ECONOMIC AND FINANCIAL INDEXES

by

ALAN G. WHITE

B.A. (Honours), University of Dublin, Trinity College, 1993
M.Litt., University of Dublin, Trinity College, 1994

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES
Department of Economics

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

April 1999

©Alan G. White, 1999
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of ECONOMICS

The University of British Columbia
Vancouver, Canada
Date 29/6/1999

DE-6 (2/88)
Abstract

This thesis examines the theoretical underpinnings and practical construction of select economic and financial indexes. Such indexes are used for a variety of purposes, including the measurement of inflation, portfolio return performance, and firm productivity.

Chapter 1 motivates interest in economic and financial indexes and introduces the principal ideas in the thesis.

Chapter 2 focuses on one potential source of bias in the Canadian consumer price index (CPI) that arises from the emergence of large discount/warehouse stores—the so-called outlet substitution bias. Such outlets have gained market share in Canada in recent years, but current CPI procedures fail to capture the declines in average prices that consumers enjoy when they switch to such outlets. Unrepresentative sampling, and the fact that discount stores often deliver lower rates of price increase can further bias the CPI. Bias estimates for some elementary indexes are computed using data from Statistics Canada’s CPI production files for the province of Ontario. It is shown that the effect on the Canadian CPI of inappropriately accounting for such discount outlets can be substantial.

Another area in which indexes are frequently used is the stock market. Several stock market indexes exist, including those produced by Dow Jones and Company, Standard and Poor’s Corporation, Frank Russell and Company, among others. These indexes differ in two fundamental respects: their composition and their method of computation—with important implications for their usage and interpretation. Chapter 3 introduces the concept of a stock index by asking what, in fact a stock market index is—this is tantamount to considering the purpose for which the index is intended, since stock indexes should be constructed according to their usage. Because stock indexes are most commonly used as measures of returns on
portfolios, the main considerations in constructing such return indexes are examined.

Chapter 4 uses the Dow Jones Industrial Average (DJIA) as a case study to examine its properties as a return index. It is shown that the DJIA is not the return on a market portfolio consisting of its thirty component stocks: in fact the DJIA measures the return performance on a very particular (and unusual) investment strategy, a fact that is not well understood by institutional investors. An examination of some other popular stock indexes shows that they all differ in their computational formula and that each is consistent with a particular investment strategy. Numerical calculations reveal that the return performance of the DJIA can vary considerably with the choice of basic index number formula, particularly over shorter time horizons.

Given the numerous ways of constructing stock market return indexes, the user is left to determine which is 'best' in some sense. The choice of an appropriate (or 'best') formula for a stock market index is formally addressed in chapter 5. The test or axiomatic approach to standard bilateral index number theory as in Eichhorn & Voeller (1983), Diewert (1993a), and Balk (1995) is adapted here. A number of a priori desirable properties (or axioms) are proposed for a stock index whose purpose is to measure the gross return on a portfolio of stocks. It is shown that satisfaction of a certain subset of axioms implies a definite functional form for a stock market return index.

Chapter 6 evaluates the various stock indexes in use today in terms of their usefulness as measures of gross returns on portfolios. To this end the axioms developed in chapter 5 are used to provide a common evaluative framework, in the sense that some of the indexes satisfy certain axioms while others do not. It is shown that the shortcomings of the DJIA as a measure of return arise from its failure to satisfy a number of the basic axioms proposed. Notwithstanding this, each index corresponds to a different investment strategy. Thus, when choosing an index for benchmarking purposes an investor should select one which closely matches his/her investment strategy—a choice that cannot be made by appealing to axioms alone.
# Contents

Abstract .................................................................................................................. ii

Table of Contents ................................................................................................... iv

List of Tables .......................................................................................................... vii

List of Figures ......................................................................................................... ix

Acknowledgments ................................................................................................. x

1 Introduction ......................................................................................................... 1

2 Outlet Types and the Canadian Consumer Price Index ..................................... 8
   2.1 Introduction .................................................................................................... 8
   2.2 Outlet effects in the CPI ............................................................................... 9
      2.2.1 Empirical evidence to date .................................................................. 13
   2.3 Outlet types in the CPI ............................................................................... 15
   2.4 Outlet type as a price determinant ............................................................... 17
   2.5 Bias estimates ............................................................................................. 22
   2.6 Concluding Remarks .................................................................................. 29

3 Stock Market Indexes: Basic Concepts ............................................................... 31
   3.1 What is a stock market index? ..................................................................... 32
   3.2 Constructing a stock market index: preliminaries ........................................ 36
3.2.1 Adjustments to the index ........................................ 38
3.2.2 Fixed base and chaining ....................................... 38
3.3 Three considerations for a stock index ...................... 41
  3.3.1 Choice of basket ............................................ 41
  3.3.2 Treatment of cash dividends ............................ 42
  3.3.3 Chaining and Rebalancing ................................. 45
  3.3.4 Examples .................................................. 46
    1 period examples ........................................... 46
    2 period examples .......................................... 47

4 Stock Market Indexes in Practice ......................... 52
  4.1 The Dow Jones Industrial Average ......................... 53
    4.1.1 A Brief History ....................................... 53
    4.1.2 Features of the Dow .................................... 56
    4.1.3 Adjustments to the Dow Jones Industrial Average . 59
      Timing of stock splits ..................................... 64
      A 'reverse' stock split ................................... 66
  4.2 Other Indexes ................................................ 68
    4.2.1 Standard and Poor's Composite 500 (S & P 500) .... 68
    4.2.2 Value Line Indexes ..................................... 71
  4.3 What return is it anyway? .................................. 73
    4.3.1 Investment strategies ................................... 73
    Dow Jones Industrial Average ............................... 73
    Standard and Poor's Composite 500 ....................... 74
    Value Line Arithmetic ...................................... 75
    Value Line Geometric ....................................... 76
  4.3.2 DJIA recalculation ....................................... 76
5 An Axiomatic Approach to Stock Market Gross Return Indexes

5.1 Axiomatic and Economic Approaches to Index Numbers ................................................................. 85

5.2 An axiomatic framework for stock indexes ................................................................................. 87

6 Stock Market Indexes: An Evaluation .............................................................................................. 106

6.1 Axiomatic Evaluation of Indexes ..................................................................................................... 108

6.1.1 Standard and Poor’s 500 Composite ....................................................................................... 108

6.1.2 Dow Jones Industrial Average ............................................................................................... 108

6.1.3 Value Line Geometric Average ............................................................................................... 109

6.1.4 Value Line Arithmetic Average ............................................................................................... 109

6.2 Concluding Remarks ..................................................................................................................... 112

Bibliography ........................................................................................................................................ 115

A Appendix to Chapter 2 .................................................................................................................... 121

A.1 Outlet coverage ............................................................................................................................... 121

A.2 Sample turnover ............................................................................................................................. 123

A.3 Market shares of outlet types ....................................................................................................... 123

A.4 Statistical significance of average price differences .................................................................. 126

B Appendix to Chapter 5 .................................................................................................................... 128

B.1 Independence of axioms A1-A6 .................................................................................................... 128

B.2 Dependency of axioms A7-A11 ..................................................................................................... 129

B.3 Proofs of Propositions .................................................................................................................. 130
List of Tables

2.1 Outlet-Type Classifications. .............................................. 14
2.2 Commodities for examination. ............................................ 16
2.3 Sample turnover rates. ..................................................... 17
2.4 Index for men’s jeans by outlet type, 1990-1996 (Ontario). ............... 20
2.5 Index for non-prescribed medicines by outlet type, 1990-1996 (Ontario). 21
2.6 Percentage discounts across the outlet types. ................................ 22
2.7 Levels outlet substitution biases based on unit values. ...................... 23
2.8 Adjusted and official index for audio equipment, 1990-1996 (Ontario). . 25
2.9 Adjusted and official index for non-prescribed medicines, 1990-1996 (Ontario). 27
2.10 Adjusted and official index for other household equipment, 1990-1996 (Ontario). 29
3.1 Calculating a 3-stock index ............................................... 50
3.2 Index values based on a 3-stock index ................................... 51
4.1 DJIA Components: 1884 .................................................. 54
4.2 DJIA Components: 1928 and 1998 ....................................... 55
4.3 DJIA versus Laspeyres index ............................................. 57
4.4 Stock splits in the DJIA .................................................... 60
4.5 Stock splits in the DJIA .................................................... 62
4.6 Timing of stock splits in the DJIA ....................................... 65
4.7 A consolidation in the DJIA .............................................. 67
4.8 Index formulae and investment strategies ................................ 77
4.9 DJIA basket updates: 1980-1998 ........................................ 79
4.10 Index values: S&P 500, DJIA, $DJIA_L$, $DJIA_L^{(TR)}$, $DJIA_{EW}$, $DJIA_{EW}^{(TR)}$ 81

