as by changes in their prices. Unit value indices cannot therefore be expected to provide good measures of average price change over time" (para 16.13).

Amazingly nobody realized that the same criticism would also apply to chain indices, advocated, however, with great vigor by just the same SNA 93.

In a broader sense we may also see another interesting inconsistency in accepting the idea of a "pure" price comparison: there is usually no problem whenever a single commodity is concerned, but for some reason or other things seem to change fundamentally when a *basket* of $n > 1$ commodities, also called a "composite commodity" is concerned.[279]

We will discuss this issue once more in chapter 8. Let us sum up as follows:

> Most people will readily reject the idea that a rise or decline in the price of a single ($n = 1$) commodity can be discovered by comparing two completely different commodities. But curiously the same is *not* true for comparing different composites of a *number* $n > 1$ of commodities. It is not only found legitimate, it is even urgently required to base such a comparison on a selection of goods with a *varying* (constantly updated) structure.

The most frequently overlooked inconsistency is of course this: It is rather $\bar{P}_{0t}^{LC}$, but not P_t^{LC} which has to be compared with P_{0t}^{L}. Not only the arguments of type **A** are based on this confusion, we also find this logic in some other arguments, like **B1**, **B3**, **C1** (arguments of group **C** also ignore the fact that $\bar{P}_{0t}$ has cumulative weight, unlike the link P_t^{LC}, which has a single weight only) and in a sense also in **D2**.

Finally a most fundamental inconsistency is addressed in the theorem of Funke et al. (see sec. 4.2a) according to which the ideas of

- chainability, allowing a consistent aggregation over time by chaining, and
- the constant adjustment (up–dating) of weights

[278] See also sec. 0.2c (footnote 10) for the SNA quotation.

[279] We again refer to this argument in fig. 8.2.2 below. A good example for this inconsistency is STUVEL (1989) who elaborates in great detail, at the very beginning of his book the importance that "we indeed are comparing like with like". He also says "The same is true for composite commodities, i.e. collections or 'baskets' of goods and services of unchanging composition" (p. 3), and "…it is necessary that the physical composition of the basket does not change" (p. 5). But he does not seem to be concerned with indices combining *different* baskets, nor is he concerned with chaining (for *this* reason at least).

are incompatible, in the sense that there is no index formula known with variable weights and simultaneously satisfying the circular test.[280]

Onesidedness and playing down of counter–arguments can be seen by the fact that as good as none of the known disadvantages is mentioned in the parade of arguments above (sec. 6.1):

- there is no mention for example of problems with consistency in aggregation (over commodities), that is the decomposition of the index formula in order to distinguish different sources of variation, or to additivity in volumes, and
- not much attention is given to problems of understandability of the chain approach[281], or to problems of data collection, let alone the increased c ost official statistics has to incur when a direct Laspeyres index is replaced by a chain index.

Interestingly and most regrettable: there is little or no "cost–benefit–analysis" made in order to show whether a move from a direct Laspeyres index to a chain index will really be worth the indisputedly higher cost of collecting information (esp. on weights)[282].

It is not even well understood to which extent empirical results of chain indices might differ from those of direct indices[283], such that arguments of type **D** will apply[284]. These arguments are not only "measurement without theory", but they rather represent what might be called *"measurement aiming at desired results"*. Arguments of this kind are

- unlike a priori considerations (axioms) not *logically* compelling, but rather related to specific data;
- whether they become true depends on empirical conditions (such as monotonical versus cyclical price movement) for the data, and finally

[280] Even so the combination of both ideas certainly has some intuitive appeal. This inconsistency also sheds some light on the multiplication mystery (what is directly incomparable is nonetheless comparable indirectly), which is essentially built on the same combination of two ideas (chaining, and variable baskets) which are clearly not reconcilable. Thus this most popular argument possibly cannot hold water either.

[281] By this it is not only meant the interpretation of the "message" of the chain index, for example in terms of a "pure" price comparison, there are also no attempts known in giving an (economic) interpretation to the *product* of links. We have such interpretations in the case of direct indices, for example in terms of an expenditure ratio or a mean of relatives. Interestingly none of these two interpretations of an index formula applies to chain indices. The formula now consists of a product but there is no effort made to give an interpretation of it.

[282] We are going to discuss some of these problems in chapter 8. Due to the need of updating the weights in much shorter intervals in time the chain approach is markedly more demanding and costly. But according to FORSYTH and FOWLER (1981), p. 230, this is the price for the "overwhelming advantages" of chain indices.

[283] What can be taken for sure is not much more than simply this: the price level will be more affected by changes in quantities consumed than it used to be in case of the direct Laspeyres price index.

[284] There will be no special rebuff of this type of arguments in part b of this section, thus we think we had discussed the group **D** arguments in some detail above already.

- to assess whether the conditions in question are met, some kind of index calculation has to be made. Hence a formula has to be chosen and applied to data in order to decide on a second step in the choice of the appropriate index formula.[285]

In conclusion arguments of type **D** are unsatisfactory for a number of reasons. There simply is not sufficient evidence that chain indices will produce smooth time series under realistic conditions. Experiments with chain indices rather reveal that "chain indices can be more biased than their direct counterparts, which is both contrary to popular belief and to the purpose of linking" (SZULC (1983), p. 555).
Arguments of type **D** not only refer to the numerical "behavior" (instead of the conceptual foundation) of chain indices which greatly depends on the type of data to which chain indices are applied, but they also require conditions[286], under which the defects, the direct Laspeyres index is blamed for would be less likely. This leads us to a second general remark.

2. Criticism of the fixed basket approach is lost from perspective, or: For which situations superiority of chain indices has to be proved?

In view of the undoubted increase in computation difficulties and cost there should be some *significant* improvements to justify a move from fixed base direct indices to chain indices. Furthermore these indices should be advantageous especially in those cases in which the direct index approach fails, or is a target of criticism, that is when

- comparisons over *long* intervals in time are wanted, and
- consumption patterns change *rapidly* and *fundamentally* in response to changes in prices.[287]

Thus to be worth the increased burden and expense of compilation chain indices should guarantee a *substantial* progress to direct indices in just these fields, where direct indices face their specific weaknesses and (allegedly) intolerable disadvantages.
But surprisingly, many proponents of chain indices seem to have forgotten what was originally under discussion. The most prominent representative of this logic is the "only–valid" argument **A2** stating that chaining is primarily good in the short run. By the same token Stuvel made objections[288] to chain indices for *long* term comparisons, but as to *short* time series they "might be preferable, since the direct comparisons in such a series are between adjacent years" (STUVEL (1989), p. 73).

[285] Moreover it is possible that a part of the prices combined in a price index changes monotonically and another part is influenced by some cyclical movement.

[286] Such as smooth and non–cyclical changes, a short rather than long time series, a current year's basket not too different from the base year's basket etc.

[287] In case of slow changes of minor importance we can hardly blame the direct index approach for making a comparison on the basis of an *irrelevant* historical basket.

[288] This is because chain indices cannot be interpreted in terms of Stuvel's "additive model" of decomposing the value change into a change of prices and of quantities respectively.

Table 6.2.1: Summary table of the twelve arguments for chain indices

Name of argument	Short description of the argument	Dealt with in section
	A: links	
A1: why not accelerate	Sooner or later the base (even of direct indices) has to be renewed, chaining is simply the limiting case[a].	4.1d, 4.2b
A2: only valid information	Only the direction and amount of *change* from year to year is valid, not the *level*.	3.7a, 7.3a
A3: multiplication mystery	What is directly incomparable is nonetheless indirectly comparable (Martini). Smaller error in computation (Banerjee).	3.7c, 4.1b
	B: the "base"	
B1: moving (rolling) comparison	Base is constantly changed (rolling)[b], better use is made of time series data and some valuable additional information is provided.	3.3c, 3.5, 4.1f, 4.2b
B2: independent of the base	No need to choose a base in chain type of comparisons, or base is automatically the preceding period (year).	3.3c, 4.1f+g, 4.2b
B3: growth factor	Better growth rates (as calculated on the basis of non–obsolete weights) of volumes or prices.	3.3c
	C: weights	
C1: updating of weights	Chaining results in "indices whose weighting structures are as up–to–date and relevant as possible" (SNA).	4.2, 4.4
C2: new developments	Less problems with price quotations, esp. with making quality adjustments.	3.2b, 3.7a
	D: results	
D1: smoothness, less inflation	Smoother development except for the case of cycles (regular fluctuations) in prices.	3.2f, 3.6, 6.2a
D2: less formula problems	Laspeyres Paasche gap (LPG) will be smaller in the case of chain indices.	3.2f, 3.3b, 3.6, 6.2a
D3: goodness of fit	High value of a coefficient of determination r^2 in an econometric model.	3.6, 6.2a
	E: deflation (SNA recommendations)	
E: SNA on deflation	Theoretically the preferred method[c].	3.3c, chap. 5, 6.1e

[a] To the extent the frequency of updating of weights is increased a direct index tends to a chain index.
[b] There is also the idea that chain indices will display less abrupt changes after (more frequent) base renewal.
[c] The desire to continuously update weights as soon as possible overshadows all other considerations.

This being so, the advantage of chain indices over indices like P^L_{0t} would be rather limited since short series are not what is criticized in case of P^L_{0t}. Chainers seem to have lost sight of their original issue, the critique of P^L_{0t}, according to which it is primarily the updating (instead of the constancy) of the basket that matters. In this light it is not fair

- to claim superiority of the chain approach over the direct approach when the interval between 0 and t is short, and weights have not changed dramatically[289], and
- to ignore the fact that chain indices are not the only indices in which weights are constantly updated (an alternative is the use of "superlative" direct index formulas, as mentioned in sec. 1.1, in which also the current year weights enter)[290]

As stated above (sec. 1.1) most of the chain index rhetoric is ensuing from a critique of P^L_{0t}, applicable only under certain conditions. It should be noted, however, that the critique does not apply to certain "superlative" direct indices.
Thus arguments which focus mainly on the renewal of weights, are incapable of establishing superiority of chain indices over *all* kinds of direct indices (they are suitable at best relative to P^L_{0t}), and cannot be proof of a general desirability of chaining.

3. Solution vs. dissolution of a problem
Some arguments for $\bar{P}^{LC}_{0t}$ owe their appeal directly to the comparison with P^L_{0t}, in that they explicitly pretend to solve problems we usually have to face in the direct index approach. With this group of arguments we have belong in particular

- the pretension that a chain index has solved the problem of choosing a base, and
- the alleged solution of problems with making quality adjustments,

that is arguments **B2** and **C2** respectively. They are indebted to a great deal to some vagueness of the term "solution". As will be shown in connection with a rebuff of arguments **B2** and **C2**, below doubts arise as to the nature of the promised "solution":

> In the chain index approach "base" has another meaning, such that there no longer is anything which makes the choice of the base period a problem. Furthermore once the fixed basket (constant over time) is abandoned, there is also no longer any need to ensure comparability over more than just two adjacent periods, $t-1$ and t, and hence there is no problem in making adjustments for changes in quality or other price determining characteristics. It turns out that the "solution" of problems (typical for the fixed basket approach), like **B2** and **C2** simply consists in dissolving the problem altogether.

[289] Note for example that also arguments, like **D1** and **D2** actually require conditions under which the specific defects for which the direct Laspeyres index is blamed are less likely.
[290] The idea of updating weights not necessarily requires a chain index. The Boskin Commission for example no less stressed the importance of a rapid and constant updating of weights, did not decidedly recommend chain indices, but rather (direct) "superlative" indices, using weights from both periods, 0 and t respectively, like for example a Fisher or Törnquist direct index.

We now turn to specific weaknesses of some of the arguments listed in sec. 6.1 which cannot easily be identified as belonging to one of the three general principles above, but yet found worth mentioning.

b) Special remarks to selected arguments of chainers

Argument A1 (why not, limiting case)

There are basically two drawbacks of this argument, firstly a generalization of index designs is made such that the direct index approach is placed in a bad light[291], and secondly a direct index linked together still differs from a chain index by degrees, if not fundamentally.

1. As pointed out above, the guiding principle of the fixed–basket–approach is comparability within an interval rather than across intervals. The primary concern is keeping constant what might disturb comparability, not chaining, to arrive at a comparison of two periods *very* far away. Constancy is the rule, chaining is the exception. The latter applies to *all* direct indices, not only to the fixed–basket–approach. Direct indices referring to different base periods in general will not be multiplied. By contrast in the chain approach links are *necessarily* chained together in order to derive an index. What was an exception is now the rule, and what used to be the rule (keep things constant) is now abandoned as completely as possible. Bluntly speaking, argument **A1** is a misinterpretation of the direct index approach and its intentions.

2. Even if chaining of direct indices were done, for example a new base defined at $t = 5$, and the direct indices at base 0 were chained to those at base 5 there would still be a remarkable difference: $\bar{P}^{LC}_{09}$ is the product of nine factors, using information given in no less than ten sets of prices and nine sets of quantities, such that altogether 19 vectors affect the result

$$\bar{P}^{LC}_{09} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \frac{\sum p_2 q_1}{\sum p_1 q_1} \quad \ldots \quad \frac{\sum p_8 q_7}{\sum p_7 q_7} \frac{\sum p_9 q_8}{\sum p_8 q_8},$$

while $P^{L*}_{09} = P^{L}_{05} P^{L}_{59}$ is the result of chaining two direct indices, and hence the product of *two* factors only. P^{L*}_{09} is influenced by five vectors $\mathbf{p}_0, \mathbf{p}_5, \mathbf{p}_9, \mathbf{q}_0$ and $\mathbf{q}_5$ only. There are 14 more vectors entering the formula $\bar{P}^{LC}_{09}$. Different data for $\mathbf{p}_1, \ldots, \mathbf{p}_4, \mathbf{p}_6, \ldots, \mathbf{p}_8$ as well as for certain quantity vectors will yield different results for $\bar{P}^{LC}_{09}$. But they will not affect P^{L*}_{09}.

3. Finally there is an interesting exception of the why–not–logic: for adherents of chaining who find the idea that annual chaining is better than chaining at five year intervals convincing, it should be surprising that monthly chaining is *not* better than annual chaining, or weekly chaining is not better than monthly chaining.

[291] In sec. 6.3 we identify the drawing of unfair parallels or making of certain unfair generalizations as one of the five tactics to advocate skillfully for an index approach.

By the way, it is again the base in the *link* P_t^{LC} and not the base of the chain that is updated every year, and where an updating might be accelerated, the base of the *chain* $\bar{P}_{0t}^{LC}$ is always 0, as it is in P_{0t}^{L}.

Argument B1 (moving comparison, additional information)
As to the claim, that a chain index represents a superior type of analysis of time series, we should emphasize once more that apparently the nature of the difference is misunderstood. What makes the difference between $\bar{P}_{0t}^{LC}$ and P_{0t}^{L} is not a "moving" or "rolling" pattern in the comparison. We have a binary comparison (of two states, 0 and t) in both cases, in $\bar{P}_{0t}^{LC}$ no less than in P_{0t}^{L}. What counts is only that P_{0t}^{L} is not affected by prices and quantities of other periods than 0 and t, while $\bar{P}_{0t}^{LC}$ is, because it is path dependent.
Moreover, as noted above (sec. 6.1) already, to construct an index capable of comparing periods 0 and t indirectly, i.e. by accounting for all intermediate periods 1, ..., $t-1$ in addition to 0 and t does not necessarily mean that a chain index has to be adopted.
The problem with the strange "valuable additional information" is not only, that nobody seems to be able to specify the nature and value of this "information", allegedly gained by chaining, but there also seems to be a false inference underlying this argument[292]:

if there is more empirical data (reflecting more phenomena) entering a formula, i.e. we have more *data input*	⟶	**then** the resulting statistic is more "informative", that is then we also get more *information output*

This is blatantly erroneous, and as pointed out in sec. 4.1f and 4.4.d statistics affected by *many* (possibly uncorrelated) influences are in general just *inferior* rather than superior to statistics which are determined by a few known factors only.

Argument B2 (independence of the reference base)
We have already remarked that the independence (or rather irrelevance) of the *reference* base (RB) and thus the pretended "solution" of the problem of the choice of the base period, seems to contradict the greatly increased emphasis put on the *weight* base (WB). It is widely accepted that there is no point in comparing indices having different bases in the sense of different RB, but surprisingly there is not much scruple to study time series of indices like $\ldots \bar{P}_{0,t-1}^{LC}, \bar{P}_{0,t}^{LC}, \bar{P}_{0,t+1}^{LC} \ldots$, in which in fact each element has a different WB. Note that

1. the irrelevance of the RB is simply derived from the fact that the chain is *defined* by multiplication of links, and

[292] The following misunderstanding seems to be quite widespread. So we find the additional–information and the better utilization of all time series data argument in the quoted (in sec. 3.3c) article of the four Dutch authors AL, BALK, DE BOER and DEN BAKKER (1986), not renown as notorious defenders of chain indices. This shows that argument **B1** seems to be subliminally convincing.

2. the difference regarding the relevance or irrelevance of RB and WB respectively, gives rise to the question of the meaning of “base” in the chain index context.

As to 1 the interpretation given to eq. 6.1.1, or equivalently to

$$\bar{P}^C_{04}/\bar{P}^C_{03} = \bar{P}^C_{14}/\bar{P}^C_{13} = \bar{P}^C_{24}/\bar{P}^C_{23} = P^C_4 \tag{6.2.1}$$

is usually that a chain index of any type (therefore the symbol C), not only the Laspeyres chain index (LC) is no more tied to one base (say 0) than to another (say 1, or 2). Thus there is no need to bother with the problem of choosing an appropriate base period. But it should be remembered that (as shown in sec. 4.1g) irrelevance of the base is *not* inherent in the data, but rather resulting from construction of the index. It is ensuing from restrictive assumptions of proportionality imposed on the time series $P_{01}, P_{02}, \ldots$[293]. Moreover as noted in sec. 4.1d different subdivisions of a given interval between 0 and t will yield different results of $\bar{P}^C_{0t}$.