6.1 Evaluation of indexes based on axiomatic approach 110

A.1 Percentage distribution of outlet sample by outlet type for non-prescribed medicines, 1990-1996. 122
A.2 Percentage distribution of outlet sample by outlet type for photographic services and supplies, 1990-1996. 122
A.3 Percentage distribution of outlet sample by outlet type for audio equipment and other household equipment, 1990-1996. 122
A.4 Department store categories and associated basic classes. 124
A.5 Market shares of outlet types. 125
A.6 t-statistics for differences in average prices across outlet types. 126
List of Figures

2.1 Index for men's jeans by outlet type, 1990-1996 (Ontario) ............... 19
2.2 Index for non-prescribed medicine by outlet type, 1990-1996 (Ontario) ... 21
2.3 Adjusted and official index for audio equipment, 1990-1996 (Ontario) .... 25
2.4 Adjusted and official index for non-prescribed medicines, 1990-1996 (Ontario). 27
2.5 Adjusted and official index for other household equipment, 1990-1996 (Ontario). 28

4.1 DJIA and $DJIA_L$, 1980-1998 ........................................... 79
4.2 DJIA and $DJIA_L$, 1980-1998 ........................................... 80
4.3 DJIA and $DJIA_L$, 1983-1987 ........................................... 80
4.5 DJIA, $DJIA_L$, $DJIA_{EW}$, 1980-1998 .................................. 83
Acknowledgments

Many people have contributed to bringing this thesis to fruition. First, a huge thanks to my thesis supervisor Erwin Diewert. When I first began studying at UBC I never thought I would hear an index number being described as superlative—it seems only appropriate that I use such a word to describe Erwin's excellent supervision. Erwin's constant encouragement, intellectual stimulation, good humour and friendship have made my graduate days both rewarding and memorable.

My thesis has also benefited enormously from interactions with Alan Kraus. Alan gave generously of his time to extensively comment on the financial aspects of my thesis—it was a pleasure working with him. Chuck Blackorby also provided some provocative comments.

My classmates at UBC have enriched my graduate days, not only by helping me with my work and commenting on my research but also by providing support and friendship during some difficult times. I cherish the times we have spent pondering life and other things over my daily ritual of going for coffee (and an occasional beer). A particular mention to Dolors Berga, Steve Morgan, Mark Neumann, Philip Oreopoulos, Stephen Whelan, and Kevin White. Thanks also to the secretaries in the economics department for putting up with me for so long.

A final thanks to my Mum and Dad, whose moral support and love over the years have propelled me to reach this point in life. I will be eternally grateful for all they have done for me.
Chapter 1

Introduction

In 1995 the U.S. Senate Finance Committee commissioned a group of leading experts on price measurement issues to study the U.S. Consumer Price Index (CPI) produced by the Bureau of Labor Statistics (BLS). Specifically the commission was to report on whether or not the CPI possessed any inherent biases or mismeasurements. The final report of the Boskin Commission\(^1\) concludes that the U.S. CPI is biased upward by approximately 1.1% per annum. Thus, relative to a 'true' cost-of-living index, the Laspeyres fixed-basket U.S. CPI overstates changes in the cost-of-living over time.

The CPI is the most-used and well-known measure of inflation or changes in the cost-of-living. Annual changes in the CPI are used to adjust social security benefits, and wage contracts are often based on formulae that explicitly include changes in the CPI. When one considers the many uses of the CPI, indications that it exhibits upward bias must be given due consideration. Upward biases in the CPI have enormous consequences: the fiscal implications for the U.S. are documented in Boskin, Dulberger, Gordon, Griliches & Jorgenson (1996)—welfare recipients are being overcompensated for changes in the cost-of-living when incomes are indexed to the CPI. Considerable savings to the federal government could be realized by indexing incomes to the bias-adjusted CPI. If the general thrust of the Boskin Commission Report is accurate and is applicable to many countries' CPIs, this

\(^1\)The Commission was chaired by Michael J. Boskin. The other members were Ellen R. Dulberger, Robert J. Gordon, Zvi Griliches, and Dale W. Jorgenson.
will involve a reevaluation and reexamination of traditional views of economic history. For example, real income growth may be larger than economic historians currently calculate; stagnant real wages will be replaced by rising real consumption levels; and the productivity slowdown of the last two decades may disappear. In light of possible CPI mismeasurements inflation targets may have to be adjusted accordingly.

The Boskin Commission has fulfilled a very useful purpose—it has called into question the accuracy and validity of the most closely monitored statistic produced by statistical agencies. Since prices are at the heart of economics, accurately measured prices and price indexes are central to economic measurement. The Boskin Commission has highlighted the strengths and weaknesses of the CPI as a cost-of-living index. The Report will no doubt generate interest in price measurement issues and more general economic measurement issues for some time to come. Indeed the BLS has already taken initiatives to improve the U.S. CPI—these are described in Abraham, Greenlees & Moulton (1998).

Potential measurement errors in economic and financial statistics and their misuse and misinterpretation are not limited to consumer price indexes—it is likely that other economic and financial statistics, indexes and measures are prone to mismeasurement or misuse. Economic statistics and measures are used by both government and industry to understand current economic conditions, to predict future conditions, and to develop, implement and evaluate various economic policies. Business and individual decisions on spending and investment are influenced by barometers of financial market performance. The (negative) consequences then, of using numbers without due consideration to their construction could be substantial.

The present thesis examines some features of economic and financial indexes from both a theoretical and practical perspective. More specifically, in the wake of the Boskin Commission and concerns over the accuracy/inadequacy of statistics, the objective of this thesis is to investigate the construction, validity, accuracy and interpretation of select economic and financial indexes. The implications of mismeasurement and misinterpretation are also briefly discussed. The examination of measurement undertaken in this thesis is conducted along
the following lines—both by asking a number of questions and making some observations:

1. What is being measured, and for what purpose? Once this has been established, economic theory and practical considerations can be used to formulate an ‘ideal’ target measurement benchmark.

2. How can theory and practical concerns be borne to bear on the implementation and approximation of these targets, subject to theoretical, data, and resource constraints? The measurement procedures should have sound theoretical foundations and yet have an easy-to-understand intuition from an applied perspective.

3. How do the measures deal with changes in the economic environment in which the measurement takes place?

4. Relative to an ‘ideal’ situation what (undesirable) properties do current practice measures possess?

5. What are the possible negative consequences of a faulty measurement framework, and how can these measurement frameworks be improved upon (by appealing to both theoretical and practical considerations)?

Although not every aspect of the above five points are addressed in every chapter, these are the issues that one should keep in mind when examining and evaluating index methodology.

Chapter 2 presents an empirical analysis of one potential source of bias in the Canadian consumer price index—the so-called outlet substitution bias. Such a bias arises from statistical agencies inadequately accounting for shifts in retailer patronage from traditional retailing outlets to ‘discount’ outlets offering lower prices. Attention is focused on three elementary indexes—non-prescribed medicines, audio equipment, and other household equipment (i.e. small electrical appliances). The data are taken from the Statistics Canada CPI production files for the province of Ontario, and cover the period 1990-96. It is shown that the entry of (discount) outlets with lower price levels, the fact that discount outlets appear to
deliver lower rates of price increase over time, and unrepresentative sampling have caused an overstatement in the rate of price inflation for these elementary indexes.

Another area in which indexes are used for measurement purposes is the stock market. Several stock market indexes exist, including those produced by Dow Jones and Company, Standard and Poor's Corporation, Frank Russell and Company, among others. Notwithstanding the fact that several different stock indexes exist, all suffer from the same apparent lack of grounding in theory. The indexes produced by the different companies and stock exchanges differ in two fundamental respects: their composition and their method of construction—with corresponding implications for their interpretation and application. Nowhere in the relevant literature on stock indexes is there any strong justification for using one index number formula for a stock index over another. Apart from differences in composition and weights used, there is no rigorous discussion of what precisely differentiates (say) the Dow Jones Industrial Average (DJIA) from the Standard and Poor's Composite 500 Index (S & P 500) in terms of their properties. Much of this void in the literature arises from an inadequate discussion of the uses of such indexes—the intended use of an index will dictate the manner in which it is constructed. Indeed it is unclear from the literature that stock index users understand precisely what the different indexes actually measure.

It should be clear from the above that the issue of mismeasurement is not necessarily (and cannot be) the sole focus of attention in examining stock market indexes. After all, mismeasurement is a relative concept, i.e. an index is mismeasured relative to an 'ideal' benchmark. In the consumer price index situation the Konsis cost-of-living index has been embraced as the conceptual target measure for the CPI (see Diewert (1983)[page 168]). In fact the Final Report of the Boskin Commission begins with the strong recommendation that “the BLS should establish a cost-of-living index (COLI) as its objective in measuring consumer prices” (Boskin et al. 1996)[page 2]. What is lacking from the literature on stock indexes is a discussion of what ‘good’ stock indexes are, what properties they should possess, and the manner in which such stock indexes should be constructed. The remainder of this

2A recent but very informal discussion of stock index properties is found in Fortune (1998).
thesis endeavours to shed light on these issues.

Chapter 3 attempts to provide an answer to the question “what is a stock market index?” This is tantamount to asking the purpose for which the index is intended since stock indexes (and any index, for that matter) should be defined and constructed according to their usage. It is seen that the primary use of stock market indexes is as measures of (gross) returns on particular portfolios of stocks—the term *stock market gross return index* is adopted for this particular use. It is *this* important application which is examined in the remaining chapters of the thesis. Some of the basic features of the stock market which must be taken into account when constructing such indexes are discussed. Chapter 3 shows that incorporating such features into the construction of stock market indexes require fundamentally different considerations from those that arise in the pricing of a fixed basket of goods and services over time (as in the standard consumer price index situation). In particular, an investor's return derives from both the price appreciation of stocks and the receipt of cash dividends. A *total return* index takes both capital gains and income gains fully into account—such an index requires an appropriate treatment of cash dividends.

The best-known and most-quoted indicator of financial market performance is the Dow Jones Industrial Average produced by Dow Jones and Company. It is for this reason that chapter 4 examines the construction and computation of the world’s most famous financial index. The extent to which the considerations raised in chapter 3 in forming a *stock market gross return index* are implemented in constructing the DJIA is explored. It is shown that the DJIA possesses some unusual properties from the perspective of its use as a return measure. In light of these properties it is asked what the DJIA actually does measure. It is shown that the DJIA is not the return on a ‘market’ portfolio consisting of its thirty component stocks: in fact the DJIA measures the return performance on a very particular (and unusual) investment strategy, a fact that is not well understood by institutional investors. As such then, statements such as “the market is up” (as measured by the DJIA) are both vague and somewhat misleading.

The DJIA is compared with other major stock indexes, in terms of both its composition...
and calculation formula. In particular the calculation of the DJIA is compared with the S & P 500. An examination of the various different indexes shows that each is consistent with a particular type of investment strategy. Chapter 4 concludes with an empirical comparison of the different index number formulae used in computing return indexes. To this end, the data used in computing the DJIA are used. It is shown that the formulae lead to very different measured returns (ignoring transactions costs and taxes), particularly over shorter time periods.

Given the numerous ways of constructing stock market return indexes one is left to determine which to use, and to ascertain which one is ‘best’ in some sense. Chapter 5 formally addresses the choice of appropriate formula for a stock market index, whose primary purpose is to measure the gross return on a given portfolio of stocks. With few exceptions (notably Balk (1988)) the theoretical underpinnings of stock indexes are not well developed. Nowhere in the literature is there any strong justification for using one formula over another, nor is there any clear framework for differentiating among the various indexes in terms of the various properties that they possess. To this end, a number of properties (axioms) for a stock market gross return index are proposed—these are considered to be desirable a priori. The test or axiomatic approach to index numbers as developed by Eichhorn & Voeller (1983), Diewert (1993a), and Balk (1995) is adapted here. The axioms (or properties) are suggested in light of common stock market phenomena such as the general rising/falling of stock prices, stock splitting, the issuance of cash dividends, mergers and acquisitions etc. Many of the axioms are standard and have direct analogues to those in the measurement of price inflation for a fixed basket of goods over time, while others are given new interpretations. Some of the axioms do not have analogues in the consumer price index situation. It is discovered that satisfaction of a subset of the proposed axioms implies that the index number formula assumes a specific functional form—it is this index that is labeled ‘best’ from the point of view of this axiomatic approach to stock indexes.

The thesis concludes with an evaluation of the major types of stock market indexes in use today, when their purpose is to measure the gross return on a portfolio of stocks.
The axiomatic approach developed in chapter 5 provides a unifying framework in which to evaluate the various indexes, in the sense that some indexes satisfy certain desirable properties while others do not. Indexes that fail to satisfy basic requirements are considered poor measures of return. It is shown in chapter 6 that the shortcomings of the DJIA as a return measure arise from its failure of a number of the basic axioms proposed. In particular the S & P 500 dominates the DJIA in terms of satisfying the most number of desirable properties—it is in this sense that it is thought to be a superior index. Notwithstanding the satisfaction/failure of axioms it is stressed that each of the major types of indexes corresponds not only to a different portfolio of stocks, but to a different investment strategy. Thus, when choosing an index for benchmarking purposes an investor must choose one which closely matches his/her investment strategy, a decision that cannot be made by appealing to the satisfaction of axioms alone.
Chapter 2

Outlet Types and the Canadian Consumer Price Index

2.1 Introduction

The Report of the Boskin Commission (Boskin et al. 1996) concludes that the U.S. consumer price index (CPI) produced by the Bureau of Labor Statistics (BLS) overstates the rate of inflation relative to a 'true' cost-of-living index. While several empirical studies of CPI bias exist for the U.S., bias estimates and systematic analyses of their sources are lacking for Canada. This chapter investigates one potential source of bias in the Canadian CPI.

The starting point of the construction of the Canadian CPI is the calculation of indexes for CPI basic classes, i.e. subaggregate indexes at the lowest level of aggregation that measure price changes for a single good or service, or a homogeneous group of goods or services. Calculating such indexes requires the collection of price quotes from a number of retailing outlets for goods/services of similar quality. However, consumer substitution among outlets in favour of those offering lower prices and better value, and having different price movements could cause a bias in these indexes. Furthermore, the judgmental sampling of outlets can result in outlet samples that are unrepresentative, resulting in a further bias in these indexes.
The growth in the popularity of lower-priced (discount) outlets during the 1980’s/1990’s has been phenomenal. Reinsdorf (1993) describes such shifts in retailer patronage for the U.S.; Swain (1994) documents similar trends for Canada. It could be argued that discount outlets offer an inferior quality of service to consumers, and therefore that price differences simply reflect quality differentials. Revealed preference theory would indicate that since such outlets have grown in popularity, consumers do experience non-negligible price differentials (net of quality adjustments). When consumers shift to (quality-adjusted) lower-priced outlets the average price for the good in question falls. To the extent that the good remains priced at the higher-priced outlet, or to the extent that no direct comparison is made between prices at the new and the old outlet, the CPI component misses the price decline and overstates the rate of inflation. Biases may also arise from the fact that different outlet types may have different price movements in addition to level differences.

This chapter examines the extent of these outlet effects problems in the Canadian CPI and is organized as follows. Section 2.2 explains the concepts of outlet substitution bias or outlet effects, and describes two phenomena leading to potential overstatements of CPI basic classes relative to a ‘true’ cost-of-living index. Section 2.3 explains the different outlet types examined in this paper and the extent to which the Statistics Canada outlet sample inadequately represents these different outlet types. The remaining sections report on empirical evidence on index calculations. Section 2.4 reports on differential price movements and levels across outlet types. Section 2.5 finds empirical evidence of the two types of bias discussed in section 2.2 for a number of CPI basic classes, indicating that official CPI subaggregate indexes grow at faster rates than more appropriate alternatives. Section 2.6 offers some concluding remarks.

2.2 Outlet effects in the CPI

Outlet substitution biases or outlet effects are defined here as mismeasurements that arise in the CPI for failing to adequately account for shifts in retailer patronage to outlets with
different price movements and/or levels. Once an outlet sample has been chosen a statistical
agency can aggregate price and quantity information over many outlets by calculating a unit
value for each commodity (over all outlets). Alternatively outlet-specific unit values can be
aggregated using an index number formula such as the Laspeyres. These are considered in
turn.

Method one: unit values across outlets

The unit value price $u_i^t$ of an item $i$ at time $t$ sold at $N$ outlets for prices $p_j^t$ and quantities
$q_j^t, j = 1, \ldots, N$ is $u_i^t = \frac{\sum_{j=1}^{N} p_j^t q_j^t}{\sum_{j=1}^{N} q_j^t}$ (see Hill (1993), Diewert (1995) and Balk (1998) for
more on unit values). Unit value prices (across outlets) give a more accurate summary of
the average transaction price than an isolated price quotation from each outlet. They are
appropriate when aggregating across goods which are homogeneous. Purchases of such goods
are primarily influenced by price—as such these goods are not distinguished by their point-
of-purchase. Independence from place-of-purchase is thus consistent with the requirement of
homogeneity. Balk (1998) notes that unit value indexes (which are ratios of unit value prices)
suffer from two drawbacks when used as price indexes: they are sensitive to the units of
measurement (commensurability property) and equality of prices in both periods does not
guarantee an index of unity (identity property). In the present context, the requirement of
homogeneity means that price quotes are being collected for the same unit of item so that
the failure of commensurability poses no problems here.