As regards 2 the point becomes apparent when we ask questions like these: What makes the choice of the base period difficult in the direct index framework? Is it possible to choose a “wrong” (inadequate) base in the chain index framework?

The numerical result of a direct index, like P^L_{0t} is said to reflect the price level in period t measured *in terms of* the level in 0. The value of P^L_{0t} is expressed “in percent (in units) of the base period value”. Hence the price level in period 0 acts as a sort of yardstick, such that it matters which period is taken as the “base”. In this sense we may also require that the weights are taken from 0, as this is the period chosen as a *standard* (100%). And thus it is also not surprising that in contrast to eq. 6.2.1 growth rates between to periods (3 and 4 for example) referring to different base periods (e.g. periods 0, 1, or 2) will *not* be equal.

In a sense the notion of a “base” is fundamental for all “comparative”, or “relative” figures as indices with necessity are, and this also implies some kind of *constancy*. Or in other words:

> “A change through time can be recognized only by means of some element which persists unchanged ... by means of a measuring rod which is transported through time without deformation” (AFRIAT (1977), p. 3).

In the chain index approach, however, the “base” RB is simply the beginning of the chain, and WB is no longer a fixed measuring rod. Once a decision is made on the starting point (RB) the sequence of weights (WB) is also *uniquely* determined. We cannot legitimately say that P^{LC}_{0t} represents a level in percent (in terms of the structure of some “basket”) of the *base* period 0, since there are now structurally changing baskets of various periods that are affecting the result. And we cannot choose a “wrong” base (in the sense of “inadequate” weights) either. To sum up:

[293] This incidentially throws cold water on the argument (**B1**) of making better use of the information given in the (time series) data

> In a sense the choice of a base is not solved but rather dissolved by the simple reason that the "base" no longer has the significance it used to have in the direct approach. In the case of chain indices there is simply nothing left (like disproportionality of time series with different RBs, or a non–automatic determination of the WB, once the RB is fixed) which might make the choice of a base period important and at the same time difficult.

Argument B3 (growth factor/rate)

Interestingly there is a lot of scruple to calculate actual growth rates of e.g. volume as this ought to be compiled on the basis of the *most recent (relevant)* prices, but on the other hand, there is no scruple to compare successive growth rates in which this structure of prices is constantly changing, or to compare successive price levels in which this structure of quantities (as weights) is constantly changing. As for the period to which quantities refer a comparison of eqs. 6.1.2 and 6.1.3 shows that $\bar{P}^{LC}_{0t}$ may be preferable to P^{L}_{0t}. However, this argument would not apply to the direct Paasche index[294] (instead of P^{L}_{0t}) to the same extent because

$$\frac{P^{P}_{0t}}{P^{P}_{0,t-1}} = \frac{\sum p_t q_t \sum p_0 q_{t-1}}{\sum p_0 q_t \sum p_{t-1} q_{t-1}}. \qquad (6.2.2)$$

Hence if the sole objective were to calculate growth rates (of volumes [quantities] or prices) on the basis of non-obsolete weights a direct Paasche index could do the job equally well. There is no need for a chain index approach for this reason.

Argument C1 (updated weights = better weights)

The problem with all arguments focusing on the most recent (e.g. previous year) weights which thus purport to be the most representative and most relevant weights, or simply *the better* basket of a chain index, is

- the idea that a chain index has a *single* (and better) WB as a direct index P^{L}_{0t} or P^{LC}_{t}, is false (in reality, $\bar{P}^{LC}_{0t}$ has a *cumulative* WB),

- there is no clear concept to define, or measure the *degree* of "representativity" or "relevance", hence the argument in itself cannot give any hint concerning the best frequency of a renewal of weights.

Moreover the length of the time-interval to which a link should refer, is also determined by other considerations (e.g. cyclical price movement). Given that a price index ought to reflect new *quantity* weights (as required by argument **C1**), then what is

294 The same is true for certain "superlative" direct indices in which weights are taken from both periods, 0 *and* t as well. Hence it should be remembered that **B3** again does not give sufficient reason to prefer chain indices over *all* sorts of direct indices

the task of a *quantity* index? In a way argument **C1** raises the question of the "division of labor" between a price index and a quantity index respectively.[295]
Finally there are two objections to be made from the point of view of official statistics:

- Argument **C1** is clearly two edged; the other side being the increased demand for empirical studies in consumption: experience shows, that compiling representative *weights* repeatedly in short intervals is much more burdensome, expensive, and liable to errors than reporting *prices* on a regular (e.g. monthly) basis.[296]
- An index design allowing for weights, types of goods, outlets and many other features of an index to change constantly is open to various kinds of manipulation.[297]

Argument C2 (quality adjustments less difficult)
The argument should rather read as follows: *quality adjustments no longer necessary*. As mentioned above chainers maintain that chain indices "automatically", or "naturally" "incorporate" various sorts of changes, like changes in the qualities and types of goods, outlets etc. which to account for is a difficult problem for the fixed basket approach of the Laspeyres direct index P_{0t}^{L}. Thus $\bar{P}_{0t}^{LC}$ has "solved" this problem.
This is a typical "solution vs. dissolution" argument and it is particularly unfair, because

- it is left vague what is meant by the term "incorporate", or it rather turns out that new products and new qualities are "incorporated" without any adjustments, since as the basket is allowed to (or even bound to) change constantly with the passage of time there is no point in ensuring comparability of the basket in t with the basket in 0;
- it is on the other hand exactly praised as one of the main advantages of a chain index, *not* to be in need of ensuring comparability for more than two adjacent periods only[298];
- thus the problem, the solution of which is praised, in fact is no longer existent once the fixed–basket approach is substituted for the chain approach, and what counts more: comparability is discarded which is hardly a progress over a method aiming precisely at this comparability in a time series.

[295] We already mentioned in sec. 2.1c this problem of the "division of labor" between a price index, bound to reflect a change in *prices* (at constant *quantities*) and a quantity index in the context of a quotation of FORSYTH and FOWLER (1981), p. 234 criticizing the direct Laspeyres price index as " unacceptable because it utterly fails to represent the changes taking place over the time span" (by which they meant the changes in*quantities* [in the consumption pattern]).
[296] See sec. 3.7 for more details.
[297] See sec. 8.1 for more details.
[298] When it is an advantage not to aim at long term comparability we cannot construct another additional advantage consisting in the alleged "solution" of a problem which exists only because of aiming at long term comparability.

> To make adjustments for changes in the quality of goods is necessary in the fixed basket approach in order to ensure comparability over time. When the fixed basket is abandoned, and comparability is no longer an issue, there is no need for more than just two adjacent periods, $t-1$ and t to be comparable. In short: the problem has simply disappeared.

To sum up, the "solution" of a most tricky problem involved in keeping a basket fixed or comparable consists in replacing the method which gave rise to the problem in question, by a method in which comparability is irrelevant altogether, and in which comparison is based on a number of different baskets rather than on one only.

The Laspeyres type direct index P^L_{0t} has been criticized for keeping constant quantities, and thus for using weights which become progressively irrelevant. This also makes the index increasingly more difficult to compile with the passage of time. But this is done only for the sake of comparability. And note, the problem, P^L_{0t} is blamed for, and $\bar{P}^{LC}_{0t}$ allegedly has "solved" would no longer exist once comparability is taken less seriously (as is the case in $\bar{P}^{LC}_{0t}$). Hence it is deceptive to speak of a "solution".

> It is true that the compilation of a chain index in practice is less troublesome as less care has to be taken for comparability of commodities over time. The increase in convenience has to be contrasted, however, with the fact that chain indices require more expensive empirical studies of consumption, and that now comparability, which the direct index is decidedly aiming at, is abandoned and there is a cumulation of weights affecting the index.

In conclusion $\bar{P}^{LC}_{0t}$ has not solved a problem of P^L_{0t}, it has only replaced one problem by a new type of problem, and it is far from clear wherein the advantage of this should consist.

6.3 Five rules to argue in favor of chain indices

To round off our rebuff of the arguments of chainers we finally present some facetious "rules" to advocate the cause of chain indices. At this stage we may imagine some kind of a "game" which has already been played in index theory for decades[299] and the rules of which in principle seem to be as follows:

[299] We frankly admit that some of the arguments put forward to support our standpoint may also be inspired by the "rules" of this game as set forth above.

- make use of vagueness (up to obscurity[300]) and of arguments appealing to emotions,
- fight considerations in terms of "models", suggest they would contradict "reality",
- mix up concept– and result–oriented considerations in arguing for / against a formula,
- praise / play down selectively advantages / disadvantages of index formulas[301],
- find an appropriate parallel to other approaches, or generalization of approaches.

We now comment on the "rules" one after the other.

1. Appeal to emotions, vagueness, and mysticism

Much of the liking or disliking of a formula is clearly owed to certain emotions and to the ensuing vagueness of "criteria" to justify a formula. When a choice has to be made among two approaches, one flexible and referring to the latest weights and the other which makes use of "old" and unchanged weights, we can easily predict that the number of those voting for "inflexible" and "out–dated" will tend to zero. It is a safe bet that criteria having a strong intuitive appeal like for example "making more use" of existing data, providing valuable extra–information,[302] or representing a "balanced", or "symmetric" solution, a "compromise" or so, allowing for more "relevant" and "representative" weights, etc. will always be found convincing. With this kind of rhetoric the choice is predetermined. This is true in particular if no mention is given to consequences of such a choice, and if – as is the case frequently – the alternative is set out in a conspicuously incomplete and onesided manner.

Statements related to the variable "time" certainly have an emotional charge as can be seen by the so called "axiom of simultaneity", invented by U. P. Reich (see sec. 7.3a). Accordingly we should move from thinking in "levels" to thinking in "changes", present time gains dominance over past time to an extent that it even fell into oblivion that an index by definition always has to make some kind of comparison with the past. In Reich's world, all that counts is timeliness and speed. And in his affection for up–to–dateness, the need for some kind of "base" almost disappeared altogether and he runs into mere absurdity. Another offspring of this obsession with speedy adjustment is mysticism according to which for example (indirect) comparability emerges from (direct) in–comparability (Martini–type of reasoning).[303]

301 for example the "multiplication mystery" or the notion of "valuable additional information".

301 more distinct: exaggerate advantages and play down possible disadvantages of your favorite formula.

302 It is unlikely to find somebody who prefers to make *less* use of data, or to do without certain valuable information delivered free as a kind of by-product. Such ideas clearly *sound* convincing, though we hardly hear anything about the kind of information in question and the kind of insight gained from them.

303 An idea found explicitly or implicitly in arguments **A3**, **B1**, and also **C1** and **C2**.

As it is with all emotional notions the underlying idea is difficult to describe in tolerably exact or operational terms. To our knowledge there is no agreed upon measure of "representativity" or "relevance" of weights developed, let alone an attempt made to quantify the degree of harm done by using "less" (to which degree?) relevant weights. Nonetheless the consensus is universal that the latest weights are urgently needed, and the obsession with the updating of weights is epidemic. Even so it is seldom if ever questioned, and most probably also largely unreflected.[304]

2. Confounding "model" and "reality"

Another rule consists in knowingly mistaking an analytical device for a statement describing the real world. Strictly speaking the only variable exclusively representing "raw material" of observations is the value index V_{0t}, whereas all sorts of decomposing it into a price and quantity component have to rely on hypothetical reasoning of one kind or the other. Many people have difficulties accepting the need for a "model" or "hypothetic" considerations. Notwithstanding there is absolutely no point in creating the impression that

- we could do without some kind of "model", and that
- all "assumptions" of a model have got to be as "realistic" as possible.

To set out an example for the need for a model in statistical measurement: It is a well known fact that "life expectancy" e_x at the age of x cannot be defined and measured by asking people how long they expect to continue to live when their age is x years. The only possible method instead is to make use of the model of a "life table" population in which death risk depending on age is kept constant (the same for each cohort, such that the same age means the same risk, irrespective of the birth cohort to which one belongs).

Without such a model though clearly in contradiction of observation and real world conditions (such as increased longevity as a result of progress in medicine from which it follows that probabilities of dying are *not* constant but falling) measurement of life expectancy would be impossible. It is nonesense to say, life expectancy were incorrectly or "inaccurately" measured, because the assumptions of the underlying model are unrealistic. However exactly this is done in the case of the Laspeyres-index, where the "assumption" of a constant basket has much the same function of a practical tool as has the constancy of death risks.[305]

[304] It is also conspicuous that none of the axioms introduced in index theory expressly refers to abovementioned concepts, like "representativity", "relevance" or so, possibly by the simple reason that it will be difficult to express them in exact mathematical terms.

[305] To derive life expectancy at birth, that is e_0 we have to account for that amount of mortality all those people, who are presently at an age of 0, 1, 2, ... 100, face. We assume that such figures will also prevail when the person just born in question will reach the respective age, that is in 1, 2, ..., 100 or so years later. Thus conditions observed at the moment are assumed to remain constant for up to about 100 years. And there is no criticism. But the analogous constancy of a "basket" for only some five years or so, just as much assumed for the sake of simplifying (if not even rendering possible after all) analysis is almost universally condemned.

Once the instrumental nature of the "fixed basket" as a tool is recognized it is clear that the Laspeyres formula cannot, or should not be attacked with such trivialities like "the American economy is flexible and dynamic"[306]. Nobody would deny this, nor is there any need to do so if we decide to defend the Laspeyres principle. It is bluntly nonsense to insinuate that this principle requires to ignore facts and that in case of other index formulas there is no recourse to some type of model.
The following statement rules out or "outlaws" certain models in which prices and quantities are independent: "The very concept of a basketful of commodities, the content of which remains unchanged while prices change, is ... in *contradiction* to ideas which are basic to the central body of price theory .... If it is possible to work out anything like a rational theory of price of living we must take account of the fact that in an ordinary market quantities will change when prices change" (FRISCH (1954), p. 408, emphasis by Frisch). Such reflections are popular, yet if we would consistently follow them, we would have virtually no basis for deflation or distinguishing with one another a price component from a quantity component contributing to (and making up in combination) the observed change of value. Hence such purism leads nowhere.[307]
Consider a situation in which, in accordance with this kind of "theoretical" reasoning exclusively, such statistics were permitted to be calculated in which prices and quantities correlate precisely as they correlate empirically. Consequently all isolated variations of only one variable are also banned. Is such a situation desirable? Definitely not. Apart from values, value ratios and unit values there were no other statistics admitted. Above all there were no price indices, neither in a direct nor in a chain version, because some hypothetical element was involved in all such calculations. From this it follows that

- to assume a fixed basket does not mean that a consumption is not responding to changing prices, or that the economy is static, and therefore
- such an assumption, made for analytical purposes only, should not be ridiculed with such trivial statements that the American economy is dynamic, or that an isolated change of one variable is unlikely to occur in reality.

In other words, such a pseudo theory to criticize the "realism" of a model is totally sterile and dispensable. On the other hand there is always necessarily some kind of a model (with more or less realistic assumptions) underlying the measurement of a phenomenon not directly observable, like for example "price level" or "inflation".[308]

306 ADVISORY COMMISSION (1995) (1996) (or simply "Boskin Commission").

307 In this sense HABERLER (1927), p. 67 wrote (in our translation) most scrupulously, alluding to the factor reversal test: "It is essentially absurd to search for the respective causal contributions of two factors to the result of a multiplication", or (p. 66): "What is the economic sense of ... asking which part of a change in total turnover is attributable to a change of prices, and which part can be assigned to a change of quantities?"

308 Recourse to some kind of model is particularly desirable in case of measurement of economic phenomena, because not much can be said about the "meaning" of an empirical result without a clear idea of what is intended to measure. This leads to the constant request for an underlying *economic* background

The idea underlying the so called "true cost of living index" (COLI) is for example that a minimum cost combination of goods exists to attain a certain utility level, and that households will suffer from inflation exactly to the extent in which these minimum cost are rising. This concept[309] may be found useful in order to arrive at a reasonable index function, but without going into detail it should be noticed that

- the assumptions made in this model are by no means more realistic than the concept of a fixed basket, and likewise
- the COLI would *not* be proved irrelevant or antiquated if we succeeded in showing that more and more households were unable to maximize utility in reality, a behavior which is however, "assumed" in the framework of the COLI[310].

3. Empirical results played off against conceptual foundation

There is a tendency to mix up concept–oriented and result–oriented aspects in justifying the choice of a formula, just like there is a tendency not to keep clearly distinct "tool for measurement" on one hand, and "description of reality" on the other. As a rule a conceptual foundation of a formula is desirable, and is given by showing how the formula in question fits to what is intended to be measured. But this in itself does not give rise to expect a certain numerical result of the index formula when applied to data.

There is no unique relation between the two criteria, conceptual foundation and empirical results. The latter can never be a meaningful criterion as such, because it is always possible to find a formula able to yield a constantly low inflation rate, to display a "smooth" development, or to remain consistently within the span of the Laspeyres and Paasche formula, whatever the data may be. Thus such a kind of criteria can never be convincing, nor can they make up for a lack of conceptual basis[311]. Interestingly much of the justification of chain indices is of exactly this kind, however. Furthermore, thinking in terms of numerical results which may be interpreted as some kind of average is possibly a still widespread after–effect of Fisher's ideas.[312]

---

of index formulas. However, we rarely meet people who complain about lacking microeconomic foundation of the covariance, for example. Interestingly it is just in the case of indices that some kind of theoretic underpinning in terms of economic theory is found indispensible.

309 See also sec. 8.2 .

310 Things change, however, once the empirical result of a traditional index function, like Laspeyres or Paasche is identified as the value the COLI would be taking. This can only be done if "utility maximizing" behavior is assumed.

311 It is also conspicuous that not much conceptual consideration can be found to justify the typical feature of the chain approach, that is the operation of "chaining" (multiplying links). The whole chapter 4 above is devoted to precisely such problems. There is not much *theory* of chaining in articles of chainers, but rather much allusion to vague, emotive and intuitively appealing criteria like "relevance", "flexibility" and the like.