To illustrate the nature of the outlet substitution bias that can arise if the CPI ignores
shifts in retailer patronage to lower-priced outlets, consider the following example. There are
two outlet types: high-priced ($H$) and low-priced ($L$), both selling a similar good or service.
$L$ gains a quantity share from $H$ between the base and comparison periods. If a unit value
index is adopted as the appropriate concept for a price index, then the ‘true’ (average) price
index over both outlet types would be given by

$$P_u = \frac{u^1}{u^0} = \frac{(1 - s^1)(1 + i) + s^1(1 + i)(1 - d)}{(1 - s^0) + s^0(1 - d)}$$

(2.1)

where $p_H = 1 + i$ is the price index for $H$, $s^t, t = 0, 1$ is the quantity share of $L$ in period $t$,
and $d$ is the discount offered by $L$ over $H$. If the CPI ignores the new outlet then the index is given by $P_H$. Thus an approximate outlet substitution bias $B_t$ (i.e. caused by consumers substituting lower-priced outlets for higher-priced outlets) is given by

$$B_t \equiv P_H - P_u = \frac{md(1 + i)}{(1 - s^0) + s^0(1 - d)} < s^1d(1 + i) \text{ for } s^0 > 0,$$

(2.2)

where $m = s^1 - s^0$ is the gain in quantity share of $L$ over $H$ between the base and comparison periods. This bias exists until such time as the lower-priced retailers’ quantity share has stabilized. The bias in equation 2.2 assumes that quality differentials across the two outlets are negligible and that the discount $d$ is constant between the base and comparison periods (so that both $H$ and $L$ have the same price movement). The market share gain is a good approximation of $m$ for ‘small’ and constant $d$ between the base and comparison periods. The bias in equation 2.2 is referred to as a levels outlet substitution bias and it is an upper bound on the true bias.

Statistics Canada does not ignore the emergence of new outlet types, but its procedures for introducing these outlets into the sample may still cause the CPI to miss the decline in average prices. The technique used by Statistics Canada for introducing a new outlet into the sample is known as splicing. Suppose that there are only two outlets (old and new), and that the new outlet ($L$) replaces the old outlet ($H$) in the sample at time 1. Let $p_{tj}^0, q_{tj}^0$ denote the price and quantity sold (respectively) of the good at outlet $j = L, H$ at time $t = 0, 1, \ldots$, where price and quantity information for both the new and old outlets are collected at $t = 1$. Changes in the index are proportional to price changes in the old outlet prior to and including time 1, after which they are proportional to changes in the new outlet. This procedure prevents any change in the composition of the outlet sample from directly affecting the index value by implicitly assuming that price level differences between

$^1$Equation 2.1 generalizes a formula due to Diewert (1998)[page 50] who assumed that the low cost outlet did not exist in the base period.

$^2$Note that if the base period is prior to the existence of the discount outlet then $s^0 = 0$ so that $P_u = (1 - s^1)(1 + i) + s^1(1 + i)(1 - d)$ and $B_t \equiv s^1d(1 + i)$ which is the substitution bias in Diewert (1998)[page 51].
the outlets are entirely attributable to quality differences. If \( \frac{p_H}{p_L} > 1 \) (the good in the old outlet is more expensive than the good in the new outlet) at the time of introduction, the official CPI components miss the decline in the price of the good. If, however, the price differences are due to quality differentials, then splicing is the correct procedure to use.

If the unit value price at \( t = 1 \), \( u^1 \) is used instead of either period 1 price at the old or new outlet then the chained sequence will reflect this fall in average price. If \( p_L^1 = p_H^1(1 - d) \) then the divergence between the standard price index at time 1 and that based on unit values is given by \( mdP_H^1 \), where \( P_H^1 \) is the period 1 index for the old outlet, and \( m \) is the quantity share of the new outlet in period 1.\(^3\) The extent of an upward bias here depends crucially on both the strategic price reaction of incumbents and on the quality differentials across the different outlets.

**Method two: outlet-specific unit values**

In the United States 20% of the outlet sample is rotated each year—such a rotation keeps the sample current in a rapidly changing environment and helps reflect consumer expenditure patterns. Just as consumers are free to alter their pattern of consumption among goods as relative prices change, so too are they free to shift their retailer patronage in response to differential mixes of prices/services offered by new entrants and existing incumbents. Outlets are chosen on the basis of the Continuing Point of Purchase Survey (CPOPS) using a probabilistic rule, where the probability of an outlet being selected is proportional to consumer expenditures. At Statistics Canada outlets included in the sample are chosen judgmentally. Infrequent outlet sample reviews may result in samples which are increasingly unrepresentative via-à-vis available market share information. This may result in a second type of bias.

This second approach to estimating bias involves aggregating outlet-specific unit values and quantities using an index number formula such as the Laspeyres.\(^4\) Here a bias may arise

\(^3\)Note that, unlike in equation 2.2, \( m \) is not the quantity share gain in the sense that price and quantity information for the new outlet were not collected in the base period. However \( m \) is the quantity share gain of the new outlet if it did not exist in the base period.

\(^4\)Typically outlet-specific unit values are single price quotes assumed to be representative of the average
from the fact that different outlet types may have different price movements (as opposed to level differences), and if the outlet sample is unrepresentative. Suppose, as before that there are two outlet type indexes $P_H$ and $P_L$ (assume for example that $P_H > P_L$). If $L$ has a base period market share of $m\%$ then the true price index which gives proportionate weighting to the outlet type indexes is $P_A$, given by

$$P_A = (1 - m)P_H + mP_L.$$  \hspace{2cm} (2.3)

If the last outlet sample review occurred prior to the emerging dominance of the discounters then the official CPI is $P_H$ yielding an approximate outlet effect of $B_m = P_H - P_A = m(P_H - P_L) > 0$. This is called a movements outlet substitution bias.$^5$ Note that this type of bias only picks up that part of the outlet substitution bias due to the discount outlets lowering their prices faster rather than reflecting level differences across outlets. Consequently this second approach allows for quality differences across outlets since price levels across the different outlets are not compared directly. If $P_H = P_L$ then exclusion of type $L$ from the outlet sample does not bias the index. This is similar in spirit to the new goods bias (see Baldwin, Després, Nakamura & Nakamura (1997))—to the extent that the pricing cycle of excluded new goods is the same as old included goods, exclusion of new goods will not bias the index.

2.2.1 Empirical evidence to date

Reinsdorf (1993) examines the extent of the outlet substitution bias for the U.S. CPI. Comparing the prices of incoming and outgoing samples (since the BLS rotates its outlet sample), he obtains an upward bias of 0.25% for the food-at-home and gasoline components of the U.S. CPI. For food, the linked CPI indexes rise 2% per year faster than the corresponding transactions price at an outlet at a particular time.

$^5$A bias can result if the two different outlet types have different price movements or if the market share weights do not reflect actual consumer expenditure patterns. Thus this bias may also be called a sample selection bias or an unrepresentative outlet sample bias.
average price (AP) series produced by the BLS, while for fuel it is a more modest 0.9%. These results should be treated as upper bounds on this type of bias since no adjustment was made for potential changes in quality. Although Reinsdorf discusses "new outlet" bias he later notes that "formula bias" is the principal source of CPI-AP discrepancies (Reinsdorf 1998)[pages 181-183].

The only estimates of outlet substitution bias for Canada are Crawford (1993), based on a number of conjectures about the key parameters in equation 2.2 (although he doesn’t explicitly use it). He estimates that 40% of the CPI basket is prone to potential shifts to lower-priced outlets, that \( m = 2\% \) per annum and \( d = 10\% \), yielding a maximum bias of 0.08% per year. Investigations into outlet substitution biases using unit value indexes calculated from scanner data include Saglio (1994) (0.8% “point-of-purchase” effect for milk chocolate bars in France), Dalén (1997) (1% upward bias for fats, detergents, breakfast cereals and frozen fish in Sweden) and Bradley, Cook, Leaver & Moulton (1997) (3% upward bias for coffee in the U.S.).

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Examples</th>
<th>SIC coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>&quot;Major&quot; Department Stores</td>
<td>Eaton's, The Bay</td>
<td>6411(Department Stores)</td>
</tr>
<tr>
<td>T2</td>
<td>&quot;Discount&quot; or &quot;Junior&quot; Department Stores and General Merchandisers</td>
<td>Zeller's, Wal-Mart</td>
<td>6411 (Department Stores), 6412 (General Stores), 6413 (Other general merchandise stores)</td>
</tr>
<tr>
<td>T3</td>
<td>Retail Chain (&quot;high-end&quot;)</td>
<td>Holt Renfrew, Sony Store</td>
<td>All potential SICs excluding 6411, 6412, 6413</td>
</tr>
<tr>
<td>T4</td>
<td>&quot;Discount&quot; Retail Chain</td>
<td>Future Shop, Pharma Plus</td>
<td>All potential SICs excluding 6411, 6412, 6413</td>
</tr>
<tr>
<td>T5</td>
<td>Independents (typically local or specialty)</td>
<td>Wackid Radio, McLeod Pharmacy</td>
<td>All potential SICs excluding 6411, 6412, 6413</td>
</tr>
</tbody>
</table>

Table 2.1: Outlet-Type Classifications.
2.3 Outlet types in the CPI

A summary of the outlet type classifications that are used here to investigate outlet substitution bias in the Canadian CPI is shown in table 2.1. The classifications, which are based on Statistics Canada retail trade surveys such as the Annual Retail and Wholesale Trade Survey (RTS), the Annual Retail Chain and Department Store Survey (ARC), trade publications such as Discount Stores News (U.S.) and Retail Chain Guide (U.S.), and the author's own perceptions, are chosen to be exhaustive and mutually exclusive. Categories T1 and T2 are based on definitions in Statistics Canada (1997)[page 79], and their associated Standard Industrial Classification (SIC) codes are shown in table 2.1. Total retail chains (i.e. both T3 and T4) are defined in Statistics Canada (1997)[pages 79-80]. The delineation of retail chains into 'high-end' and 'discount' is based primarily on general market perceptions. For example, specialized stores with a reputation for quality and service fall into category T3. Category T5 covers a relatively heterogeneous group of outlets. Independents include what might commonly be referred to as 'mom-and-pop' style stores, local specialty stores and firms not meeting the definition of a retail chain.

The present analysis focuses attention on a few specific CPI basic classes. These were chosen so as to include all of the following scenarios: 1) a range of outlet types is priced and price differentials are thought to exist; 2) the range of outlets priced does not give proportionate (i.e. to market share) representation to 'discounters'; 3) the items are those that are available at many outlet types and many SIC codes. The final choice of basic classes is shown in table 2.2, and outlet effects are estimated for a subset of these. These basic classes are not representative of the entire CPI universe of goods, in part because discounting and shifts to lower-priced outlets have been more pronounced for these commodities. Nonetheless the phenomenon described here may also be indicative of retailing trends for such items as hardware, some furniture and sporting items, books, office supplies and general food purchases. These items constitute an approximate combined 30% weight in the CPI.

Appendix A.1 documents examples of the distribution of the outlet sample by outlet type, based on the actual outlet samples used by Statistics Canada in its data collection for
<table>
<thead>
<tr>
<th>Basic classes</th>
<th>% CPI weight&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other household equipment&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.37</td>
</tr>
<tr>
<td>Men's pants</td>
<td>0.36</td>
</tr>
<tr>
<td>Non-prescribed medicine</td>
<td>0.22</td>
</tr>
<tr>
<td>Photographic services and supplies</td>
<td>0.32</td>
</tr>
<tr>
<td>Audio equipment</td>
<td>0.27</td>
</tr>
</tbody>
</table>


<sup>a</sup>1992 basket at 1992 prices  
<sup>b</sup>These include kettles, food mixers etc.

Table 2.2: Commodities for examination.

Ontario. For example, for the basic class photographic services and supplies, the percentage distribution has remained relatively stable, while for others (such as audio equipment) there has been a very slow shift to discounters over the seven year period. In general the samples are unrepresentative vis-à-vis available market share information.

In order to give a more complete picture of the evolution of the outlet sample, estimates of the rate of sample turnover (i.e. the rate at which new outlets are being brought into the sample and old outlets are being dropped) are calculated. Table 2.3 lists some annual turnover rates (ATRs) which attempt to measure inertia in the outlet sample—their calculation is described in appendix A.2. In general, the ATRs are low and contrast with the 20% annual outlet sample rotation in the U.S. CPI.

Obtaining accurate and reliable market share information at the detailed commodity level is difficult. One of the main obstacles to constructing market share information is the fact that Statistics Canada (and other) trade surveys are industry surveys rather than commodity surveys. Nonetheless, some approximate market share information for certain commodity groupings is available, and these are tabulated in appendix A.3. The sources and construction of the data are also described. The data indicate that discount department stores have experienced gains in market share, particularly for the basic classes audio equipment and
Table 2.3: Sample turnover rates.

<table>
<thead>
<tr>
<th>Basic classes</th>
<th>Annual Turnover Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other household equipment</td>
<td>5</td>
</tr>
<tr>
<td>Non-prescribed medicine</td>
<td>5</td>
</tr>
<tr>
<td>Men’s pants*</td>
<td>4</td>
</tr>
<tr>
<td>Photographic services and supplies</td>
<td>5</td>
</tr>
<tr>
<td>Audio equipment</td>
<td>8</td>
</tr>
</tbody>
</table>

*Men’s jeans were taken as representative of this basic class.

other household equipment.

An examination of the tables in appendix A.1 and appendix A.3 indicates a number of pricing gaps or missing information items in the distribution and coverage of price quotes by outlet type for the period. The term pricing gaps, as used here, has two possible meanings: 1) an outlet type has no (or few) price quotes for a given basic class, even though it is felt that the items are sold at such outlet types (even in the absence of concrete market share information), and 2) there is little correspondence between an outlet type’s share of price quotes and available market share information. As an illustration of the second type of pricing gap: over the 1990-96 period discount department stores have captured (on average) 50% of total (i.e. \( T_1 \) and \( T_2 \)) department store sales for audio equipment and yet have received (on average) less than 20% of the department store allocation of price quotes for same (tables A.3 (appendix A.1) and A.5 (appendix A.3)). The consequences of such unrepresentative sampling will be examined in section 2.5.

2.4 Outlet type as a price determinant

Kokoski (1993) and Turvey (1997) report evidence that the coefficient estimates of outlet type in a hedonic price regression are statistically significant, so that outlet type is a relevant price determining characteristic for certain items. Even then, it is difficult to attribute the
outlet type coefficients to pure price movements or to something else like quality differences across outlets (such as sales service and range of products offered). Such a detailed analysis is beyond the scope of this analysis since information on outlet characteristics was not available. In light of the possibility that discounters offer goods/services of lower quality, calculated bias estimates presented here for the levels outlet substitution bias must be treated as upper bounds. Notwithstanding this, revealed preference theory would indicate that since such outlets have grown in popularity consumers do experience non-negligible price differentials (net of quality adjustments). Estimated biases would be lower if such discounters offer significantly lower quality in the services and the product lines to their customers. The movements outlet substitution bias does allow for quality differences across outlets.

In order to examine whether outlet types display differential price movements some outlet specific indexes were constructed for certain commodities. The price data used were taken from the internal monthly production files used by Statistics Canada in calculating the CPI. All prices had been adjusted (by Statistics Canada) for appropriate sales taxes, and in some cases the prices had had (judgmental) quality adjustments made to them when, for example, the unit of measurement or item specification had changed. Each price quote was identified by its point of purchase and these were classified into types $T1 - T5$, facilitating the calculation of indexes by outlet type. Two examples of outlet type indexes are shown here. These indexes are item specific indexes and the month-to-month price movements for item $i$ between months $t - 1$ and $t$, $P_{t-1,t}^i$, $t = 2, \ldots$ are calculated as geometric means of price relatives (Jevons index $P_{JE}$), based on matched samples or

$$P_{t-1,t}^i = P_{JE}(p_{t-1}, p^t) = \Pi_{j=1}^N \left( \frac{p_{j}^t}{p_{j}^{t-1}} \right)^{\frac{1}{N}},\quad (2.4)$$

where $N$ is the matched number of outlets in the sample in months $t - 1$ and $t$. The chained sequences for each item $i$ (for 84 months, covering the seven year period 1990-96) are given by

$$P_{i}^{84,1} = P_{i}^{84,83} \times P_{i}^{83,82} \times \ldots \times P_{i}^{3,2} \times P_{i}^{2,1},\quad (2.5)$$

18
Notes: The source of the price data is the Statistics Canada monthly CPI production files, 1990-1996. Pricing at types T3 and T5 did not begin until 1995, and are not shown here.

Figure 2.1: Index for men’s jeans by outlet type, 1990-1996 (Ontario)

where each $P_{i-1,t}^c$, $t = 2, \ldots, 84$ for commodity $i$ is calculated as in equation 2.4, and the indexes are rebased to January 1990 = 100. In general, items have different price movements across outlet types.

Figure 2.1 shows price indexes for men’s jeans. The brands may not be identical across the different outlets—a characteristic which consumers view as important. Nonetheless, price movements for ‘jeans’ for given outlet types do differ. For men’s jeans, prices at T1 have been rising annually at an average rate of 4.4%, while those at T2 have been falling at an annual rate of 3.3% (see table 2.4). The implication for the CPI basic class men’s pants is that if consumers favour shopping at T2 over T1 then a discount department store index better reflects the pattern of price movements that consumers experience. If this is the case then the official CPI subaggregate men’s pants overstates the true rate of price inflation for men’s pants. Figure 2.2 shows the price index for non-prescribed medicines. Prices at T1 rose annually at an average rate of 2.6%, while those at T2 showed no average change over
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>100.0</td>
<td>103.6</td>
<td>104.0</td>
<td>116.5</td>
<td>126.2</td>
<td>126.0</td>
<td>129.6</td>
</tr>
<tr>
<td>T2</td>
<td>100.0</td>
<td>98.6</td>
<td>102.9</td>
<td>109.5</td>
<td>100.8</td>
<td>90.7</td>
<td>81.7</td>
</tr>
<tr>
<td>T3</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>100.0</td>
<td>106.9</td>
</tr>
<tr>
<td>T4</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>96.0</td>
<td>—</td>
<td>98.9</td>
<td>98.8</td>
</tr>
<tr>
<td>T5</td>
<td>100.0</td>
<td>94.5</td>
<td>—</td>
<td>101.6</td>
<td>—</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>average growth rate</td>
<td>4.4%</td>
<td>-3.3%</td>
<td>6.9%</td>
<td>-0.2%</td>
<td>0.0%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Pricing at this type began in 1995.