312 The same is true for the tendency to prefer "compromise"–formulas, resulting from crossing of formulas and / or weights, which – most interestingly – is done exactly when sufficient reasons cannot be found for either of the two formulas or weights respectively.

4. Managing conflicting requirements and ignoring counter–arguments

Most of the formulas and solutions proposed and which still play a part in index theory are astoundingly old. A "rediscovery" is not infrequent in this field. The reason for this is probably that there are a great number of conflicting "axioms", "tests" or so, that can be "discovered" and found plausible or reasonable. Not surprisingly axiomatic theory cannot come up with one single formula passing all tests and being the "best" without any doubt[313]. As there are always criteria met by an index and at the same time those that are violated, a very popular game, of course is to underline heavily the importance of the first and to play down the significance of the second type of axioms. This game can be played infinitely, because no rule can be established according to which a certain axiom should be regarded as more important and thus should be met at the expense of another which might be more readily sacrificed. Since there are always conflicting requirements we might stipulate[314] it is illusive to expect a "solution" of the index problem, or a "proof" showing that this or that kind of reasoning is the only tenable one.

In our criticism of the chain index approach, for example we found it irremissible to ensure comparability, to keep a *price* index, clearly separated from a *quantity* index, and finally to give much emphasis to aggregation (of the index function as well as of volumes). With a different preference order concerning the objective of measurement, however, we might of course attain an entirely different position. In a sense there is always an element of "taste", or "intuitive appeal", and plausibility in play. But what we ought to do in any case, is to work out and make explicit our criteria for evaluation, to aim at an as complete as possible account of all advantages and disadvantages of a formula, and to avoid onesidedness. This is rather the opposite of rule number 4.

5. Inappropriate parallels and generalizations

We have already quoted an example of this strategy in sec. 1.2, showing that Eurostat apparently tried to reconcile chainers and non-chainers with a certain interpretation of eqs. 1.2.1 and 1.2.2, according to which chain indices are the more general concept (MAKARONIDIS (1999)).

In a similar manner we found the SNA stating that "sooner or later" each type of index has to be linked together to form a chain (see sec. 6.1a). The common feature of considerations like these is that they place the direct index in the bad position of a rather strange and unfounded *exception* from the rule. In applying the rule the chain index happens to turn out to be the much better reasoned, more general, and also more "natural" concept. There is for example no obvious motivation or any good reason given why a direct index should *not*

[313] It would be naive to expect another result. Thus contrary to some often expressed opinion this is absolutely irrelevant in deciding about the usefulness of the axiomatic approach. It is sometimes maintained that this approach has proved useless because of not leading to *one* best (in *all* respects) formula. This is nonesense, much the same as saying mathematics is useless, because it cannot solve the squaring of the circle.

[314] We may also invent new "axioms" for the sole purpose of supporting our position. Reich's "axiom of simultaneity" (sec. 7.3a) is a good example for this.

- constantly make adjustments of quantities also (since prices are adjusted) in eq. 1.2.2, while chain indices make adjustments of both, prices as well as quantities, and

- why P_{0t}^{L} should *not* update weights more frequently, which is not only possible, but also common practice in the case of chain indices.

In both cases the chain index appears to be doing what is more logical and more natural, whereas the direct index absolutely unintelligibly refrains from doing so. The trick of tactic number five is that it seems strange *not* to argue for chaining.
We encounter the same kind of an unfair treatment of certain index formulas when by some reason or other the "economic theory"–index or COLI–index is deemed to be the most general and fundamental concept, such that all indices should be evaluated in terms of approximating the only valid COLI–concept. Mr. Laspeyres and Mr. Paasche did not start with the notion of the "true cost of living index", and it was not their intention to find an upper bound (P_{0t}^{L}) or a lower bound (P_{0t}^{P}) respectively of this index. The rationale of P_{0t}^{L} and P_{0t}^{P} is different and we cannot do justice to such formulas if we value them from the "economic theory" point of view only. Seen this way index formulas, like P_{0t}^{L} and P_{0t}^{P} will necessarily end up as inferior indices. This is particularly regrettable for two reasons:

- P_{0t}^{L} did not really have a fair chance, because the focus of the "theory" is on aspects with no relevance for the motivation of the formula, and in turn all those ideas crucial for the justification of the formula of P_{0t}^{L} are irrelevant from the point of view of this theory.

- The economic theory view is all but a legitimized arbitrator as it is in itself not beyond doubts.

In conclusion P_{0t}^{L} as opposed to $\bar{P}_{0t}^{LC}$, is not correctly valued, if P_{0t}^{L} is primarily seen as an extreme approximation of the COLI, or as an index, which is by some oddity only imperfectly and too infrequently updated. With overarching "theories" of this kind we can easily manage a good result for $\bar{P}_{0t}^{LC}$ at the expense of a fair evaluation of other indices, like P_{0t}^{L} for example.

7 Divisia index and its relation to chain index numbers

The key element of Francois Divisia's index approach in 1925 was to introduce time as a continuous variable. It is assumed that two functions, $p_i(\tau)$ and $q_i(\tau)$ exist for each commodity ($i = 1, \ldots, n$) at *any* point in time. Since in practice conditions like these are unlikely to be found, the significance of Divisia's methodology for official statistics is most limited, and will probably remain so in future. On the other hand there is certainly a kind of nimbus surrounding Divisia's index as being something of theoretical perfection, in the sense that an index formula used in practice should approximate this type of index. Another point is the assertion that chain indices simply "translate" Divisia's index into the reality of price observation at discrete points in time. The purpose of this chapter is to show that there is no reason to regard the Divisia index as a theoretical "ideal" and that the relationship between this approach and chain indices is not such that it is capable of providing a theoretical foundation or justification of the latter.

7.1 Procedure to derive Divisia's "integral index"

Let $P(\tau)$ denote the price level function varying continuously over time and let $Q(\tau)$ denote the quantity level function defined analogously. Both $P(\tau)$ and $Q(\tau)$ are assumed to be functions of n continuous price functions $p_i(\tau)$ or the corresponding quantity functions $q_i(\tau)$ respectively for each individual commodity $i = 1, \ldots, n$ at each *point* t in time as mentioned above. It is a mere matter of definition that a value function $V(\tau)$ exists as follows

$$V(\tau) = \sum_{i=1}^{n} p_i(\tau) q_i(\tau), \qquad (7.1.1)$$

denoting an *absolute* value at τ, a point in time (τ continuous). $V(\tau)$ is not a value index measuring value at time $\tau = t$ relative to $\tau = 0$, but a value *level*, just like $p_i(\tau)$ and $q_i(\tau)$ also denote absolute individual ($i = 1, 2, \ldots, n$) levels. After aggregating over i we get $P(\tau)$ and $Q(\tau)$ which are absolute aggregative levels of prices and quantities. As τ is just a point in time, there is no problem in postulating functions $P(\tau)$ and $Q(\tau)$ to have the property that their product is

$$V(\tau) = P(\tau) Q(\tau). \qquad (7.1.2)$$

Unlike the functions $p_i(\tau)$ and $q_i(\tau)$ the levels $P(\tau)$ and $Q(\tau)$ are unobservable. They eventually will lead to a "price index" and "quantity index" respectively. Eq. 7.1.2 only *defines* the "levels" $P(\tau)$ and $Q(\tau)$ implicitly by stipulating a relation between $V(\tau)$, $P(\tau)$ and $Q(\tau)$. It is worth mentioning that:

- in eq. 7.1.2 no instruction is given, how $P(\tau)$ and $Q(\tau)$, and the "indices" derived from them should be calculated in practice, and

- contrary to a frequently expressed belief it is not true, that "factor reversibility" [315] somehow enters into eq. 7.1.2 as an "assumption".

The elegance of Divisia's approach rests on the idea of imagining a point in time τ, or a sufficiently small interval from τ to $\tau + \Delta\tau$ which simply is so short that no third (structural) component will set in. The interval is so short that there is no room for a structural component (consisting in *correlated changes* of P and Q) to be effective and observable. Hence factor reversibility is an implication of time (τ) being a continuous variable,[316] but no longer valid when a discrete time approximation to the continuous time approach is made.
The next "logical" step is to consider differential changes of $V(\tau)$ according to eq. 7.1.1

$$dV(\tau) = \sum_i q_i(\tau)dp_i(\tau) + \sum_i p_i(\tau)dq_i(\tau). \tag{7.1.3}$$

Dividing both sides by $V(\tau)$ as given in eq. 7.1.1 leads to

$$\frac{dV(\tau)}{V(\tau)} = \frac{\sum q_i(\tau)dp_i(\tau)}{\sum q_i(\tau)p_i(\tau)} + \frac{\sum p_i(\tau)dq_i(\tau)}{\sum p_i(\tau)q_i(\tau)},$$

and

$$\frac{dV(\tau)/d\tau}{V(\tau)} = \frac{\sum q_i(\tau)dp_i(\tau)/d\tau}{\sum q_i(\tau)p_i(\tau)} + \frac{\sum p_i(\tau)dq_i(\tau)/d\tau}{\sum p_i(\tau)q_i(\tau)}. \tag{7.1.4}$$

The next idea is to do the same in the case of eq. 7.1.2 (using the product rule of differentiation)

$$\frac{dV(\tau)}{d\tau} = Q(\tau)\frac{dP(\tau)}{d\tau} + P(\tau)\frac{dQ(\tau)}{d\tau} \tag{7.1.3a}$$

or

$$\frac{dV(\tau)/d\tau}{V(\tau)} = \frac{dP(\tau)/d\tau}{P(\tau)} + \frac{dQ(\tau)/d\tau}{Q(\tau)}. \tag{7.1.4a}$$

This eq. simply states the well known fact that the (continuous time) growth rate (or logarithmic deriviate) of $V(\tau)$, a *product* of $P(\tau)$ and $Q(\tau)$ is the *sum* of the growth rates (rates of change) of the two factors, $P(\tau)$ and $Q(\tau)$ respectively.

---

[315] Eq. 7.1.2 only postulates the existence of two functions, P and Q such that their product equals V. It turns out that they are specified in such a way that one function is derived by interchanging arguments of the other function. Hence we may say that factor reversibility holds *by definition* for points in time or infinitesimally small intervals, but not necessarily for a long interval from 0 to t.

[316] It should be noted that there is no specific assumption in eq. 7.1.1 and 7.1.2 apart from τ being continuous, a fact responsible for the widespread belief that Divisia'a approach is something "natural", or indisputable.

An obvious way of finding an index P or Q is to relate the right hand sides of eq. 7.1.4 and 7.1.4a, and to identify the first (second) term as the growth rate of the unknown price level (or quantity level)

$$\frac{dP(\tau)/d\tau}{P(\tau)} = \frac{d\ln P(\tau)}{d\tau} = \frac{\sum_i q_i(\tau)}{\sum_i q_i(\tau)p_i(\tau)} dp_i(\tau)/d\tau$$

$$= \sum w_i(\tau)\frac{dp_i(\tau)/d\tau}{p_i(\tau)} = \sum w_i(\tau)\frac{d\ln p_i(\tau)}{d\tau}, \qquad (7.1.5)$$

where weights $w_i(\tau)$ are expenditure shares at point τ (and hence of course changing with time) and summation takes place over n commodities ($i = 1, \ldots, n$).
In the same manner the growth rate of the function $Q(\tau)$ is a weighted arithmetic mean of growth rates of n functions $q_i(\tau)$

$$\frac{dQ(\tau)/d\tau}{Q(\tau)} = \frac{d\ln Q(\tau)}{d\tau} = \sum w_i(\tau)\frac{d\ln q_i(\tau)}{d\tau}. \qquad (7.1.6)$$

Thus the two functions $P(\tau)$ and $Q(\tau)$ are defined

- using their relation to $V(\tau)$ (eq. 7.1.2), and
- via their continuous time growth rates derived from differential equations (eqs. 7.1.5 and 7.1.6).

Calling $\frac{dP(\tau)}{d\tau}/P(\tau)$ the growth rate of the *price* level (and $\frac{dQ(\tau)}{d\tau}/Q(\tau)$ the growth rate of the *quantity* level) is justified on the following grounds (sometimes called Divisia's axiom[317]):

> Assume that the quantities in eq. 7.1.4 do not change such that $dq_i(\tau) = 0$ for all i, then
> - the change of the quantity level (that is of $Q(\tau)$) according to eq. 7.1.6 should also be zero, or in other words
> - the change of volume (in eq. 7.1.4) should equal the change of prices.
>
> This applies mutatis mutandis for the assumption of no change in prices $dp_i(\tau) = 0$.

It is this consideration that allows to separate the two differentials (in prices and in quantities) and to identify them as growth rate of price and quantity level respectively. Interestingly apart from the length of the time interval in question this idea is not much different from interpretation given to P^L_{0t} in terms of the "pure price comparison":

[317] VOGT (1979), p. 50.

P^L_{0t} measures how cost (value) had changed when quantities had remained constant. A change in prices taken in isolation (quantities kept constant) should result in $V_{0t} = P_{0t}$ and $Q_{0t} = 1$.
But to assume this is not enough to derive P^L_{0t} and this again shows that such highly general considerations will not suffice to justify index constructions. Assume that *both* types of variables will change ($dp_i(\tau) \neq 0$ and $dq_i(\tau) \neq 0$) then of course Divisia and Laspeyres will offer widely different solutions. Note also that, in contrast to other index constructions Divisia's $Q(\tau)$ and $P(\tau)$ are in fact conceived as some absolute (though not directly observable) figures and that growth of $Q(\tau)$ is in fact to be derived from growth of $P(\tau)$ and vice versa by interchanging arguments in certain functions.[318]
Another way of looking at Divisia's method is indicated in eq. 7.1.5 and 7.1.6:

In essence Divisia's method consists in constructing a (continuous time) growth rate of the (absolute) price level $P(\tau)$, or quantity level $Q(\tau)$ respectively as a weighted sum of n growth rates; weights $q_i(\tau)$ being shares in total value and varying with time.

It became fashionable to call all kinds of aggregates, the growth rates of which are gained from averaging growth rates using variable weights "Divisia". In this sense there is for example a discussion of the use of "Divisia monetary aggregates" expected to be superior to simple (unweighted) sums of money stock components.
The next step to arrive at Divisia's index is usually explained as follows: To get at price index the differential equation (eq. 7.1.5) is to solve ("integrate") for P yielding

$$P(t) = P(0) \exp \left\{ \int_0^t \frac{\sum q_i(\tau) \frac{dp_i(\tau)}{d\tau}}{\sum q_i(\tau) p_i(\tau)} d\tau \right\}$$

$$= P(0) \exp \left\{ \int_0^t \sum w_i(\tau) \frac{d \ln p_i(\tau)}{d\tau} d\tau \right\}, \tag{7.1.7}$$

and thus the price index is given by

$$P^{Div}_{0t} = \frac{P(t)}{P(0)}. \tag{7.1.7a}$$

Correspondingly Divisia's quantity index is

$$Q^{Div}_{0t} = \frac{Q(t)}{Q(0)} = \exp \left\{ \int_0^t \frac{\sum p_i(\tau) dq_i(\tau)}{\sum p_i(\tau) q_i(\tau)} \right\}. \tag{7.1.8}$$

318 The incorrect common interpretation of this is that the factor reversal condition is satisfied (or imposed on) the pair of Divisia indices (whereas in our view factor reversibility in the growth rates is simply a consequence of continuous time; such that there is no place for a "mixed term" (like in eq. 7.2.3 below) on the right hand side of eq. 7.1.4).

The name "integral index" stems from the fact that the pair of Divisia indices is derived from solving (integrating) differential equations. The problem, however, is that the integration suffers from *lack of path invariance*: the solutions (integral functions) of the "line integrals" in eq. 7.1.7 and 7.1.8 depend on the path connecting 0 and t. By contrast the integration

$$V_{0t} = \frac{V(t)}{V(0)} = \exp\left\{\int_0^t \frac{dV(\tau)/d\tau}{V(\tau)} d\tau\right\} = \exp\left\{\int_0^t \frac{dV(\tau)}{V(\tau)}\right\} = \exp\left[f(0,t)\right] \quad (7.1.9)$$

depends on the endpoints 0 and t only not on the (shape of the) path connecting them. Thus

$$V_{0t} = V_{0k}V_{kt} = \exp\left\{\int_0^k \frac{dV}{V} + \int_k^t \frac{dV}{V}\right\} = \exp\left\{\int_0^t \frac{dV}{V}\right\} \quad (7.1.9a)$$

or $f(0,t) = f(0,k) + f(k,s)$. By contrast to $f(\cdot)$ the corresponding functions in eq. 7.1.7 and 7.1.8 are *not* path invariant.[319] Both approaches, Divisia's and chaining have path dependency in common which is often viewed as a major disadvantage as ensues from it (HULTEN (1973), p. 1017f):

- "that for at least one of the paths the index does not return to its initial value",
- that it (the index) "can become arbitrarily large (or small)", and that
- "a multiplicity of index values may be associated with any given point in the set of variables to be indexed" (that is in the comparison of, for example $\mathbf{p}_t$ with $\mathbf{p}_0$).[320]

All these properties (e.g. violation of identity, reaction to cyclical movement in prices, and a binary comparison not uniquely determined) are well known to possess chain indices as well, another reason for "chainers" to believe in an inherent relationship between the two types of indices, Divisia and chain indices. However, it should be noted that not only chain indices but also certain direct indices can be derived

- under specific assumptions concerning the functions $p_i(\tau)$ and $q_i(\tau)$, or
- from various types of discrete time approximations to the continuous time Divisia index

[319] Divisia himself was already aware of this drawback and he also proposed the compilation of chain indices as a discrete approximation.