Table 2.4: Index for men's jeans by outlet type, 1990-1996 (Ontario).

the same period (see table 2.5). Similar trends were observed for other items such as kettles, food mixers and cassette recorders.

While *price movements* are the relevant concept/consideration for the CPI, a comparison across outlet types indicates that absolute levels of prices *do* differ for the item specifications priced for the CPI. Table 2.6 tabulates approximate discounts for four basic classes for the seven year period. The discounts are computed by calculating arithmetic average prices over the different outlet types. The price levels are ranked from most expensive to least. The most expensive outlet has a 0% discount, and all other percentage discounts are expressed relative to the most expensive outlet for that given basic class. Average price dispersion among the different outlet types can vary from year to year: those shown in table 2.6 show the average dispersion for the period 1990-1996. While these prices have not been adjusted for quality differences, the results are striking—discount department stores offer considerably lower prices than independent stores, and are consistently the cheapest outlet type at which to shop.
Notes: The source of the price data is the Statistics Canada monthly CPI production files, 1990-1996.

Figure 2.2: Index for non-prescribed medicine by outlet type, 1990-1996 (Ontario)

<table>
<thead>
<tr>
<th>year</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>100.0</td>
<td>100.0</td>
<td>—</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>1991</td>
<td>100.0</td>
<td>98.0</td>
<td>—</td>
<td>97.4</td>
<td>100.8</td>
</tr>
<tr>
<td>1992</td>
<td>86.9</td>
<td>100.4</td>
<td>—</td>
<td>97.7</td>
<td>100.8</td>
</tr>
<tr>
<td>1993</td>
<td>86.1</td>
<td>103.6</td>
<td>—</td>
<td>99.4</td>
<td>101.6</td>
</tr>
<tr>
<td>1994</td>
<td>101.6</td>
<td>101.5</td>
<td>—</td>
<td>103.1</td>
<td>103.1</td>
</tr>
<tr>
<td>1995</td>
<td>113.5</td>
<td>101.3</td>
<td>—</td>
<td>102.2</td>
<td>103.2</td>
</tr>
<tr>
<td>1996</td>
<td>—</td>
<td>100.3</td>
<td>—</td>
<td>102.9</td>
<td>101.9</td>
</tr>
</tbody>
</table>

average growth rate 2.6% 0.0% — 0.5% 0.3%

*Major department stores were dropped from the sample in 1995.

Table 2.5: Index for non-prescribed medicines by outlet type, 1990-1996 (Ontario).
**2.5 Bias estimates**

Recall that savings from switching to lower-priced outlets are not reflected in current CPI procedures. Method one in section 2.2 above explained the calculation of a levels outlet substitution bias based on unit value indexes. Measuring such a bias is a tricky empirical issue, since the magnitude of the bias depends crucially on the strategic price response of incumbents and the quality differential between the new and old outlets. If the incumbents match the lower prices of the discounters, then the CPI procedures would reflect those price changes. If however, the lower prices reflect a lower level of retailer services, then direct comparison of price levels would overstate the gain in consumer welfare from the emergence of such outlets.

An approximate calculation of the effect of discount outlets can be estimated from the calculations on average prices and market shares based on equation 2.2 in section 2.2. Here the focus is on the bias resulting from T2, i.e. the \( d \) in equation 2.2 refers to the percentage discount offered by T2 over all other outlet types, and \( m \) is the annual gain in market share\(^6\) of T2 over the 1990-1996 period.\(^7\) For example, the average annual gain in market share of type T2 for the audio equipment basic class was 2%. Prices at discounters were

---

\( ^6 \)Recall again that market shares approximate the quantity shares in equation 2.2 for ‘small’ \( d \), where \( d \) is constant in the base and comparison periods.

\( ^7 \)Tests for the statistical significance of differences in average prices across the outlet types were carried out. These are shown in appendix A.4.
Table 2.7: Levels outlet substitution biases based on unit values.

approximately 24% lower than prices at all other outlet types. These imply a maximum bias for audio equipment of about 0.5% per year (see table 2.7). Biases for other basic classes are shown in table 2.7. If it is assumed that 40% of the CPI basket is prone to such effects (as Crawford (1993) does) and using the estimates of \( d = 20\% \) and \( m = 2\% \) (for all basic classes on average) would imply a maximum bias of 0.16% per annum, twice as large as Crawford who assumed \( d = 10\% \). In fact the likely range of bias estimated in table 2.7 (assuming that 30% of the CPI basket is subject to outlet substitution) is 0.1 - 0.15% per annum. The bias estimates for the individual basic classes shown in table 2.7 compare to the 0.25% outlet bias of the total 1.4% per annum discrepancy between the CPI and AP series for food discussed by Reinsdorf (1998). The absence of chaining in the AP series, which would otherwise prevent the outlet sample composition from affecting the series, means it is less susceptible to outlet bias. Reinsdorf’s bias lies at the lower end of the range in table 2.7, and is based on comparing the price levels at old and new outlets in sample rotation. This bias is due primarily to splicing. Such a problem is likely to be smaller in magnitude for the Canadian CPI due to its infrequent outlet sample review.

Note that the estimates in table 2.7 are based on a unit value concept and it is assumed that there is no quality differential between discounters and other outlets—thus the bias should be treated as an upper bound on the levels outlet substitution bias. Levels outlet

<table>
<thead>
<tr>
<th>Basic class</th>
<th>Outlet sub. bias (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>other household equipment</td>
<td>0.34</td>
</tr>
<tr>
<td>non-prescribed medicines</td>
<td>0.16</td>
</tr>
<tr>
<td>photographic services and supplies</td>
<td>0.31</td>
</tr>
<tr>
<td>audio equipment</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Note: these estimates are based on equation 2.2 in section 2.2. The calculations assume an average level of \( s^0 = 0.2 \) and an average inflation rate of \( i = 0 \), both consistent with the empirical evidence. The estimate for photographic services and supplies is based on an assumed annual market share gain by discount department stores of 1%.
substitution bias estimates based on the unit value approach implicitly assume that items should not be distinguished by their point of purchase, and given that consumers quite often base their purchases almost entirely on the price level, this is not an unreasonable assumption. Although the average prices are not adjusted for quality differentials the percentage differences between T2 and T5 are striking.

Unrepresentative sampling combined with a lower rate of increase in prices at discount stores have resulted in a movements outlet substitution bias. Recall from section 2.2 that the second method for estimating an outlet effect involved aggregating over outlet types indexes, giving explicit consideration to the differential price movements across the outlet types.

Basic classes are composed of representative commodities that are expected to be reasonably homogeneous from the point of view of price change. These *micro-indexes* or *elementary indexes* are month-to-month ratios of geometric means of price relatives, and longer term price changes are calculated by linking the price indexes as in equation 2.5. This is the current practice at Statistics Canada and was introduced with the last CPI update in January 1995, before which time the ratio of arithmetic mean prices index (Dutot index) was used. Since basic classes typically consist of a number of representative commodities, the price index (at time $t$) for a basic class (on time base 1990) $P_{t/90}^{bc}$ is a Laspeyres index given by

$$P_{t/90}^{bc} = \frac{\sum_i P_i^{t/90} \times w_i}{\sum_i w_i},$$

where the summation is over the representative commodities and the $w_i = p_{90} q_{90}$ are the estimated dollar weights obtained from the 1990 *Family Expenditure Survey*, where 1990 is both the basket reference period and the base period. The formula in equation 2.6 is used in a recursive way, meaning that the index for any given subaggregate can be expressed, in turn, as a weighted arithmetic average of indexes for the components of this subaggregate. As an example, a provincial index is a weighted arithmetic average of the subprovincial (strata) indexes.\(^8\)

\(^8\)Because of the way in which it is chained when its basket is updated, unlike a fixed-basket index, the CPI (at higher levels of aggregation) cannot be interpreted as a weighted arithmetic average of its constituent subaggregates, since its longstanding non-revision policy has resulted in the CPI being chained forward.
Notes: The source of the price data is the Statistics Canada CPI monthly production files, 1990-1996.

Figure 2.3: Adjusted and official index for audio equipment, 1990-1996 (Ontario).

Table 2.8: Adjusted and official index for audio equipment, 1990-1996 (Ontario).
The above procedure was modified so that each basic class index is calculated for each type, i.e. computed as if a particular outlet type were the only type from which quotes were collected (quotes from other outlet types were dropped from the index calculation). An additional tier of aggregation was added above the strata level—a basic class index for a province is a weighted arithmetic average of the outlet type indexes, where the weights reflect estimated market shares of the different outlet types, as in equation 2.3 in section 2.2.

The basic class audio equipment is currently based on pricing for cassette recorders and stereo systems. Two serious gaps in pricing coverage occurred for audio equipment in Ontario for the 1990-1996 period—discount department stores have captured 50% of department store sales for audio equipment and yet have received less than 20% of the department store allocation of price quotes. Furthermore stereo systems were not priced at T2. Two adjustments were made to the index to take account of these facts: 1) proxy pricing was used for stereo systems by imputing it from the price movements of cassette recorders (for T2); 2) the outlet type indexes were aggregated using market shares as weights. Figure 2.3 shows the adjusted index compared with the corresponding official one. From the index levels shown in table 2.8 it is seen that if these adjustments had been made, the index for audio equipment would have fallen on average by 1.8 percent per year between 1990 and 1996 rather than by 1.4 percent as the official series does—a movements outlet substitution bias of +0.4% per year. This result is attributed primarily to the faster rates of price decline at T2 and their gain in market share over the period.

The basic class for non-prescribed medicines is based on pricing for such items as anti-septic mouthwash and cold/cough remedies. For the 1990-1996 period the most apparent problem for non-prescribed medicines was the lack of coverage at supermarkets, whose approximate market share was 30%. The price movements at supermarkets were imputed from those of discount pharmacy chains, whose price movements they are thought to follow most closely. When the outlet type indexes are aggregated the index for non-prescribed medicines rather than backward. However, any composite index in the new CPI series (or unlinked series) is calculated as a weighted arithmetic average of the price indexes of its subaggregates.
Notes: The source of the price data is the Statistics Canada CPI monthly production files, 1990-1996.

Figure 2.4: Adjusted and official index for non-prescribed medicines, 1990-1996 (Ontario).

<table>
<thead>
<tr>
<th>year</th>
<th>Official index 1990=100</th>
<th>Adjusted index 1990=100</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>1991</td>
<td>105.6</td>
<td>99.6</td>
</tr>
<tr>
<td>1992</td>
<td>105.7</td>
<td>98.0</td>
</tr>
<tr>
<td>1993</td>
<td>107.8</td>
<td>99.0</td>
</tr>
<tr>
<td>1994</td>
<td>108.5</td>
<td>101.9</td>
</tr>
<tr>
<td>1995</td>
<td>107.6</td>
<td>104.0</td>
</tr>
<tr>
<td>1996</td>
<td>107.8</td>
<td>103.3</td>
</tr>
<tr>
<td></td>
<td>average growth rate</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.3%</td>
<td>0.5%</td>
</tr>
</tbody>
</table>

Table 2.9: Adjusted and official index for non-prescribed medicines, 1990-1996 (Ontario).
Notes: The source of the price data is the Statistics Canada CPI monthly production files, 1990-1996.

Figure 2.5: Adjusted and official index for other household equipment, 1990-1996 (Ontario).

would have risen on average by 0.5 percent per year between 1990-1996 rather than by 1.3 percent as the official series does (see figure 2.4 and table 2.9).

The basic class for other household equipment is based on pricing for items such as food mixers and kettles. One feature of this sample is the disproportionately high allocation of price quotes to T1, while T2 command approximately 50% of the market. When the outlet type indexes are aggregated the index for other household equipment would have fallen on average by 1.1% per annum rather than by 0.3% as the official series does. These indexes are shown in figure 2.5 and table 2.10. The higher rate of price decline for the adjusted index is caused by the higher rates of price decline for types T2 and T4 whose combined market share is approximately 70%.

Note that the large drops in prices over the period 1991-92 (figures 2.3, 2.4, and 2.5) correspond to Wal-Mart's (formerly Woolco) entrance into the Canadian retail market.
<table>
<thead>
<tr>
<th>year</th>
<th>Official index</th>
<th>Adjusted index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1990=100</td>
<td>1990=100</td>
</tr>
<tr>
<td>1990</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>1991</td>
<td>100.0</td>
<td>96.1</td>
</tr>
<tr>
<td>1992</td>
<td>98.2</td>
<td>93.7</td>
</tr>
<tr>
<td>1993</td>
<td>98.0</td>
<td>92.7</td>
</tr>
<tr>
<td>1994</td>
<td>98.4</td>
<td>93.8</td>
</tr>
<tr>
<td>1995</td>
<td>98.2</td>
<td>93.6</td>
</tr>
<tr>
<td>1996</td>
<td>97.9</td>
<td>93.7</td>
</tr>
</tbody>
</table>

average growth rate   -0.3%   -1.1%

Table 2.10: Adjusted and official index for *other household equipment*, 1990-1996 (Ontario).

### 2.6 Concluding Remarks

Between 1990 and 1996 certain *basic class* indexes averaged higher rates of inflation than alternate indexes based on either a unit value approach or on one obtained by aggregating over outlet type price indexes. The entry of lower-priced (discount) outlets, the fact that discount outlets appear to deliver lower rates of price increases over time, and unrepresentative sampling have contributed to the deviations uncovered in this paper.

Aggregating up to the All-Items CPI the biases are smaller since the effects discussed here are not representative of the entire universe of goods/services covered by the CPI. If 30% of the CPI basket is prone to such outlet effects, then a plausible range for the levels outlet substitution bias for the Canadian CPI is $0.1 - 0.15\%$ per annum, assuming that quality differentials are negligible. Biases arising from the movements outlet substitution bias and unrepresentative sampling lie in the range $0.12 - 0.24\%$ per annum for the All-Items CPI. These biases are approximately additive, resulting in an overall outlet substitution bias and unrepresentative outlet sample bias for the Canadian CPI in the range $0.2 - 0.4\%$.

A number of recommendations/observations arise from the work. Given the apparent importance of the outlet characterization, *better* and *more* use of available market share information should be used in the outlet selection process. A negative consequence of failing
to do this is an outlet sample that is unrepresentative by type, resulting in large upward biases. The Statistics Canada outlet sample needs to be updated on a more timely basis. In some circumstances, unit value indexes are appropriate and can capture declines in average prices when a new lower-priced outlet is included in the outlet sample, and can help reduce biases from splicing. While the results are striking, the biases are based on limited evidence for Ontario, which may be indicative of trends in the rest of Canada. Additional research and more data are needed to corroborate the findings here. Further work on hedonics needs to be undertaken to make appropriate quality adjustments to prices facilitating direct comparisons across outlets.

An implication of these findings is that the CPI may have tended to overstate changes in the cost of living over the past few years. Consequently growth rates in productivity and real wages may in fact be a lot healthier than official statistics suggest. Trends in retail trade figures indicate that the impressive and sustained growth of warehouse/discount outlets may be a very real and pervasive source of mismeasurement in the Canadian CPI.
Chapter 3

Stock Market Indexes: Basic Concepts

Several indexes of stock market performance exist. Perhaps the best known indexes are the family of Dow Jones Averages produced by Dow Jones and Company. Other stock market indexes include those produced by Standard and Poor's (e.g. Standard and Poor's 500 Composite Index (S & P 500)), the American Stock Exchange (AMEX indexes), Wilshire Associates (e.g. Wilshire 500 Equity Index), Value Line Inc. (Value Line Averages), Frank Russell and Company (Russell Equity Indexes), and the Toronto Stock Exchange (e.g. Toronto Stock Exchange 300 Composite Index (TSE 300)). Further examples are listed in Ross (1992)[pages 583-586], Berlin (1990) and Schwert (1990). While each index shares the common characteristic of aggregating price (and sometimes quantity) information pertaining to common stocks, the manner in which this is done is not uniform across indexes. Each index differs in composition, aggregation formula, and methodological constructions—with corresponding implications for interpretation and usage.

Stock indexes have gained widespread acceptance and are monitored by economists, bankers, investors and speculators. Indeed, there is a tendency to accept such indexes on faith. As early as the beginning of this century economists were questioning the construction of stock indexes. Mitchell (1916)[page 625] notes
the paucity of explanations and warnings has encouraged readers to use or mis-use the results without undergoing the mental toil of criticism or the moral strain of doubt. As for the cautious minority, they have been foiled by this simplicity of presentation; they have been given few materials wherewith to determine the representative value of the original quotations, to judge the appropriateness of the methods used, or to compare the results of rival series.

This concern motivates the need to test the value of these popular indexes. One approach is to put them side by side and evaluate them against the potential concepts of stock indexes and examine the differences between them.

This chapter explores some of the basic uses and features of stock market indexes. Section 3.1 attempts to provide an answer to the question: "what is a stock market index?" As noted in chapter 1, this is tantamount to asking for what purpose the index is designed, since, in general, its intended purpose will dictate and define the design and meaning of the stock market index. Different types of indexes are required depending on whether the index is supposed to represent quantity, price, value, or some other aspect of the stock market. Section 3.2 introduces some preliminary ideas and concepts relevant to forming stock indexes. Section 3.3 discusses three basic considerations that need to be addressed when constructing a stock market gross return index, whose primary purpose is to measure the gross rate of return on a portfolio. A more formal discussion of the construction and properties of stock market gross return indexes is undertaken in chapter 5.