[320] This consequence of path dependency gave rise to many investigations (we cannot discuss here, however) in index formulas resulting from taking other paths than the empirical path, and in conditions under which the Divisia index becomes path independent (or "non cycling"). An interesting approach, though highly complex and theoretical (and hence not having received much attention), in this respect is the so called "natural index" of A. Vogt with a straight line path (VOGT (1978b), (1979)). See also USHER (1974).

and therefore some *direct* indices may be seen as having an interpretation in terms of Divisia as well. In some textbooks for example, it is shown that the direct Laspeyres index can be derived from some specific (though farfetched) conditions (BANERJEE (1975), p. 126). Consider in τ the same (or for all i proportional) quantities as compared with 0, or assume a (most unlikely) path of quantities such that $q_i(\tau) = \lambda q_i(0) = \lambda q_0$. This gives

$$\frac{dP(\tau)}{P(\tau)} = \frac{\sum q(\tau)dp(\tau)}{\sum q(\tau)p(\tau)} = \frac{\sum \lambda q_0 dp(\tau)}{\sum \lambda q_0 p(\tau)} = \frac{d\sum q_0 p(\tau)}{\sum q_0 p(\tau)}, \tag{7.1.10}$$

where the subscript i, denoting the commodity, is deleted for convenience. Integration of the price differential

$$\int_0^t \frac{dP(\tau)}{P(\tau)} d\tau$$

subject to constant or proportional quantities[321] leads to

$$\ln P(t) = \ln \left[\sum q_0 p_t\right] + C \qquad \text{where } P(t) = P(\tau = t) \tag{7.1.11}$$

with an arbitrary constant C, the value of which can be determined by assuming prices p_0 to enter $P(0)$ as do prices p_t with $P(t)$. This means

$$\ln \frac{P(t)}{P(0)} = \ln P(t) - \ln P(0) = \ln \frac{\sum p_t q_0}{\sum p_0 q_0} \tag{7.1.11a}$$

such that we finally arrive at the familiar Laspeyres index

$$\frac{P(t)}{P(0)} = \frac{\sum p_t q_0}{\sum p_0 q_0} = P_{0t}^L.$$

Of course P_{0t}^L is path invariant as opposed to P_{0t}^{Div}. There are also attempts to show that other direct indices, like Paasche[322] (P_{0t}^P) or Fisher (P_{0t}^F) might also be regarded as special (making some peculiar assumptions, however) cases of the Divisia index.

7.2 Discrete time approximations and chain indices

Divisia's approach differs from other methods of finding index functions at least in three inter–related aspects:

- the focus is on each point of the interval from 0 to t, not only on the end points,

321 Or: in which the individual price changes are weighted with constant base period weights.

322 The method is similar to the way in which the Laspeyres index was derived: assume $q(\tau) = \lambda q(t)$ for all τ.

- to describe the movement of individual price differentials $d\ln p_i(\tau)/d\tau$ are considered, that is *increments* instead of ratios, like p_{it}/p_{i0}, or log changes $\ln\left(\frac{p_{it}}{p_{io}}\right) \approx \frac{p_{it}}{p_{io}} - 1$, and finally it turns out that
- weights $w_i(\tau)$ used in aggregation over commodities (individual price changes) in order to get a price index P (or to get $\frac{dP(\tau)}{d\tau}/P(\tau)$) are not only taken from either or both points 0 and t but they also relate to all points τ in time $0 \leq \tau \leq t$.[323]

This makes clear that a certain resemblance between Divisia's index and chain indices exists. The problem is, however, that there are *various* different discrete approximations to Divisia's index. Aside from path dependence this is one of the main shortcomings of Divisia's approach and thus subject to a lot of criticism.
In what follows we will refer to two approximations only, Laspeyres and Paasche, that is to $\bar{P}_{0t}^{LC}$ and $\bar{P}_{0t}^{PC}$. Substituting forward differences $\Delta V_t = V_{t+1} - V_t = \sum p_{t+1}q_{t+1} - \sum p_t q_t$ for the differential dV (and correspondingly Δp_{it} and Δq_{it} for dp and dq) leads to

$$\Delta V_t = \sum_i q_{it}\Delta p_{it} + \sum_i p_{it}\Delta q_{it} + \sum_i \Delta p_{it}\Delta q_{it}, \tag{7.2.1}$$

an equation equivalent to eq. 7.1.3, but with a mixed element $\sum_i \Delta p_{it}\Delta q_{it}$. It is reasonable to define P_t and Q_t in such a way that

$$\frac{\Delta P_t}{P_t} = \frac{\sum q_t \Delta p_t}{\sum q_t p_t} \quad \text{and} \quad \frac{\Delta Q_t}{Q_t} = \frac{\sum p_t \Delta q_t}{\sum q_t p_t} \tag{7.2.2}$$

and omitting subscripts i for convenience. Using eq. 7.2.1 this gives

$$\frac{\Delta V_t}{V_t} = \frac{\Delta V_t}{\sum q_t p_t} = \frac{\Delta P_t}{P_t} + \frac{\Delta Q_t}{Q_t} + \frac{\sum \Delta p_t \Delta q_t}{\sum q_t p_t}, \tag{7.2.3}$$

or

$$\frac{\Delta V_t}{V_t} = \left(\frac{P_{t+1}}{P_t} - 1\right) + \left(\frac{Q_{t+1}}{Q_t} - 1\right) + R_t, \tag{7.2.3a}$$

where the residual term

$$R_t = \frac{\sum \Delta p_t \Delta q_t}{\sum q_t p_t}$$

will tend to zero and will henceforth be neglected. Thus a growth factor of a "price level" defined on the basis of this kind of reasoning is

$$\frac{P_{t+1}}{P_t} = \frac{\sum q_t p_{t+1}}{\sum q_t p_t} = P_{t+1}^{LC}, \tag{7.2.4}$$

[323] In view of "chainers" especially this third aspect (updating weights) makes Divisia's methodology attractive.

which is the Laspeyres link, and the corresponding *index* (comparing period t with 0) is then

$$\frac{P_{t+1}}{P_0} = \frac{P_1}{P_0}\frac{P_2}{P_1}\cdot\ldots\cdot\frac{P_{t+1}}{P_t} = P_1^{LC}P_2^{LC}\ldots P_{t+1}^{LC} = \bar{P}_{0,t+1}^{LC}, \tag{7.2.5}$$

and analogously

$$\frac{Q_{t+1}}{Q_t} = Q_{t+1}^{LC} = \frac{\sum q_{t+1}p_t}{\sum q_t p_t}, \tag{7.2.4a}$$

and

$$\frac{Q_{t+1}}{Q_0} = \bar{Q}_{0,t+1}^{LC}. \tag{7.2.5a}$$

In a similar manner we may derive Paasche chain indices $\bar{P}_{0t}^{PC}$ (and $\bar{Q}_{0t}^{PC}$) by using *backward* differences $\Delta^* P_t = P_t - P_{t-1}$ and $\Delta^* p_{it} = p_{it} - p_{i,t-1}$ respectively[324]

$$\frac{\Delta^* P_t}{P_t} = 1 - \frac{\sum q_t p_{t-1}}{\sum q_t p_t} = 1 - \frac{P_{t-1}}{P_t} = 1 - \left(\frac{P_t}{P_{t-1}}\right)^{-1} = 1 - \left(P_t^{PC}\right)^{-1}. \tag{7.2.6}$$

On the basis of such considerations chain indices are often regarded as being practical (discrete time) approximations of Divisia's index, implicitly claiming chain indices to have some desirable properties in common with Divisia's index, such as factor reversibility. But (in analogy to eq. 7.1.4)

$$\frac{\Delta V_t}{V_t} = \frac{\sum p_{t+1}q_{t+1}}{\sum p_t q_t} - 1 \neq \frac{\sum q_t \Delta p_t}{\sum q_t p_t} + \frac{\sum p_t \Delta q_t}{\sum p_t q_t} = \left(Q_t^{LC} - 1\right) + \left(P_t^{LC} - 1\right) \tag{7.2.7}$$

such that the pair of Laspeyres chain index numbers will fail the factor reversal test. But again in view of the famous "antithetic" relation it is not surprising that

$$\frac{V_t}{V_{t-1}} = P_t^{LC}Q_t^{PC} = P_t^{PC}Q_t^{LC} \tag{7.2.8}$$

holds. Hence a relationship valid in the continuous time approach (for differentials) need not entail a favorable property which also applies to the discrete time framework (i.e. it also holds for differences instead of differentials).

7.3 Uses and limitations of Divisia's approach

The basis Divisia's index and chain indices have in common, is the idea of subdividing an interval in time into small sub–intervals and measuring a price level by integrating

[324] See ALLEN (1975), p. 181, BANERJEE (1975), p.127. The analogous derivation of Fisher's ideal index given in these textbooks needs assumptions, however, for which an interpretation is difficult to find.

its growth rates related to the sub–intervals of observation, (ideally of infinitesimally small length). The usefulness of this idea and the significance of its consequences[325] is contentious, however, and in what follows we present two widely different opinions.

a) Reich's "axiom of simultaneity"

It is sometimes useful to have a look at an extreme position. The German author U.P. Reich for example went as far in his enthusiasm for most "recent" weights and most "timely" results, that the fact almost entirely fell into oblivion, that an index always has to make some kind of comparison with the past. Reich presented a veritable eulogy of the SNA recommendations and of Divisia's index, and he introduced an "axiom of simultaneity"[326], apparently of his own making to support his view, and according to which the following should hold[327]:

> "An index designed to measure change in prices and volume has got to be an aggregate of data of exclusively that point in time to which it refers. More distinctly stated, when the task is to measure price and volume change in period t data have to be taken from period t and no other period. This is exactly what is done in case of a chain index. It should be underlined that absolute values as well as changes of variables are related to a *point* in time. The mathematical ideal of this notion is the Divisia index".

It is not correct to identify the chain index as an index (or the only index) complying with this strange axiom[328]: this criterion could at best be met by the *link*, *not* by *the chain* (product). To better understand Reich's, ideas we should examine what made him propose his axiom and praise the SNA as the ultimate solution of all index problems we have been impatiently awaiting for decades.

According to Reich we should overcome the duality of *two* formulas, Paasche for deflation and Laspeyres for price level measurement which is in his view intolerable. Likewise an index design depending on which period is chosen as a *base* period is inacceptable for him.

---

325 In our view the above mentioned idea is not that revolutionary as often believed, and more important, it cannot make all other conceptualisations of the measurement problem under consideration worthless.

326 The exact German expression "temporale Eindeutigkeit" (being "unequivocal" with respect to time) is somewhat misleading, and denoting more the consideration leading to this idea than what is required by this "axiom".

327 REICH, (1998), emphasis added. In what follows we present our translation of the German text.

328 Conspicuously Reich refrains from presenting his axioms in terms of a functional equation. It is difficult to see which reasonable index function, if any at all will be able to meet the requirement of this axiom. The axiom also does not give any indication of how to make a distinction, between a price index and a quantity index (or "volume index" according to Reich), and it is obviously left vague what is really intended to measure. To give reasons for a certain formula it is definitely not enough to say to which period the data entering the formula should relate.

In short Reich wanted both, factor reversibility[329], and a measurement without reference to a "base period", and he thought the solution was given by the Divisia indices P_{0t}^{Div} and Q_{0t}^{Div} as defined in eq. 7.1.7a and 7.1.8 respectively. Hence Reich came up with a mere glorification of Divisia's pair of index functions, being factor reversible[330], and exclusively in keeping with his axiom of simultaneousness according to which all data should refer to the same *point* in time. Why does this index seem to come closest to his idea? What made Reich admire the Divisia index as a kind of ultimate solution to all index problems?

In essence Reich's mistake to consider eq. 7.1.4 exclusively and to completely forget about integration, is the same phenomenon as looking at the link and disregarding the fact that in a chain index these links are always "linked together" (multiplied) in order to form a time series. It is not the link but the chain, and correspondingly not the growth rate (eq 7.1.5) but the integral (eq. 7.1.7, and eq. 7.1.7a) that has to be compared with P_{0t}.

More than that. By some strange reasoning it appeared to Reich that the continuous time approach did away with all sorts of "levels" (leaving only "changes" and rates of "growth" of such levels), and in a sense with time altogether. He worked himself up into such statements as the following: We should move from thinking in terms of levels to thinking in "monetary changes", where "all variables are functions in time, belonging to the same point in time" such that Divisia's index is "the only index describing reality", and the only index satisfying his axiom of simultaneity.

On the other hand all sorts of calculations making use of data relating to past periods are only "hypothetical calculations" that can be made "but such statements are not describing facts". In his view it is also the practice of price statistics which for a long time have already tacitly followed this line of reasoning, because what counts is a timely result and nothing but a timely result. To adopt chain indices therefore in Reich's view is simply to draw the logical conclusion from a need for timely data, which is overdue and should be done as promptly as possible.[331]

The reason for discussing Reich's position here in such detail, is because it seems to represent some kind of logical endpoint of all desire to have an update of weights as frequent as possible. As often in cases like these, where an idea has been followed up to a logical end, there is some kind of change into plain absurdity.

---

[329] Reich welcomed the SNA's preference for Fisher's formula in which he saw a "unified" approach making use of one formula only instead of the coexistence of two formulas (Laspeyres and Paasche). There lies perhaps also the motivation for his term "Eindeutigkeit" (uniqueness, unequivocality).

[330] For Reich factor reversibility is a "postulate" instead of an "axiom", because of allegedly being a mere consequence of Divisia's index approach (of eq. 7.1.4).

[331] Of course Reich also tried hard to play down the well–known disadvantages of the Divisia and chain approach, and he takes two of them into consideration, path dependence, i.e. lack of circularity (transitivity) and non–additivity (lack of structural consistency in volumes). To do away with the latter he simply recommended the real income approach, using one single deflator instead of the volume approach to deflation. This is another advice of the "solution – vs. – dissolution" type (see p. 233 above).

When the aim of an index P_{0t} is to compare two situations, denoted by 0 and t no matter how distant these situations are away from one another, it is really difficult to understand why data should exclusively relate to t and data of 0 should be avoided. Take for example an index comparing two countries, say France (F) and Spain (S). Nobody would believe the index would gain validity inasmuch as more data relating to F and less data of S were taken. But in the case of time, things seem to change, present time obviously gains an emotional dominance over past time, and in order to make lags as small as possible a sequence logically dissolves into simultaneity, and speed becomes all that counts. Reich's axiom is a good example for what may happen when the affection for up–to–dateness and speed in adjustment is taken in isolation, and followed to its logical endpoint:

The original idea of an index making a *comparison* may well disappear altogether, and there is also some temptation to invent new axioms for the purpose of supporting ones position, and giving it a kind of theoretical dignity by relating the axiom to the Divisia index.

b) Diewert's skeptical view

Divisia's index seems to rest on little or no restrictive assumptions, and it is hence giving the impression of a most general and unassailable approach. Thus the right, as claimed by chain indices, to use Divisia's name as a reference is often seen as a major advantage. Diewert's positions, with which we agree, consist mainly in showing that the assumptions made in Divisia's approach are all but indubitable, and that there is no exclusive right of chainers to refer to Divisia. According to Diewert most of the doubts concerning the usefulness of the "continuous time approach" for practical purposes of official statistics are related to

1. the meaning of continuous observations in economics, and

2. the solutions found in approximating Divisia's index using discrete data indicate that an unequivocal "approximation" does not exist.

In what follows the position taken by Diewert concerning both topics will be quoted (DIEWERT and NAKAMURA (1993)). In taking infinitesimally small intervals, he said, not only many problems of index construction disappear, such as factor reversibility[332], but also the object of observation itself, since "...the smaller we make the unit of time or space within which production or consumption takes place, the less actual production or consumption there will be to observe, and comparisons between these tiny units will become meaningless" (p. 3). It is also questionable, for example whether the notion of a minute– or hour–inflation will ever be useful.

[332] It is often conjectured that Divisia's solution of the factor reversal test is a most elegant one (see for example NEUBAUER, (1995)), but it should not be overlooked that the solution looses much of its elegance when discrete approximations are made.

> "The problem with this approach is that economic data are almost never available as continuous time variables .... Hence for empirical purpose it is necessary to approximate the continuous time Divisia price and quantity indexes by discrete time data. Since there are many ways of performing these approximations, the Divisia approach does not seem to lead to a definite result" (p. 23).

Due to the fact that a number of index formulas, like Laspeyres, Paasche and Törnquist and many others may be viewed as approximations to Divisia's integral–index it is not surprising that widely differing positions in index theory feel authorized by Divisia's approach. "Chainers" are not the only people who believe in having found a theoretical support in Divisia, a situation casting some doubt on the usefulness of the approach altogether. More important, since the approximations "can differ considerably, the Divisia approach does not lead to a practical resolution of the price measurement problem" (p. 43). Needless to say that we agree with Diewert's skeptical view and disagree with Reich's position.

8 Official statistics and concluding remarks

In this final chapter some aspects of assessing chain indices are added which should not be disregarded, if (as decided) chain indices were used in official statistics. Furthermore in sec. 8.2 we will try to pull it all together, and to characterize once more the main differences between the chainers' and our (non–chainer's) view, especially by expounding once more the idea of pure price comparison.

8.1 Aspects of interpretation and official statistics

This section is devoted to practical aspects relevant mainly to assessing the suitability of chain indices for official (price) statistics. Interestingly the chain index or the Divisia index issue has been discussed mostly in terms of index *theory* only, widely neglecting criteria relevant for *official* statistics. With regard to conditions (national) statistical offices face, the following considerations seem particularly noteworthy:

1. chain indices will require more statistical surveys and thereby will be more costly, both directly for Statistical Offices and indirectly for respondents facing an increased response burden;
2. understandability and neutrality of the approach are preconditions for public acceptance, which is particularly important in the case of inflation measurement;
3. a change of methods has far–reaching implications as official statistics have to provide a whole *system* of indices (not only price indices) which should fit together.

It is conspicuous that for example Bruce Mudgett plays down precisely the first two aspects, costs and lack of "understandability"[333]. By contrast, we are strongly convinced that just this interpretation issue is crucial and it seems no coincidence, that chainers tend to play down questions concerning the conceptual foundation of chain indices like: Is the latest basket automatically also the most "relevant" one?[334] Is a chain index reflecting rising costs of a basket? How to measure "relevance" or "representativity" of a basket? What is the empirical evidence of fixed–base weights becoming progressively more "irrelevant" as time goes on?
Chainers' fixation on the latest possible weights goes largely unquestioned. In a similar vein there is also not much thought given to the third point mentioned above,

[333] The quite often quoted words of MUDGETT (1951), p. 37 read as follows: "The present writer is not too greatly impressed with either the 'understanding' or the 'calculation' reasons for choosing index number formulae . . . As for the 'understanding' issue, we all use automobiles, telephones, radios, and many other gadgets without knowing too much about how they are made." The parallel with some technical devices, like automobiles etc. is most revealing. Mudgett completely disregards the importance of trust and confidence (in impartiality of statisticians) of the general public in the "statistics" business (a *public* good). For such criteria we have no counterpart in the case of (*private*) consumer durables, such as automobiles and the like.

[334] See sec. 4.4a for some conditions necessary for inferences of this kind to be valid.

what would a universal application of the chain index approach in official statistics really mean to us?

a) Difficulties in compilation, family expenditure surveys (FES), cost benefit studies

There is no doubt that chain indices will redound to increased costs in compilation of indices (as compared to P^L_{0t} for example) mainly due to a need for more statistical surveys. It is widely accepted that surveys of consumption patterns (so-called family [or household] expenditure surveys [FES]) are not only costly, but also laborious and not without difficulties for respondents.

Therefore in some countries, especially in those favoring the direct index approach (P^L_{0t}), as for example in Germany, such FES are carried out in five year intervals or so only (or were so at least until recently). To maintain such a practice would also be in keeping with the general principle adopted only recently in most if not all countries' national statistical offices, that

- statistical surveys should be reduced to an indispensable minimum, and that also
- the response burden involved in questionnaires has got to be minimized effectively.

But in order to provide true *chain* indices, as required by international recommendations or regulations, it will be necessary – for countries like Germany also – to give up this practice and perform FES annually. Therefore, however, a move to chain indices clearly contradicts the two points of the agreed-upon strategy outlined above (reduction of surveys and of response burden), it will definitely be more costly[335] and thus appears justified only if the *advantages* of this move were indeed substantial.

But Eurostat nonetheless took steps in precisely this direction. Moreover it will be difficult to speed up investigation in consumption as it is highly unlikely that

- FES, even if carried out annually, will be capable of providing speedily a renewed basket[336],
- an update of CPI "baskets" can be made with the help of statistics other than FES.

[335] This has to be taken into account in the cost benefit relation of a move from direct indices to chain indices. Cost undisputedly increase in both respects, directly as well as indirectly (increasing response burden of private households as well as shops contributing as data suppliers), but this is said to be justified by the "overwhelming advantages" (FORSYTH and FOWLER (1981), p. 230) or "undoubted benefits of an annually linked chain consumer price index" (FORSYTH (1978), p. 354).

[336] That is an update of the (preceding year) basket, just at the beginning of a new year, such that in each year we start right at the beginning with weights related to the year, that has just finished.

In practice baskets will be available with a delay of a few months only, possibly also only provisional baskets which are subject to revisions later[337], and baskets will in general not (yet) be sufficiently detailed into some hundreds or so items for which prices will be collected. Hence it is much more realistic that detailed weighting schemes will have a *two* year lag rather than lagging behind *one* year only, and it is also possible that only parts of the basket (weights) are renewed annually.[338]

Other statistics than the FES will be of limited value to derive CPI baskets more frequently. Production and turnover statistics for example deal with *all* sorts of sales, including those for government expenditure (collective consumption), intermediate consumption, investment and (in case of production) possibly also export. Furthermore as a rule they are also often not detailed enough to provide updates of CPI weights. Doubts may also arise with respect to obtaining a sufficiently detailed and timely update of CPI weights using National Accounts, which is a practice reportedly adopted by some European countries.[339]

Another noteworthy aspect is that chaining makes an index more dependent on (changing) results of family expenditure surveys (FES), and at the same time FES has become more and more complicated for a number of reasons. Accounting for

- more and more variety in commodities, more rapid changes in conditions of purchases, such as types of outlets, terms of payment, services rendered etc, and not the least
- the need for an enhanced international comparability and harmonisation, which also entails a growing complexity, as the constant extension of coverage[340], and refinement of measurement concepts[341] etc.calls for more and more detailed regulations

[337] In price statistics (by contrast to National Accounts) it is for a number of reasons most uncommon to have “revisions” of indices after their first being published.

[338] This will result in *weights of different age* as a full update may perhaps be possible only in longer time intervals. If an annual check of weights is done only partially an index may well combine such weights. In case of the so–called Harmonized Index of Consumer Prices (HICP) in Europe and of the European inflation rate which is calculated by combining these national HICPs – we explicitly have to put up with such a situation.

[339] In sec. 3.6 we discussed some empirical results obtained when a chain index was used instead of the corresponding direct index. We quoted a German simulation study estimating an experimental chain index CPI* (instead of the correct official CPI, Consumer Price Index) in which weights were derived from National Accounts. Of course these weights could not be broken down to such a detailed list as usually underlying a CPI. By the way, the difference between the CPI*, and the official CPI index was much smaller than expected.

[340] For example questions related to the treatment of health services, the inclusion of some more complex types of services such as subscription to the Internet, services of financial intermediaries etc., the decision for a domestic or resident concept of the CPI, the inclusion of institutional households or of certain imputed transactions.

[341] For example a decision on the harmonized treatment of owner–occupied dwelling. Another problem is the comparable measurement of consumption of health and education goods and services across countries, as budget surveys mostly report *direct* expenditures only instead of actual (total) consumption (sometimes entirely or partly funded by government or insurance institutions).

will aggravate problems with FES and thus with the compilation of a chain index. It is all but sure that the *practice* of statistics is capable of deriving valid and definite weights for CPIs updated as timely, rapidly and promptly as required by the *theory* of chain indices.

> The establishment of weights is well known to be the part of index compilation which is more difficult, costly, time consuming and burdensome to respondents, and perhaps also even less reliable and becoming more complicated compared with the regular reporting of prices. chain indices by requiring more frequent FES consist in making precisely just this relatively more problematic element more important and more influential for the results of a price index. Moreover in the European Union decisions for chain indices were apparently taken without detailed cost benefit analysis accounting for those circumstances.

To blame the Laspeyres index for being a too primitive approach ignoring important aspects of changes in a dynamic world is clearly only one side of the coin. The other is to show that detailed information on most up to date quantities is really that important (and influential for the results), and at the same time also available at reasonable cost and with acceptable precision. There is no clear demonstration of advantages of chain indices able to offset the indisputably increased cost and practical difficulties in index compilation. Moreover it is not necessarily better to follow a more ambitious measurement concept only imperfectly[342] than to provide a less sophisticated information more readily available and more likely to be understood and interpreted correctly. Or to quote a well known phrase in statistics (ascribed to J. M. Keynes): it is better to be roughly correct than precisely incorrect.

b) Acceptance, understandability, room for manipulation

An aspect which due regard should be paid to especially in official statistics is to avoid all sorts of ambiguity of the results and dubiousness about their derivation. In view of the role of official statistics aspects of interpretation, understandability and of a clear meaning of an index are of utmost importance. It is unwise to play down the importance of such aspects, as there are few if any points equally essential to official statistics as acceptance by the general public, trust in neutrality and confidence in professional integrity of statisticians.

It is sometimes said that the general public will not understand that a Laspeyres index P^L_{0t} is based on an "old" basket with obsolete weights, hence marketing for P^L_{0t} will be difficult nowadays. In repudiation of this quite common argument it is noteworthy that

[342] Attempts to more completely account for the growing complexity and dynamics of consumption in index calculations will not only introduce more aspects in which results will be less comparable, but also where mismeasurement might arise.

- on the other hand it is often said, that the advantage of the Laspeyres formula is just its simplicity (alluding to the rising–cost–of–a–fixed–basket interpretation) [343], and
- it does not seem easier to explain such features of the chain index, such as $\bar{P}^{LC}_{0t}$, like non–additivity, or path dependency in general and violation of identity in particular.

In fact if we set great store by understandability of a formula, the formula $\bar{P}^{LC}_{0t}$ has not much to recommend it:

> None of the standard interpretations of index calculations applies to chain indices. They are path dependent, have poor aggregation properties, they are not well-suited to the role of deflators, and they lack a theoretical background. Finally, and most important for an interpretation of results, they do not provide a "pure price comparison", such that a number of different influences may well affect the value of a chain.

To begin with the *"standard interpretations"*: chain indices do not measure the increase in an overall price (or expenditure) of the same basket in two situations, they rather combine several baskets. They neither represent a mean (or linear combination) of price relatives, nor do they reflect minimum costs to attain the same level of utility in exactly (and solely) two situations, 0 and t (or: under two price regimes only). Though the utmost emphasis is put on weights, $\bar{P}^{LC}_{0t}$ is *not* a weighted mean of price relatives a^i_{0t}, such that $\bar{P}^{LC}_{0t}$ does not need to satisfy the condition $\min\left(a^i_{0t}\right) \leq \bar{P}^{LC}_{0t} \leq \max\left(a^i_{0t}\right)$. Results of $\bar{P}^{LC}_{0t}$ may increase or decrease beyond limits in case of cyclical price movements, and they may also differ from unity despite identical prices in periods 0 and t respectively. It is not sure that such peculiarities are easier to understand, and accept than constancy of weights for some years in a direct index.

As to the *theory underlying chain indices*: if there is any deliberate concept going beyond the urge of using the most up-to-date weights it seems contradictory (like the relation between chaining and chainability and some other aspects as shown in chapter 4) and arguments put forward to plead for chain indices are doubtful to say the least (as shown in sec. 6.2). That chain indices do not provide a *"pure comparison"* has various dimensions. In fig. 8.2.1 (see page 274) for example it will be shown that chain index comparisons are "impure" in *two* respects as they are influenced by other variables than prices, and by more than just the two points in time, 0 and t which are to be compared.[344] In the context of official statistics, what is primarily important is the following fact:

[343] It should be noted, however, that in our days (unfortunately and wrongly) "simple" is often found disadvantageous.

[344] See sec. 8.2 for some more aspects of "purity" that should be considered.

> The principle of keeping a basket constant is first and foremost a kind of safeguard against all sorts of manipulation. In view of the distrust statistical offices have to face, they should welcome the two aspects of the Laspeyres formula: its straightforward interpretation and its built–in barrier to manipulate by making changes in the composition of the basket, possibly arbitrarily (or even politically motivated) and aiming at desired results.

official statistics should always reckon with users that tend to be skeptical and distrustful as far as impartiality of statisticians is concerned[345]. On the other hand politicians and governmental bodies are often tempted to present lower "empirical" inflation rates. [346]

There is definitely a much greater chance of successfully lowering inflation (or what is statistically measured as such), when a more flexible index design is used as compared to the rather rigid model of a fixed basket. The latter approach would make it much more difficult to manipulate and wield political influence,*if* so wanted, and this is a great advantage for official statistics that should not be underestimated. It is important that official statistics in applying strict and rigid principles is indubitably *seen* to be free from such influences. Wherever continuing strongly held and diverging views are met, like in inflation measurement, rigidity and simplicity in methods is nothing inconvenient, but rather an invaluable advantage. To make this clear, a personal remark should be added.

The present writer has learnt an awful lot about politics and statistics from a study concerning official statistics in the former GDR (V.D. LIPPE, (1999)). It was common in the GDR that inflation was prevented from existing by two rather simple devices:

- a rising price was consistently offset by an equal rise in "quality", or
- the commodity becoming more expensive was replaced by another the price of which "had dropped or at least remained the same"[347].

Of course nobody argues for chain indices *because of* their offering more opportunities to manipulate, but on the other hand – for some users at least – to fall back on more flexible methods of price statistics, if possible with the intention to lower inflation rates is not unattractive.

[345] In this respect the situation seems completely different to Mudgett's automobiles, telephones or radios etc.

[346] It should also be remembered that many ideas of the Boskin Commission (or Advisory Commission, see footnote 290) for example were decidedly motivated by the desire to lower the inflation rate. To introduce new goods and new outlets most rapidly as well as to account for substitution by more frequently changing the basket was discussed expressly under the headline "correction for overrating inflation". The title "ways to lower inflation" would describe the intentions perhaps more directly and correctly.

[347] V.D. LIPPE, (1999), p. 12; thus for the selection of goods it was important that the price *declined*. By the same token in output figures only those goods were reported where production has *risen*.

c) Applying the chain principle to all kinds of indices and to double deflation

Once chain indices are introduced for price level measurement and deflation purposes there is a strong temptation to make use of this principle in *all* kinds of indices, not only price indices but also production indices, indices of new orders and the like. Once all types of indices are chain indices, this would have significant consequences for

1. results of statistics defined as relations between (e.g. ratios of) two indices or certain methods combining two or more indices like for example double deflation, or for

2. defining growth rates, endpoints of intervals, turning points, leads / lags, phases of the business cycles etc. which is usually made on the basis of certain indices.

1. Statistics defined in terms of indices

Ratios made up of indices in the numerator and denominator are for example the *"terms of trade", "productivity", real wages* and real salaries etc. They heavily rely on the fact that each index reflects a well defined phenomenon of its own such that the combination of two such indices in the numerator and the denominator respectively has a clear meaning. In cases where an exact concept of what the indices are measuring is lacking the interpretation of the statistic will be difficult, more difficult even than in the case of a single index taken in isolation.[348]

It would be found unwise for example to measure *productivity* by comparing a measure of output based on the structure of 1995 with input based on the 1990 production structure. By the same token things become difficult when quantity (volume) indices of output and input are used, each of which reflect a complex mix of influences (not only physical growth of output and input).

A comparison of output and input *prices* is also involved in the *double deflation* method as described in sec. 2.3d, and commented below with special regard to the possibility that both price indices are *chain* indices. Conspicuously there is not much known about the consequences of a combination of chain index and double deflation, both methods renowned for possibly yielding counter–intuitive results.

2. Phenomena measured by making use of indices

When a chain production index is used for *"dating"* business cycles instead of a direct production index which one of the two sets of "turning points" (direct or chain?) diagnosed will be the more reliable one? Furthermore in sec. 3.3c we discussed the problem of two sets of *growth rat*es for GDP, using a direct approach and a chain approach respectively. It became apparent that the results may well differ substantially,

[348] Little is known for example about the possible interpretation of *terms of trade* when each of the two indices, that is the indices of export and of import prices respectively reflects to some extent its own possibly quite different "history", and the impacts of the two "paths" on the overall ratio "terms of trade" are widely obscure and difficult to disentangle.

there was no clear superiority of one method over the other, and things are likely to become even more difficult when not only one single set of growth rates (for example of GDP) is analyzed, but the relation between two or more sets of chain-based growth rates as opposed to direct-index growth rates. How for example in the case of universally applied chain indices will

- the length of certain phases, the position of turning points in business cycles, or
- the development of "structural" changes as reflected by decompositions into subindices of the same index or by comparing indices of different kind, or
- the lead–lag–structure[349] between two time series, or the relation between the growth rates of two or more different phenomena

differ from the corresponding result of the hitherto generally applied direct index approach? What do we know about the factors influencing such differences?
It is in fact surprising that much is left vague in this field, since the practice of indices always deals with a *number* of indices and with relating them to one another. Thus when superiority of a new formula is asserted one would expect that at least some thought has been given to such aspects. But this does not seem to be the case. Moreover we should view this against the background of poor performances in aggregation (over commodities) and of path dependency (i.e. poor performance in temporal aggregation) of chain indices.

A remark on double deflation

It again seems widely unknown how results of double deflation will differ when either *chain* indices of input– and output prices are used or when use is made of *direct* indices for both time series. Double deflation of value added is in both systems, the SNA and the European System of National and Regional Accounts (ESA) regarded as the only "theoretically correct method"[350] without reservation.[351] Thus on the output side each product should be deflated separately by a Producer Price Index (PPI)

[349] To draw conclusions in terms of "leads" or "lags" is for example one of the main reasons why it is found useful to compare time series of new orders (considered as *leading* indicator) with time series of turnover (they are supposed to reflect the end of a sequence beginning with demand [new orders] via production and ending with sales, and thus turnover will be *lagging* behind new orders). Assume both indicators, the leading and the lagging take the form of *chain* indices (instead of the hitherto more commonly used direct indices) known to reflect not only the conditions prevailing in periods 0 and t respectively, but also in the intermediate periods which may be different in the two series compared and which may have in a sense a cumulative effect. What will be the difference? It must be difficult to predict what will be detected as lead-lag-structure when chain indices instead of direct indices will be taken in order to find such structures.

[350] ESA, para 10.28. It is well known that this "theoretically correct" method even if applied with direct price indices of Paasche as deflators, and in the absence of all kinds of measurement problems (which in fact arise in great number, such as valuation of output for own use, government consumption, services of financial intermediaries, definition of a price and volume component in case of taxes, subsidies and the like) may well produce sometimes awkward and counter–intuitive results such as negative value added at constant prices.

[351] In documents dealing with harmonization of deflation methods a distinction was made between three types of methods, from A methods ("most appropriate methods") through C methods: ("methods which

broken down into sufficient detail. By the same token the preferred method of deflating input (intermediate consumption) is a PPI product by product, using prices collected from purchases and taking differences in prices for different purchasers into account.[352] In this context the European Commission voted for the adoption of chain weighted measures of real GDP as the U.S. statistics two years before also moved from a constant prices to a chain index measure of volume.[353]
Now assume PPIs as *chain* indices on the output (O) as well as the input side (I) enter the formula of eq. 2.3.12, that is indices $\bar{P}^{FC}_{0t}(O)$ and $\bar{P}^{FC}_{0t}(I)$ instead of $P^{P}_{0t}(O)$ and $P^{P}_{0t}(I)$. In what follows emphasis is not so much put on "the well–known 'non–additivity' problem" (as said in the aforementioned Decision)[354] and the need to "explain" non-additivity to users of the "volume" data, but rather on the *interpretation* of the underlying formula of "real value added" (VA) so defined. It is interesting to spell out in detail what substitution of $\bar{P}^{FC}_{0t}(O)$ and $\bar{P}^{FC}_{0t}(I)$ respectively for $P^{P}_{0t}(O)$ and $P^{P}_{0t}(I)$ really means.
Let $p_{it}(O)$ and $p_{jt}(I)$ denote output– and input prices respectively and quantities accordingly[355]. To make formulas simple we consider a chain of two links only. It is easy to derive the corresponding more general formula. The deflated output (DO) is given by dividing the output value, i.e. $\sum p_2(O)\,q_2(O)$ by

$$\bar{P}^{FC}_{0t}(O) = \left(\frac{\sum p_1(O)\,q_0(O)}{\sum p_0(O)\,q_0(O)}\frac{\sum p_2(O)\,q_1(O)}{\sum p_0(O)\,q_1(O)}\frac{\sum p_2(O)\,q_2(O)}{\sum p_1(O)\,q_2(O)}\right)^{1/2},$$

and in exactly the same manner the deflated input (DI) is derived by dividing $\sum p_2(I)\,q_2(I)$ by the corresponding index $\bar{P}^{FC}_{0t}(I)$, and we finally have to divide $\sum p_2(O)\,q_2(O) - \sum p_2(I)\,q_2(I)$ by the difference DO − DI giving the following implicit deflator price index:

$$P^{imp}_{02}(Y) = \frac{(1-i_2)\,\bar{P}^{FC}_{02}(O)\,\bar{P}^{FC}_{02}(I)}{\bar{P}^{FC}_{02}(I) - i_2\bar{P}^{FC}_{02}(O)} \qquad (8.1.1)$$

shall not be used"). All "double indicator methods" based on independently observing both, output and input indicators were called "A methods". The advantage of such methods was among other things seen in the possibility to derive an independent assessment of productivity changes. Consequently the "real income approach" to deflate value added using a single deflator or a single indicator of volume were ruled out, or regarded as B or C methods.

352 The problem with this, however, is that satisfactory price indices of this kind rarely exist.

353 EU Commission Decision (as concerns the principles for measuring prices and volumes) Nr. 2223/96 approved by the Statistical Programme Committee in Stockholm, 27th May 1998. By January 1996 the US Bureau of Economic Analysis (BEA) adopted a chain Fisher quantity index as measure of real growth (as recommended by SNA).

354 ESA and the Commission Decision also state " ...that disaggregated constant price data, i.e. direct valuation of current quantities at base–year prices, are compiled in addition to the chain indices for the main aggregates", and " ...it will have to be explained to users why there is no additivity in the tables. The non additive 'constant price' data is published without any adjustment. This method is transparent and indicates to users the extent of the problem", and finally: "These discrepancies cannot be interpreted as indication of the reliability of the results".

355 The subscript i for output– and j for input–goods will be omitted for convenience of presentation henceforth.

in analogy with eq. 2.3.12, or in detail

$$\frac{(1-i_2)\sqrt{\frac{\sum p_2(O)q_2(O)}{\sum p_0(O)q_0(O)}}}{\sqrt{d\frac{\sum p_0(O)q_1(O)\sum p_1(O)q_2(O)}{\sum p_1(O)q_0(O)\sum p_2(O)q_1(O)}} - \sqrt{i_0 i_2}\sqrt{\frac{\sum p_0(I)q_1(I)\sum p_1(I)q_2(I)}{\sum p_1(I)q_0(I)\sum p_2(I)q_1(I)}}}, \tag{8.1.2}$$

where i_t is the input share of output (at current prices) at time t. Of course with more than two links the result will be all the more complicated. This should also be viewed against the background that the result of a chain Fisher deflation even in the case of a "commodity flow" is in itself not easy to understand (as shown in chapter 5). The combination of two methods, in a way contradictory

- double deflation based on the existence of accounting constraints[356], and
- chain indices a method which brings with it the "well–known non additivity–problem", i.e. the violation of accounting constraints in deflation

does not seem to be thoroughly investigated and well understood by now. The results of such a procedure (use of chain indices in deflation of output and input) made mandatory for all Member Countries of the European Union are influenced by no less than three factors

a) differences in the development of output– and input–prices already identified as being responsible for some curious results of double deflation (with direct indices),

b) the "path" of prices and volumes in the intermediate intervals, $1, 2, \ldots, t-1$ as they habitually affect the value of chain indices (by contrast to direct indices), and finally by

c) the degree of violation of balancing constraints as a result of using Fisher – instead of Paasche indices and of using chain indices (see chapter 5).

Hence there is a structural influence (ensuing from problems of aggregation over time as well as over commodities) in addition to a true change in the volume component which has to be taken into account when interpreting sometimes implausible results of double deflation. To sum up:

> Interestingly there is not much discussion (and thus also knowledge) of how results of chain index methods might differ from those obtained with direct indices, whether for example quite different conclusions might be drawn as to lead–lag–structures, double deflation etc.

[356] According to which the difference between output and input should equal value added both in terms of current prices as well as constant–prices.

8.2 Pure price comparison

The idea of "pure" price comparison as introduced already in sec. 0.2c and explored in more detail in various parts of the book thereafter[357], is recognized as a sort of theoretical basis, and the be all and end all of the direct Laspeyres index. In essence chain indices are characterized by their not paying attention to comparability. Consequently the often sadly ignored or misunderstood notion of pure price comparison is a cornerstone in our critique of chain indices. In this final section of the book we therefore once more summarize the "philosophy" underlying this principle. Part a of this section is devoted to an attempt at finding a description of this principle as precise as possible (we found three such "notions" or "concepts"), and in part b and c we will give reasons why the first (part b) or the second and third (part c) notion of "pure" comparison should indeed be carefully taken into account.

a) The meaning of "pure" comparison

1. Three conceptualizations

Though we have already tried to set out the main idea of pure comparison verbally in sec. 0.2c we have to admit that the notion of "pure" in this context is unfortunately not described in sufficiently exact terms. This is all the more astonishing as there is a certain tradition (esp. in Germany) in referring to "pure price comparison", and in expecting that we all know somehow what is meant by this concept. On second thoughts we can easily see that it is difficult to describe the idea with a precision comparable for example to the formulation of axioms, like identity, monotonicity or so.

There are at least three different notions of "pure comparison" we might think of, and which will thereafter be referred to by a two letter code (to be explained later):

1. the **c.p. – concept**[358]: the index P_{0t} (or P_{AB}) should reflect solely the fact that *prices* are different at two periods, 0 and t, or two locations, A and B; or in other words, the index should be affected by *one influence only*;
2. **c.w. – concept**: as a weighted mean of price relatives an index should have *constant weights* (c.w.) for all periods $t = 1, 2, \ldots$
3. **l.p. – concept**: a price index should be *linear* (l) in the *prices* (p) of the current period (and more restrictive even: the weights in the linear combination should be kept constant).

Interestingly all sorts of chain indices are unable to meet *any* of these criteria.

In what follows each of the three notions will be considered in turn.

[357] Most of our criticism concerning chain indices is more or less directly based on the fact that this type of index is at odds with the fundamental idea of making *pure* price comparisons.

[358] c.p.= ceteris paribus.

2. Ceteris paribus (c.p.) concept of "pure"

The first notion is in a sense also the most general idea, such that the second and third concept may be regarded as special aspects of concept number 1, which might be called *ceteris paribus* principle (or c.p.–principle for short), which is Latin and means that an isolated change is under consideration and other ("ceteris") factors are kept constant ("paribus"). These other factors, called price determining characteristics (PDC) in sec. 0.2c remain unchanged, if not in reality then there are at least some provisions made to neutralize numerically the influence of a change in these factors (an example for such a numerically "accounting for" changes are the well known adjustments made in the case of quality change).

The c.p. principle aims at making sure that the logical conditions of comparability are met[359] and thus it aims at the analytical usefulness of an index design in general. The idea in itself, however, is of course not yet sufficiently concrete to gain an index formula. Furthermore the idea is not able to rule out some nonsense formulas.

An *unweighted* mean of price relatives for example, like Carli's (C) or Jevons' (JV) formula, that is

$$P^C_{0t} = \frac{1}{n}\sum \frac{p_t}{p_0}, \qquad \text{or} \qquad P^{JV}_{0t} = \left(\prod \frac{p_t}{p_0}\right)^{1/n}$$

respectively, is not affected by changes in quantities ($q_0 \to q_t$). Since there are only prices entering the formulas, and no quantities existing in them, such indices would perfectly comply with the c.p. principle.

It is of course not our intention to recommend such formulas, and consequently we have to look for a formulation of the principle which also allows for other variables than prices to vary.

It is often maintained that in order to satisfy the c.p. principle successive indices in a series $P_{01}, P_{02}, \ldots$ should be different with respect to prices p_{it} ($t = 0, 1, 2, \ldots$) only. Such a requirement is fairly restrictive, however, and it would rule out Paasche's index as we get

$$\frac{\sum p_1 q_1}{\sum p_0 q_1}, \frac{\sum p_2 q_2}{\sum p_0 q_2}, \ldots,$$

a sequence in which varying quantities $q_1, q_2, \ldots$ will leave their mark, by contrast to

$$\frac{\sum p_1 q_0}{\sum p_0 q_0}, \frac{\sum p_2 q_0}{\sum p_0 q_0}, \ldots,$$

the sequence in the case of P^L_{0t}, i.e. $P^L_{01}, P^L_{02}, \ldots$.

By this criterion we ought to exclude also all index formulas derived from the Paasche index like Fisher's index for example. By the same token all indices with averages of quantities q_0 and q_t or expenditure shares (like Törnquist's, or Walsh's index) are

[359] See page 274 below for more details.

unable to satisfy *this* particular concept of pure price comparison. The same is true for all sorts of chain indices since we get

$$\frac{\sum p_1 q_0}{\sum p_0 q_0}\frac{\sum p_2 q_1}{\sum p_1 q_1}, \frac{\sum p_1 q_0}{\sum p_0 q_0}\frac{\sum p_2 q_1}{\sum p_1 q_1}\frac{\sum p_3 q_2}{\sum p_2 q_2}, \ldots,$$

taking for example $\bar{P}^{LC}_{02}, \bar{P}^{LC}_{03}, \ldots$. The c.p. principle is obviously not met.
On the other hand we implicitly make use of this idea when a distinction is made between price and quantity indices for example. A value index is said to be inappropriate for certain purposes because it is affected by both, changes in prices as well as changes in quantities (the same argument is also used for discarding the "unit value index"[360]).
It is also often argued as follows: to proceed from a value index to a (true) price index, we have to keep quantities constant and allow prices to vary, and conversely we get a quantity index by allowing a variation of quantities while prices are constant. But what is the meaning of "keeping constant" a variable in the index formula framework?

The c.p. concept of "pure comparison" in the sense of "successive values $P_{01}, P_{02}, \ldots$ of a price index should differ with respect to prices only" may well be

- restrictive on the one hand by ruling out indices like Paasche's, Fisher's, etc., and
- too lax on the other hand by accepting unweighted indices, like Carli's or the index of Jevons.

But there is some intuitive appeal in this idea. In making a distinction between a price and a quantity index respectively, or in rejecting unit value indices for example we tacitly make use of this particular c.p. concept of "pure comparison".

The idea is not only restrictive in that it throws water on some well established indices like Paasche, Fisher or Törnquist but it also covers some other aspects of index compilation which are not, however, manifest in the formula itself. It is for example in principle not in accordance with the c.p. idea when the collection of goods or the selection of reporting outlets is changed, or when the index P_{0t} is influenced not only by the situation in periods 0 and t, but also by other periods. Such aspects should be under control because they will have an impact on the result of the index calculation, and even worse, this impact is not separable numerically from the influence of prices. The idea is that as we have to put up with observational instead of experimental data, influences other than prices should at least be accounted for properly. This is common practice for example in the case of adjustments for quality changes.

[360] This type of index is, as everybody knows, supposed to be inappropriate (in place of a price index) precisely because it is affected also by changes in variables other than prices.

chain indices seriously impinge on such ideas in a number of ways, among other things also by depending on the path connecting 0 and t (see fig. 8.2.1, page 274). Moreover there are implications of this concept of comparability which are readily accepted when a single commodity is under consideration while they are strangely neglected when a basket of commodities is considered.

3. The c.w. – concept of "pure" comparison

We now come to the second and third notion of "pure price comparison" which both allude to the formal structure of an index (average, linear function) and which should be referred to for short as

- second concept: constant weights (c.w.), and
- third concept: linearity in prices (l.p.)

as opposed to the c.p. concept. For chainers there seems nothing quite as important as weights in index formulas. Thus the c.w. idea is important, because most if not all of the discussion concerning chain indices is dealing with weights and the period to which they are related. The c.w. concept of pure price comparison requires an index as a weighted mean of price relatives where weights are constant in the course of time. It is not only the existence of constant weights, but also the mean value property which characterizes this concept, a property which is also central to aggregation and disaggregation of index functions. Hence it is reasonable to consider the constant weight issue as an element of "pure" comparison in its own right.
According to this criterion we should accept, for example, in addition to P_{0t}^{L} the following index

$$P_{0t}^{HB} = \frac{\sum p_0 q_0}{\sum \frac{p_0}{p_t} p_0 q_0} \tag{8.2.1}$$

which simply is the harmonic (H) mean of price relatives weighted with constant base-period (B) expenditure shares. However there are several reasons by which this formula[361] would hardly be advantageous and recommendable although it complies with this particular notion (i.e. the c.w. concept[362]) of pure price comparison. Furthermore: while poor index formulas such as the P_{0t}^{HB}-index[363] are included some possibly quite meaningful formulas are excluded. By this c.w. – criterion (if not already by the c.p. – notion) of "pure" comparison ruled out are for example all those indices which account for substitution by entering either the variable (as t varies) quantity q_{it} or both quantities, the fixed quantity q_{i0} and *and* the variable quantity q_{it} into the formula.

[361] It has been recommended by the German statistician W. Neubauer and the present author has devoted a short article to the analysis of this index and the Palgrave index, both repeatedly advocated and brought into play by Neubauer. Cp. V.D.LIPPE (2000a)
[362] As well as the c.p. concept.
[363] See V.D. LIPPE (2000b) for the deficiencies of this formula.

The c.w. concept of pure comparison does not, of course, deny the triviality that the result of price index calculations will differ depending on which weights are taken.[364] The c.w. concepts only states that weights should be taken from the *same* period for all successive periods $t = 1, 2, \ldots$ of an interval of a certain length. It also means that *within this interval* an index should not be affected by a change of weights and no less by a change in the selection of goods, outlets etc. that is by a change in the "domain" of definition of an index. If an index reflects both, changes in prices as well as changes in weights, the two effects should at least be numerically separable. Furthermore:

Contrary to some common prejudice a *price* index (as opposed to a quantity index) does not convey more relevant information when it depends on (variable) quantities in addition to prices. There is also no increase in relevance as the index is more frequently rebased. This applies in particular to an index formula which reflects *cumulatively* all prior weights as it is the case with chain indices, which therefore fail the c.w. criterion.

chain indices miss the c.w. criterion the more severe, the more frequent an update is made as in such cases an even greater number of quantity vectors affects the result of a price index (as shown in sec. 4.2).
On the other hand using constant weights, as done for example in the case of P^L_{0t} but also of P^{HB}_{0t} and lengthening the period in which weights are kept constant, in order to have truly comparable successive values of a price index for periods $t = 1, 2, \ldots$ cannot be the be all and end all of a reasonable index. As any other notion of "pure" price comparison the c.w.–concept cannot be the only criterion, and it is not sufficient to justify the choice of a particular formula because it complies with the c.w.–concept.

4. The l.p. – concept of "pure" comparison

It is also in line with the idea of "pure" price comparisons that there is a clear and comprehensible relationship between individual prices and the resulting price index[365]. Such a situation is desirable as it is the paramount aim of price index compilation to infer somehow from an index to an underlying movement of prices. In the case of linearity (additivity) we have a simple function $f(\cdot)$ relating $\Delta P_{0t} = P_{0t} - P_{0,t-1}$ and the individual price changes, that is $\Delta p_{it} = p_{it} - p_{i,t-1}$ for all commodities $i = 1, 2, \ldots, n$, and

$$P_{0t}(\mathbf{p}^*_t) = P_{0t}(\mathbf{p}_t) + P_{0t}\left(\mathbf{p}^\Delta_t\right) \tag{8.2.2}$$

where $\mathbf{p}^*_t = \mathbf{p}_t + \mathbf{p}^\Delta_t$ holds *if and only if* the index function $P_{0t}(\cdot)$ is linear in the prices[366] (price vectors $\mathbf{p}$).

[364] This is not the point and it were by no means an improvement if an index formula were in fact independent of weights or erroneously *seen* to be independent. This applies to chain indices simply because, in this case weights are *automatically* taken from the successive periods, and therefore there is no problem in choosing a weight base or of handling an abrupt change of weights.

[365] See footnote number 212 for this advantage of linearity.

[366] To be more precise: linear in *reference (current) period* (t) prices. Linearity in base period prices is defined correspondingly, see sec. 2.2.d.

Moreover we found a number of interesting tenets and theorems which are applicable only to the case of *linear* indices, as for example the theorem of L. von Bortkiewicz (see fig. 2.2.6, page 74). It is therefore reasonable to require within the framework of a "pure" comparison that a price index should be linear in prices and a quantity index should be linear in the quantities. This l.p. criterion is met for example in the cases of P^L_{0t}, P^P_{0t}, but conspiciously *not* in the cases of the direct indices P^F_{0t} and P^{HB}_{0t}, and it is all the more missed in the case of chain indices (of all sorts).

In chapter 5 we have set great store on the idea that deflation (in volume terms) should result in linearity in quantities[367] (sec. 5.2a, eq. 5.2.5). It is among other things for this reason that deflation with Fisher's index (direct or chain) has been criticized above. It should be noted that the two notions of pure comparison are independent as shown in tab. 8.2.1 according to which it is possible to fulfill the c.w. criterion and at the same time to violate the l.p. criterion and vice versa.

While the c.w.–notion makes sure that we can meaningfully aggregate over commodities or sub–indices to compile an all–item–index, the l.p.–notion plays a similar part in aggregating over time in addition to aggregating over commodities. As shown in sec. 3.5 there is of course a great difference in terms like

$$P^L_{0t} = 1 + \frac{\sum q_0 \Delta p_1}{\sum q_0 p_0} + \cdots + \frac{\sum q_0 \Delta p_t}{\sum q_0 p_0}$$

as opposed to the corresponding expression in $\bar{P}^{LC}_{0t}$ yielding

$$\left(1 + \frac{\sum q_0 \Delta p_1}{\sum q_0 p_0}\right) \cdot \left(1 + \frac{\sum q_1 \Delta p_2}{\sum q_1 p_1}\right) \cdot \ldots \cdot \left(1 + \frac{\sum q_{t-1} \Delta p_t}{\sum q_{t-1} p_{t-1}}\right).$$

Table 8.2.1: Independence of two notions of pure price comparisons

linearity in prices (l.p.)	**constant weights (c.w.) notion** satisfied	**constant weights (c.w.) notion** violated
satisfied	Laspeyres P^L_{0t}	Paasche P^P_{0t} Walsh P^W_{0t}
violated	HB–index P^{HB}_{0t} Cobb Douglas P^{CD}_{0t} log. Laspeyres DP^L_{0t}	Fisher's direct index P^F_{0t}, all chain indices

Good reasons can be given for each of the three notions of pure price comparison. Before showing this it should be noted that chain indices achieve the feat to be unable

367 This has been stipulated as the PQC (pure quantity comparison) criterion of deflation. It is by the way primarily due to failing PQC that the HB–index (equation 8.2.1) is not recommendable.

to satisfy *any* of the three criteria we might think of when trying to make clear what pure price comparison might comprise.

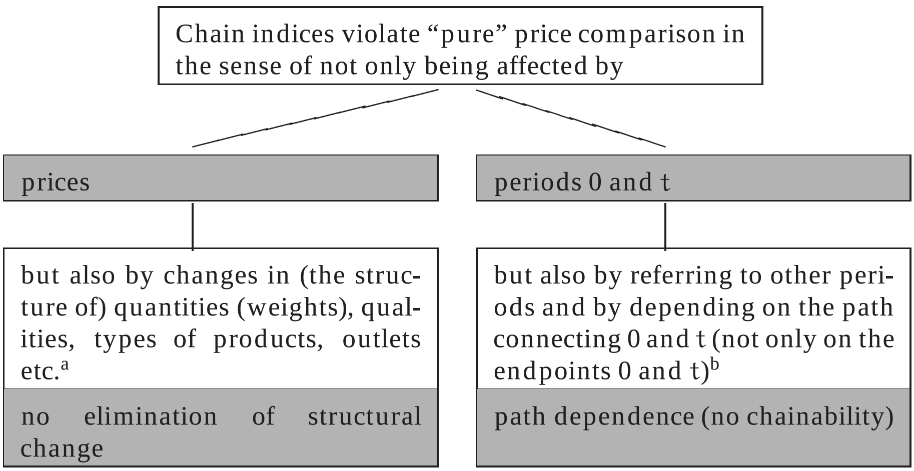
b) Analytical purpose: elimination of structural change and isolation of known determinants in an index

In essence comparability of any two situations requires a distinction between

- common aspects (CA) on the one hand, and
- aspects with respect to which the situations differ (DA) on the other hand,

by which a distinction also is made between aspects which should affect the result (like DA), and those which should not have an influence (like CA).

Figure 8.2.1: Comparability and chain indices (part 1. Dimensions of comparability)



[a] this applies to unit value indices as well for example
[b] in essence "pure" comparison is the idea of making a comparison uniquely determined, but as chain indices are "path dependent" the result of comparing 0 and t is not uniquely determined (as it were in the case of a direct index comparing these two situations *only*), but rather depending on the shape of the time series of prices and quantities connecting these two endpoints.

There is no point in comparing "Thursday" with a "thumb", because these things do not have much in common except for the first three letters in their spelling. Hence for two things to be comparable there should be a minimum number of CAs and not too many of DAs.

Ideally we should have one DA only. By “pure” is meant that situations to be compared should differ in only *one* aspect in order to avoid difficulties (ambiguities) of interpretation, and to *make sure that like is compared with like*. A result influenced by an unknown *mix* of a *number* of influences (DA) is much more difficult to explain, to analyze and to work with. This is the *analytical* purpose of making *pure* comparisons, stressed especially in sec. 4.4d, where we tried to show that much if not most of statistics is devoted to precisely this task of making things comparable[368]. Isolation of a DA and “control” (constancy) of CAs (by elimination of structural changes) is also one of the reasons for developing “indices” as a method in statistics instead of simply comparing averages $\bar{x}$ or unit values.

As shown above (sec. 0.3) we

- prefer a wage *index* to a comparison of simply two *average*–wages ($\bar{x}_t$ and $\bar{x}_0$), or
- a price index to a unit value index (rightly rejected even by the SNA).[369]

chain indices clearly have more resemblance to comparisons of average wages or unit values in the sense that the impact of structural change is incompletely or inadequately eliminated.

> The *analytical* purpose of “pure” comparison is to eliminate the influence of structural change in a comparison and thus to make a comparison uniquely determined by only very few (ideally a single) influence(s). It requires common aspects (CA) and isolation of the DA, or *the* very influence, the variation of which is under consideration.

In the absence of a “common denominator” (or influences kept constant) there is no meaningful comparison possible. Moreover it is only by keeping certain aspects constant (as for example the selection of goods and services, or the weights given to them) in an index that the results of index calculations have certain *known determinants*. [370]

Known determinants and comparability not only among countries, but also over time is particularly important whenever an index is used as a target variable, not only as an informational variable.

---

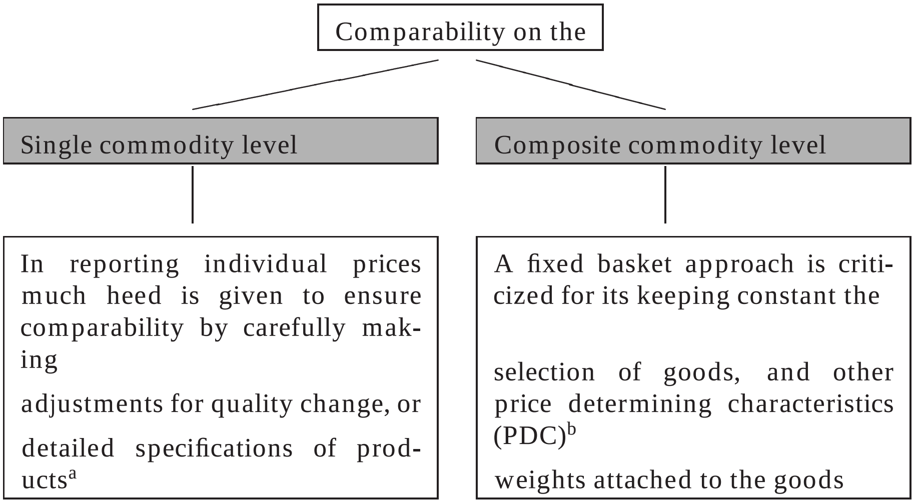
[368] This is why such well known procedures like elimination of “disturbances”, control of sources of variations, isolation of systematic influences, as well as the inspection of data in various disaggregations (breakdowns) always plays an important part in statistics.

[369] See sec. 0.3b and ex. 0.3.1 for a demonstration of this need for “standardization” i.e. making comparisons on the basis of the *same* weights, and with the intention of eliminating the influence of structural change. A *wage index* by contrast to a comparison of *average wages* (or unit values) is *not* disturbed by structural changes, because the *same* weights are used in both situations.

[370] In sec. 4.4d we also pointed out that “axioms” follow the same idea: construction of some simple “stylized” data generating processes, “simple” enough in order to allow for a distinction between meaningful and meaningless results.

This is the case with the inflation rate in the context of a certain monetary policy (and to a lesser degree in the case of the ECB's monetary policy)[371], or when the results of price index or volume calculations are used for far–reaching policy decisions.[372]

Figure 8.2.2: Comparability and chain indices (part 2: Inconsistencies, comparability with respect to the number of commodities)



[a] this applies in particular to goods deemed acceptable and suitable for international comparisons.

[b] See sec. 0.2c for this notion.

It is no coincidence that non–chainers use to emphasize criteria like consistency in aggregation (over commodities). For them it is important to facilitate the decomposition of an index and the detection of sources of variation. Whenever the identification of influences is aimed at, the c.p.– or the c.w.–idea gains importance. Conspicuously such ideas are *not* central in the chain index context, and as shown in fig. 8.2.1 chain indices are in conflict with such ideas in at least two respects, that is with respect to variables (prices, quantities) as well as to intervals in time. Consequently aggregation properties of chain indices interestingly are poor exactly in these two respects: in

[371] ECB = European Central Bank as the decision maker of the European Monetary Union (EMU).

[372] An example of this is the monitoring of national fiscal policies of Member States of the European Union, where we have the so called "Excessive Deficit Procedure" as a part of the "Stability and Growth Pact". The procedure consisting in recommending more or less drastic measures which Member States ought to take, definitely calls for volume figures, which are not only truly comparable, but can also be interpreted in terms of "influences" and "effects". In this context it is e.g. not unimportant whether "volumes" reflect quantity movement only or are influenced by price movement or changes in the structure of prices in addition to quantity (volume) movement.

aggregation over commodities as well as in aggregation over time.
There is another fundamental inconsistency in the chain index rhetoric as shown in fig. 8.2.2: By contrast to the case of a *single* ($n = 1$) commodity where of course comparability (constancy) of types and quality is taken most seriously it is popular to ridicule a *basket* (case of $n \geq 2$ commodities) which is kept constant. “Representative” new weights are deemed much more important than aiming at pure price comparison. In the case of a single commodity or the critique of the unit value index (as opposed to a true price index) the need is in general strongly emphasized, that reference should be made to possibly precisely the *same* goods in both periods 0 and t in order to compare like with like[373].
But curiously the same writers who emphasize such elementary ideas often seem negligent if not oblivious to them once a *basket*, a “*composite* commodity” is concerned. The enthusiasm for an index which purports to account for the alleged need to keep pace with changes, and to reflect all such changes seems to supersede the idea of “pure” comparison. However:

> The fact that nowadays everything changes quickly does not justify to jettison principles of fair comparisons in statistics. Restrictions to what is meaningfully comparable are still valid, and we should continue to observe them carefully, simply in order to avoid vagueness and ambiguities.

Digression on annual growth rates of an index subject to constant methodological changes: the example of the Harmonized Index of Consumer Prices
In what follows we try to describe another inconsistency with respect to the concept of “pure” comparison. Conspicuously in a European harmonization project, the Harmonized Index of Consumer Prices (HICP) of the European Union, great emphasis has been put on comparability over *countries*, but apparently not on comparability over *time*. When a decision had to be taken on how to calculate annual growth rates of this index, there was knowingly no heed given to comparability and this was interestingly again a case in which Germany was outvoted by other Member States. Hereby the widely differing views in European Statistics concerning the importance of comparability once more became apparent.[374]
As it is well known and a most reasonable strategy, the HICP is going to be implemented in a stepwise procedure by constantly widening the coverage of the index. In a step taken in January 2000 some additional types of insurances had been incorporated in the HICP–basket and therefore a decision had to be made how to correctly measure the increase of the price level in general, and the level of the relevant sub–index (expenditures on insurances) in particular in the first months in 2000 against

[373] This seems to go without saying for example in the case of quality adjustment or the making of detailed specifications of commodities in order to ensure international comparability, but interestingly not when the issue of weights is concerned.
[374] Germany possibly is a country in which official statistics puts a particularly great store by comparability and where therefore a chain index is not very popular.

the respective months in 1999 where these insurances were still excluded and hence prices were not yet reported.
It was the (minority) position of the German Statistical Office that the compilation of growth rates should account for the greater coverage. Eurostat's growth rate of the insurance–subindex of the HICP in January 2000 amounted in the case of Germany to 7.1% compared with 5.3% found by the German statisticians, who reported the 7.1-figure to Eurostat "with much scruple and hesitation" only.
The problem in question of course has relevance only for the growth rates in 2000. Beginning with January 2001 the annual growth rate will no longer suffer from differences in the coverage. Hence the problem will exist temporarily only with each new round of expanding the coverage. But as there is an ongoing process of harmonization of price indices in Europe[375] we will possibly have "impure" growth rates in Europe for quite a while.
A rise or decline of such rates is not only reflecting true price movements but also changes in the *methodology* of price statistics and in the guidelines of index compilation. This is most definitely *not* in line with the idea of "pure" price comparisons and therefore also impairing the analytical usefulness of such calculations.

c) Why constant weights?

Much of what the general public and the average user of statistics expects from a measure of inflation might be expressed as follows: a figure showing to which extent rising prices contribute to rising expenditures of the average household. This kind of thinking automatically leads to the idea[376]

- of an isolated price component (or P – component) as opposed to the quantity–, or Q–component[377], and
- the idea of a basket.

What is under consideration is the rising cost of a (or *the*) basket. There is no such thing like "*the* basket" when the collection of goods and their weights is constantly changing or when the cumulative influence of many baskets is studied. In this sense pure comparison requires quite naturally *constant* weights throughout the interval under consideration[378].

[375] This applies not only to the HICP, an index where harmonization is already fairly advanced though still not yet completed, but also to other types of indices.
[376] This, e.g. the cost of a basket, is also by the way underlying the "economic theory" thinking, although in this context the basket is not fixed in physical terms but in terms of, so to say, the "amount of" utility it provides. In any case the focus is not on a multitude of *observed* baskets, as it is in the chain index logic.
[377] It is also far from clear why the isolation of the P– and Q–component in the sum of PQ products (the only observable variable) should best be done with the *same single* index formula as the SNA or the idea of factor reversibility insinuates.
[378] Note that this does not say anything about the period from which the weights should be taken, provided that the period is kept constant for the whole interval over which comparability is aimed at.

Interestingly chainers rarely if ever refer to such considerations and if so we find the idea of making "pure comparisons" clearly misunderstood or mixed up with other considerations, like the "relevance" of weights. FOWLER (1974), p.81 for example is one of the few chainers who explicitly used the term "pure price comparisons" maintaining that chain indices, which he vigorously favors, satisfy this principle. According to his definition (as quoted already on page 47 above): "Any price index between time t_0 and t_n will give an estimate of pure price change if the (quantity) weights are the same in both the numerator and the denominator of the index". A chain index being simply a "logarithmic summation of a series of pure price indices for successive unit time intervals" is of course also " a 'pure' index and not an average value index" (p. 83)[379].

It is hard to imagine a formula *not* providing a pure comparison in Fowler's definition of "pure".[380] In our view the notion of pure comparison implicitly requires that there is *one* set of quantities, which is the same in the numerator and the denominator as is the case for example with the direct Laspeyres index P^L_{0t}, an index formula in our view indeed is a perfect expression of pure price comparison.

FORSYTH (1978), p. 349 maintained that: "The purpose of a consumer price index is to isolate and measure the effect of changed prices on the consumption expenditure of the average household". This is certainly correct, but surprisingly this very sentence is continued as follows: "and this only makes sense if the change is measured on the expenditure which consumers are currently making". We then may ask, however,

- how an index in which all sorts of determinants of expenditures are allowed to vary can "isolate" the price component in expenditures, and
- if the expenditure consumers are *currently* making is most important, why not compile a direct Paasche index instead of a chain index?

From this we may conclude:

When "isolation" of influences and ensuring comparability is understood as aiming at a figure for the "price level" which is *not* affected by changes in other variables than prices (like quantities), this certainly calls for *constant* weights. On the other hand it might be desirable to always make use of *current* or most recent weights. Obviously these are conflicting requirements which calls for a compromise. It is our position that the direct Laspeyres index which is rebased from time to time *is* already the compromise we look for, and are in need for when both goals are to be achieved. The interval between renewals of the basket should be long enough to provide "isolation" and "comparability", but not too long in order to maintain "relevance".

379 The last term refers to "unit values".

380 Moreover even though an individual link may be "pure" it is by no means compelling that a "summation" of "pure" comparisons is also a "pure" comparison.

It is not only simplicity, why linearity is desirable. Can we regard for example

$$\frac{\sum p_4 q_4}{P_{04}^{FC}} = \left(\sum p_4 q_4 \sum p_3 q_4 \frac{\sum p_0 q_0}{\sum p_1 q_0} \frac{\sum p_0 q_1}{\sum p_2 q_1} \frac{\sum p_1 q_2}{\sum p_3 q_2} \frac{\sum p_2 q_3}{\sum p_4 q_3} \right)^{1/2} \quad (8.2.3)$$

in table 5.3.1 instead of $\sum p_0 q_4$ as a meaningful expression of the isolated quantity component (or “volume”) corresponding to the value $\sum p_4 q_4$? Do terms like these really represent a “pure” comparison of quantities in $t = 4$ with those in $t = 0$, not “disturbed” by a *number* of price- and quantity vectors? And why should the number of these determinants rise with the passage of time, such that we get

$$\frac{\sum p_5 q_5}{P_{05}^{FC}} = \left(\sum p_5 q_5 \sum p_4 q_5 \frac{\sum p_0 q_0}{\sum p_1 q_0} \frac{\sum p_0 q_1}{\sum p_2 q_1} \frac{\sum p_1 q_2}{\sum p_3 q_2} \frac{\sum p_2 q_3}{\sum p_4 q_3} \frac{\sum p_3 q_4}{\sum p_5 q_4} \right)^{1/2} \quad (8.2.3a)$$

although there are again two periods only we refer to in our interpretation of this term as “the volume in 5 at constant prices of 0”?[381] When the European Union opted for deflation with chain indices (see sec. 3.4.c and 8.1.c) the reasons given were mainly as follows (as quoted already above): “the most recent base year possible should be used, since in that case the weights used are most up–to–date, and the problems of disappearance of products and new products are minimized”.[382]

If the use of the “most recent”, or “most up–to–date” prices in a volume term really were the only criterion for good “deflation” there would be no point in deflating a “value” to a “volume” at all. Why should we modify $\sum p_4 q_4$, and make use of other prices than p_4, since these are the “most up–to–date” prices? Every other term than $\sum p_4 q_4$ can only be less “relevant” and more “fictitious”. Hence with respect to up–to–dateness of prices in period 4, there is nothing quite as good as $\sum p_4 q_4$, and we better refrain from deflation altogether.

This shows that the widespread fixation with variable and “most recent” weights

- eventually leads to nonesense[383], and it should also be noted that this fixation
- not necessarily calls for chain indices, but also leaves open other options, such as superlative index.

[381] Interestingly such questions are rarely considered. Even worse, there are many who readily believe in such mere figments of imagination like the alleged “better making use of the information in a time series” by chain indices.

[382] In the case of the European Union there are additional aspects worth being discussed: Statistics reflecting a *mix* of influences and needing additional explanations (due to “non–additivity”), and supplementary values deflated by conventional methods will be used for monitoring policy targets within the so called “Stability and Growth Pact”. Contributing to the mix of influences affecting “volumes” are among other things changes in the structure of prices, in conditions for price observations, or in the relative position of the country in the community. This again demonstrates that the idea of “pure” comparison is a much more reasonable criterion of deflation than “most up–to–date” prices.

[383] In our view such a nonesense is also the fact that even for the same interval in time $(0, t)$ the result of chaining is not unique, but depends on the way the interval is subdivided into subintervals.

As there are difficulties in finding, the “overwhelming advantages” of chain indices (FORSYTH and FOWLER (1981), p. 230) we may also question the pretended superiority of other methods in which it is again found inappropriate to work with a fixed basket, i.e. with constant weights.

d) Why linearity in the prices of the current period?

Criteria like c.p. and c.w. are also fulfilled by direct indices like the logarithmic Laspeyres index (or Jöhr's index) DP^L_{0t} as defined in eq. 2.3.1 or by the quadratic mean index P^{QM}_{0t} as defined in eq. 2.2.12, an index which was discussed above for demonstration purposes only, rather than seriously recommended. So such indices would be in line with pure comparison of prices at least in the sense of the c.p. and c.w. concept. They nonetheless do not need to be considered earnestly, which shows that there should be some additional criteria like linearity for example to define “pure” correctly.

Why should we add the criterion of linearity (ideally with constant weights in the linear combination of prices p_{it}), and why are such indices like DP^L_{0t} and P^{QM}_{0t} of limited use only? The idea that linearity (or “additivity” as defined in eqs. 2.2.15, 2.2.16, 2.2.19 and 2.2.20) unequivocally relates n individual prices to one single overall price level is not sufficient. Multiplicativity (eq. 2.2.18) could equally well serve this purpose[384]. A total change in the price level can be decomposed into a *product* of two factors, as there is a *sum* of two components in the case of linearity (eq. 8.2.2). Yet there is quite an interesting difference between a product and a sum: A sum, and arithmetic averaging as opposed to a product and geometric averaging is more closely related to expenditures, i.e. to amounts (*sums*) of money spent for some purpose, and thus such an approach is more akin to deflation. Hence P^L_{0t} is to be preferred to DP^L_{0t}, its “geometric” counterpart, and the l.p. notion should be seen as a legitimate third criterion of “pure” in addition to the c.p. and c.w. notion.

A short digression on the “True Cost of Living Index”

As already mentioned repeatedly a commission appointed by the U.S. Senate Finance Committee, headed by M. Boskin and therefore also known as the “Boskin Commission” (BC) came up with the assertion that the US Consumer Price Index (CPI, hitherto a P^L_{0t}–index) overstates inflation and should be abandoned due to its fixed basket (Laspeyres) approach. Though the Commission readily endorsed all popular beliefs in the alleged damage done by the “growing irrelevance” of fixed baskets it did *not* recommend chain indices[385], but the so called True Cost of Living Index (COLI) or Economic Theory Index, approximated by a (direct) superlative index such as Fisher's index or Törnquist‘s index. This shows that

[384] Needless to add that chain indices (as well as the direct Fisher index) neither meet linearity nor multiplicativity.

[385] cp. footnote 290 on page 233 in chapter 6.2 .

> we may well agree with chainers in downweighting the need of making pure comparisons and recommending to do away with "fixed weights" (the instrumental nature of which being thoroughly misunderstood), and still *not* opting for chain indices.

There is at least one advantage[386] those "superlative" *direct* indices enjoy over chain indices: they are not path dependent because the third source of variation we mentioned in sec. 1.1 does not exist in this case. This is not the place to comment in detail on the COLI approach which the Boskin Commission was quick in adopting and praising. It is more than doubtful that the notion of "constant utility" can ever replace a "constant basket" as a manageable, operational, and as close as possible to observation concept capable of guiding practical statistical work.[387] The Laspeyres approach is not in need of assuming utility maximizing behavior of households, and their ability to reach an optimum immediately after a change in prices occurred. The inflation rate gained by this approach is not built on "subjective", unobservable psychic and emotional conditions and processes of satisfaction. Nor is it built on questionable value judgements.[388] With an index theory conceptualizing the degree of "inflation" as the fair amount of "compensation" in income we inevitably embark on problems of equity. Statistics cannot solve problems which have to be decided politically, and it would be rather better to strive to get the CPI *removed* from the (political) firing line.

In conclusion: we should continue to follow the popular cost–of–a–basket idea, because the options we have, once we abandon this clear and straightforward concept are unattractive. Disadvantages abound in both, the chain index as well as the COLI approach.

---

386 If not a second advantage in that more attention is paid to a theoretical foundation of this index approach.

387 It is our view that the framework of "economic theory" indices is of little use because it does not solve any problem in price statistics and index theory but poses additional problems. It is no surprise that quality adjustment, a field in which utility deliberations come into play, is perhaps the most difficult and despite much scientific effort still one of the least satisfactorily resolved problems. In the economic theory such problems become ubiquitous and the basis of index construction consists in assessing the "amount" of unobservable utility change in deriving a measure of the change in prices, which should be in principle observable.

388 It is often said that such judgements are no longer necessary ever since Diewert developed his theory of "superlative" indices being widely independent of a concrete utility function. Surprisingly the "economic index theory" not only started with the assertion that the utility function matters, and now came up with index formulas which are supposed to be better, because they are independent thereof. Making reference to Diewert, did also not prevent the Boskin Commission from making speculations about what should provide the same degree of utility, or implicitly making some threadbare prescriptions of how consumers should spend their income. There is, in our view at least, another weakness of utility reasoning worth being mentioned. It is the fact that "utility" refers to a much broader concept than observable "goods". It is no coincidence for example that the BC discussed also problems like the increase of variety and convenience in consumption, or the impact of crimes or a lower level in education etc. on welfare (or in general on "utility") in the context of inflation measurement (!).

A final comment on political implications of the issue

As with any other issue in economics we can also in the case of price index problems safely rely on the fact that from time to time more or less the same discussions will be triggered pretending that the measurement methods hitherto applied are wrong and the damage inflicted from falsely following such wrong methods is substantial. We only recently witnessed an example of this in the case of the Boskin Commission. It will definitely not be the last event of this sort. We will always promptly find, when needed, an expert or somebody who styles himself as such, who readily calculates a lower inflation rate or a higher growth rate of real values. It is not so much the "new" method itself by which such a feat is achieved that is interesting[389], it is rather the question why just now, for whose interests, and for which purposes it was brought to the fore (or reanimated). In such cases the temptation is strong to move to methods purporting to be "new" while a stubborn and inflexible abiding by "old" methods is ridiculed. Nowadays the atmosphere in which such decisions are taken is dominated by non–statisticians who lack time, patience and perhaps also knowledge to study in detail the arguments put forward for the new method. This may favor a massive move to new methods everywhere and simultaneously worldwide.

There is, however, some hope that at least some of the decision–makers will eventually return to the old truth that for all sorts of statistics it is requisite that like is compared with like and that the conditions under which observations are made are kept constant for a certain time. As things indeed **do** change constantly we can only make calculations **as if** things were constant. We may also call it an "old truth" that conceptually different processes, such as price movement and response with respect to quantities, should rather be treated as separate phenomena and correlated than mixed up to one "impure" measure.

[389] We also find repeatedly the claim that the correct inflation rate should be compiled under the condition that certain goods are excluded. An example of this is the notion of a "core inflation rate", which came into being only recently. Interestingly such constructions continue to be regarded as a measure of the purchasing power of money, as if money were not spent for the goods excluded.

References

ACZÉL, J., AND W. EICHHORN (1974): “A Note on Additive Indices,” *Journal of Economic Theory*, 8, 525 – 529.

ADVISORY COMMISSION (1995): “Toward a More Accurate Measure of the Cost of Living,” Interim Report to the Senate Finance Committee (USA).

——— (1996): “Toward a More Accurate Measure of the Cost of Living,” Final Report to the Senate Finance Committee (USA).

AFRIAT, S. (1977): *The Price Index*. Cambridge University Press, Cambridge.

AL, P., B. BALK, S. DE BOER, AND G. DEN BAKKER (1986): “The use of chain indices for deflating the national accounts,” *Statistical Journal of the United Nations*, 4, 347 – 368.

ALLEN, R. (1975): *Index Numbers in Theory and Practice*. Macmillan Press, London.

ANDERSON, O. (1949a): “Mehr Vorsicht mit Indexzahlen,” *Allgemeines Statistisches Archiv*, 33, 472 – 479.

——— (1949b): “Über die Neuberechnung von Indexzahlen der Lebenshaltungskosten in Deutschland,” *Weltwirtschaftliches Archiv*, 62, 175 – 197.

ASTIN, J. (1999): “The European Union Harmonized Indices of Consumer Prices,” Paper presented at the Fifth Meeting of the International Working Group on Price Indeces (The Ottawa Group).

BALK, B. (1995): “Axiomatic Price Index Theory: A Survey,” *International Statistical Review*, 63(1), 69 – 93.

BANERJEE, K. (1975): *Cost of Living Index Numbers : Practice, Precision, and Theory*. Dekker, New York.

BARR, A. (1996): “Measuring growth: Is the new US approach better?,” *Economic Outlook*, pp. 18 – 22.

BOSSERT, W., AND A. PFINGSTEN (1987): “The circular and time reversal tests reconsidered in economic price index theory,” *Statistische Hefte*, 28(4), 301 – 315.

BRACHINGER, H.-W. (1999): “Inflationsmessung in der Schweiz: Konzeption und mögliche Verzerrungen,” Diskussionspapier 3, Deutsche Bundesbank, Frankfurt, Beiträge zu einem Workshop in der Deutschen Bundesbank.

CRAIG, J. (1969): “On the Elementary Theory of Index Numbers,” *Journal of the Royal Statistical Society, Series C*, 132(18), 141 – 152.

DIEWERT, W. (1998): "Index Number Issues in the Consumer Price Index," *Journal of Economic Perspectives*, 12(1), 47 – 58.

——— (1999): "The Consumer Price Index and Index Number Purpose," Paper presented at the Fifth Meeting of the International Working Group on Price Indices (The Ottawa Group).

DIEWERT, W., AND C. MONTMARQUETTE (eds.) (1983): *Price Level Measurement: Proceedings from a Conference sponsored by Statistics Canada*. Min. of Supply and Services Canada, Ottawa.

DIEWERT, W., AND A. NAKAMURA (1993): *Essays in Index Number Theory*, vol. 1. North Holland, Amsterdam, London, New York.

EICHHORN, W., R. HENN, AND S. OPITZ (eds.) (1978): *Theory and Applications of Economic Index Numbers*. Physica, Würzburg.

EICHHORN, W., AND J. VOELLER (1976): *Theory of the Price Index: Fisher's Test Approach and Generalizations*, vol. 140 of *Lecture Notes in Economics and Mathematical Systems*. Springer, Berlin.

——— (1983): *Axiomatic Foundation of Price Indexes and Purchasing Power Parities*,pp. 411 – 450, in (Diewert and Montmarquette 1983).

FISHER, I. (1922): *The Making of Index Numbers*. Houghton Mifflin, Boston.

FORSYTH, F. (1978): "The Practical Construction of a Chain Price Index Number," *The Journal of the Royal Statistical Society, Series A*, 141(3), 348 – 358.

FORSYTH, F., AND R. FOWLER (1981): "The Theory and Practice of Chain Price Index Numbers," *Journal of the Royal Statistical Society, Series A*, 144(2), 224 – 246.

FOWLER, R. (1974): "An Ambiguity in the Terminology of Index Numbers Construction," *The Journal of the Royal Statistical Society, Series A*, 137(1), 75 – 88.

FRISCH, R. (1954): "Some Basic principles of Price of Living Measurement; A Survey Article," *Econometrica*, 22(4), 407 – 421.

FUNKE, H. E. A. (1979): "Fisher's Circular Test Reconsidered," *Schweizerische Zeitschrift für Volkswirtschaft und Statistik*, (4), 677 – 687.

HABERLER, G. (1927): *Der Sinn der Indexzahlen: Eine Untersuchung über den Begriff des Preisniveaus und der Methode seiner Messung*. Mohr, Siebeck, Tübingen.

HANSEN, B., AND E. LUCAS (1984): "On the Accuracy of Index Numbers," *Review of Income and Wealth*, 30, 25 – 38.

HILL, P. (1988): "Index Number Theory and External Trade," Eurostat News–Special Edition, Proceedings of a Seminar held in Luxembourg.

HOFFMANN, J. (1998): “Probleme der Inflationsmessung in Deutschland,” Discussion paper, Deutsche Bundesbank, Diskussionspapier der Volkswirtschaftlichen Forschungsgruppe.

HULTEN, C. (1973): “Divisia Index Numbers,” *Econometrica*, 41(6), 1017 – 1025.

MAKARONIDIS, A. (1999): “Fehlerquellen und Mindeststandards beim Harmonisierten Verbraucherpreisindex,” Diskussionspapier 3, Deutsche Bundesbank, Frankfurt, Beiträge zu einem Workshop in der Deutschen Bundesbank.

MARTINI, M. (1996): “The frequency of updating of the base of consumer price index numbers,” Expertise prepared for Eurostat, mimeographed Eurostat B–3, PTF 2 / 96 / 51.

——— (2000): “Index Numbers, Axiomatic Theory and Applications,” Draft version of a forthcoming book.

MC LENNAN, W. (1998): “Introduction of Chain Volume Measures in the Australian National Accounts,” Information Paper of the Australian Bureau of Statistics.

MUDGETT, B. (1951): *Index Numbers*, Wiley Publications in Statistics. Wiley & Sons, New York.

NEUBAUER, W. (1995): “Konzeptionelle Vor– und Nachteile eines verketteten Verbraucherpreisindex,” unveröffentlichtes Gutachten für Eurostat.

OLT, B. (1996): *Axiom und Struktur in der statistischen Preisindextheorie*. Peter Lang, Frankfurt / Main.

PFOUTS, R. (1966): “An Axiomatic Approach to Index Numbers,” *Review of the International Statistical Institute*, 34(2), 174 – 185.

REICH, U. (1998): *Der zeitliche Vergleich von Aggregaten der VGR*,pp. 53 – 90, in (Reich, Stahmer, and Voy 1998).

REICH, U., C. STAHMER, AND K. VOY (eds.) (1998): *Zeit und Risiko in der VGR*. Metropolis, Marburg.

SCHIMMLER, K. (1973): “On National Accounts at Constant Prices,” *Review of Income and Wealth*, 19, 457 – 459.

SCHMIDT, B. (1995): “Der Preisindex für die Lebenshaltung aller privaten Haushalte in Gestalt eines Kettenindex,” Projektarbeit im Statistischen Bundesamt, veröffentlicht in Spektrum Bundesstatistik, Bd. 10, 1997.

SELVANATHAN, E., AND D. PRASADA RAO (1994): *Index Numbers, A Stochastic Approach*. Macmillan, Basingstoke.

SNA (1993): *The System of National Accounts*, Inter–Secretary Working Group. Washington.

STRECKER, H. (1963): "Zur Genauigkeit wirtschaftsstatistischer Grössen," *Metrika*, 6, 133 – 142.

STUVEL, G. (1989): *Index–number Problem and its Solution*. Macmillan, Basingstoke.

SZULC, B. (1983): *Linking Price Index Numbers,*pp. 537 – 566, in (Diewert and Montmarquette 1983).

TURVEY, R. (1989): *Consumer Price Indices: An ILO Manual*. Geneva.

USHER, D. (1974): "The Suitability of the Divisia Index for the Measurement of Economic Aggregates," *Review of Income and Wealth*, 20, 273 – 288.

VOGT, A. (1978a): "Der Wertindextreue–Test und eine Vereinfachung des Indexproblems," *Statistische Hefte*, 19, 131 – 139.

——— (1978b): *Divisia Indices on Different Paths,*pp. 297 – 305, in (Eichhorn, Henn, and Opitz 1978).

——— (1979): *Das statistische Indexproblem im zwei Situationen Fall*. Zürich, Dissertation.

VON DER LIPPE, P. (1999): "Kritik internationaler Empfehlungen zur Indexformel für Preisindizes in der amtlichen Statistik: Einige Bemerkungen zur ökonomischen Theorie der Indexzahlen und zu Kettenindizes," *Jahrbücher für Nationalökonomie und Statistik*, 218(3+4), 385 – 414.

——— (2000a): "Der Unsinn von Kettenindizes," *Allgemeines Statistisches Archiv*, 84, 67 – 82.

——— (2000b): "Preisindizes mit konstanten Wertgewichten: Bemerkungen zu einem Vorschlag von Werner Neubauer," *Jahrbücher für Nationalökonomie und Statistik*, 220(6), 742 – 758.

WESTERGAARD, H. (1890): *Die Grundzüge der Theorie der Statistik*. Gustav Fischer, Jena.

YOUNG, A. (1992): "Alternative Measures of Change in Real Output and Prices," *Survey of Current Business*, 72, 32 – 48.

Index