3.1 What is a stock market index?

"How is the market doing?" The answer to this frequently posed question will vary depending on what information is used, and what index is used to answer this question. More specifically though, the answer should depend on what specific question is being asked. Answers of the form "the market is up" are neither informative nor particularly precise. The stock

1Possible interpretations of this statement might relate to the volume of shares traded, the price appreciation of stocks, the market capitalization of firms, or the return on a particular market portfolio of
market, in the broadest sense of the word is comprised of many thousands of firms who have made their stock available to the public. Incorporating all stocks (and associated information) into one market indicator would be an arduous task and would not necessarily provide a satisfactory answer to any aspect of the question posed above. Consequently, subsets of numbers are aggregated together to address different aspects of the stock market's characteristics, whether to measure price inflation or trading activity.

Asking what a stock market index is is tantamount to asking for what purpose it is used, since an index should ultimately be defined and constructed according to its intended use. An earlier paper by Marshall (1927)[page 60] notes

... One fact that emerges from the discussion of wholesale price index numbers is that the method of computation depends upon the purpose which the index is meant to serve. This principle applies with at least equal force to security price indexes, yet is one which does not seem to have received sufficient recognition by some who issue or publish them. The results which ensue from the computation of security price index numbers by different methods are so wide apart that the man on the street is sure to experience bewilderment when he sets them side by side. Something ought to be done to remove this difficulty.

Surprisingly there are very few papers in the literature which adequately discuss the intended use of stock market indexes, and how this impacts on their construction. The remainder of this thesis attempts to fill this void. Some papers which discuss their uses include Balk (1988), Göppl & Schütz (1993) and Nesbitt & Reynolds (1997). The first logical step in evaluating any stock market index series is to define in a precise manner the intended use of the series. Then, one has a criterion by which to judge the merits or defects of such a series and a guide as to how to channel efforts to improve on such a series. Some of the uses of stock market indexes are listed below—this forms the backdrop for the discussion in section 3.2.
1. One of the best-known and most common applications of a stock market index is as a measure of the rate of return on a given portfolio. Suppose, for example, that an investor invests a certain amount of money in some (fixed) basket of stocks. An index will track the price performance (capital gains) of this basket of stocks over time. Suppose at the time of investment that the index\(^2\) is normalized to \(R^0 = 100\) and at the end of the investment period the index attains the value \(R^1 = 150\). The index thus indicates a net (nominal) return of 50% on this particular portfolio. Clearly there are many such stock market gross return indexes, each corresponding to a particular portfolio composition. Typically, however, published index values (such as the DJIA or S & P 500) do not correspond to the investment portfolios of any one particular individual. Rather they are hypothetical portfolios in the sense that the change in index values between any two dates represents the return on the portfolio composed exactly like the index. These market portfolios involve an aggregation over all investors who have holdings in a particular basket of stocks (the basket corresponding to the index composition). An index-linked fund matches the performance of an index by holding stocks in proportion to the quantities of shares in the index. An increased index value denotes an increase in the value of the portfolio: this is the interpretation that one should accord to such statements as “the market is up”. In other words “the market is up” when the return on the portfolio composed exactly like the market portfolio increases over time.

2. A direct application of stock indexes as return measures is their use as benchmarks for investment performance in other types of assets, such as mutual funds, pension funds, or alternative portfolios of stocks. Thus an investor whose portfolio is returning +1% over a given time period for example, may wonder if there are feasible ways of reconstituting such a portfolio and improving on this seemingly low return. It is important, in this context to note that low and high returns are relative concepts. In other words, a particular portfolio (corresponding to a particular investment strategy) is only deemed

\(^2\)R is used to denote a return index.
to be performing poorly if alternative feasible portfolios (with similar levels of risk) show appreciably higher rates of return. Making such judgments requires an appropriate benchmark, one which closely matches the investment and business characteristics of the investor and portfolio in question. Since investment fund managers are often evaluated and remunerated relative to such benchmarks, their careful and thoughtful choice is extremely important for evaluating unique manager styles of investing. A characteristic of a good index is that it be investable, i.e. that investors can duplicate (hold) the benchmark index (see for example Nesbitt & Reynolds (1997)).

3. The S & P 500 is used by the (U.S.) government as one of the ten leading economic indicators since stock prices often incorporate investors' beliefs about the future earnings of certain companies. Investors buy stocks because they believe such stocks will rise in price, not because they have already gone up. It is also believed that relative fluctuations in stock prices are related to periods of recession or growth. To the extent that individuals invest a sizable percentage of their wealth in the stock market, stock indexes also indicate changes in the aggregate wealth levels in the economy.

4. Stock indexes also serve as the underlying asset for hedging risk of a given portfolio. Suppose an investor holds a portfolio which closely replicates some market portfolio.\(^3\) Straight indexing (or approximately so) entails portfolio risks because an index on the market can fluctuate significantly from one month to the next. Options and futures contracts essentially operate by making payments when stock market indexes rise above or fall below certain threshold (breakeven) levels. The seller of a stock index future (futures contract), for example, is theoretically supposed to deliver a portfolio of shares made up in the same proportions as the index at some specified date in the future. Typically however, such contracts are cash settled. Hedging (with futures and options) reduces portfolio risk: the idea is that a loss or profit made in the cash market will be counter-balanced by a profit or loss in the futures market. An investor with a portfolio

\(^3\)At the extreme one can consider an indexed portfolio which exactly matches the performance of an investment index.
of equities can protect his investment against a market fall by selling index futures with a contract value equal to the size of the portfolio. Any fall in the value of equities will be offset by profits on the futures and vice versa. This portfolio insurance feature of futures (and options) may not always reduce risk—as with benchmarking the choice of appropriate index is important. Control of risk involves identifying that stock index which most closely mimics the performance of a given portfolio—it is this index which serves as the basis for hedging portfolio risk. A cross-hedging risk may arise if the portfolio to be hedged does not behave in exactly the same way as the underlying asset (index) on which the futures contract is written.

3.2 Constructing a stock market index: preliminaries

The construction of a preliminary stock market index is considered here, motivated primarily by its use as a measure of return on a portfolio of stocks. The purpose of this construction is twofold: first, it illustrates some of the key features particular to the construction of stock market indexes; second, it provides the basis and background for a more thorough analysis of the construction and properties of such indexes undertaken in section 3.3 of this chapter and in chapters 4 and 5. The issue of for whom the index is being constructed is abstracted from—rather the focus is on the measurement of the (nominal) rate of gross return on a given basket of stocks for the single investor in an economy (or 'the market'). What then is an appropriate concept for a bilateral stock market index—one that will measure the return on a fixed basket of stocks at some comparison date \( t \), relative to its performance at some base date \( 0 \)? Assume that there exists some domestic stock market with a fixed number \( N \) of firms trading stocks, where \( p_i^t \) is the stock price of firm \( i = 1, \ldots, N \) at trading time \( t = 0, 1, \ldots, T \). Let \( q_i^t \) be the corresponding quantity of shares outstanding held by investors in

---

4The 'market' perspective is adopted in constructing stock market indexes. Indeed the indexes introduced here can be interpreted as the return measures for any individual investor, once the portfolio being evaluated has been identified.
the aggregate. Assume also for the moment that \( q_i^t = q_i^0 \; \forall \; t, i = 1, \ldots, N \), so that each firm's quantity of shares outstanding is fixed over time (or alternatively the quantities of shares held by an investor do not change over time). Let \( p^t = (p_1^t, \ldots, p_N^t)^\top \), \( q^t = (q_1^t, \ldots, q_N^t)^\top \) denote the \( N \)-vectors corresponding to the share prices and quantities of shares outstanding respectively at time \( t \). A reasonable concept for an index of stock prices at time \( t \), \( R^t \) — hereafter referred to as a stock market gross return index — is given by

\[
R^t = R(p^0, p^t, q^0) = \frac{p^t q^0}{p^0 q^0} = \frac{\sum_{i=1}^N p_i^t q_i^0}{\sum_{i=1}^N p_i^0 q_i^0},
\]

which is the ratio of the value of the portfolio at date \( t \) compared to the value of the portfolio at date 0. The index in equation 3.1 is a Laspeyres fixed basket price index where the quantities of shares outstanding are the weights associated with the stock prices. It compares the value of a market fixed basket of stocks at time \( t \) relative to the value of the same basket of stocks at some base time 0 (the time of purchase of the portfolio). It measures the capital gains on the portfolio over time, or alternatively it measures the current value of total market capitalization (relative to some initial (base) date) that the ‘market’ bought at the beginning of some period and held for successive periods. \( R^t, t = 0, 1, \ldots, T \) also tracks the price performance of an individual or institutional investor who holds stocks in proportion to the basket \( q^0 \).

---

5In the case of an individual investor \( q^0 = (q_1^0, \ldots, q_N^0)^\top \) are the quantities of shares held of each of the \( N \) stocks in the portfolio.

6The gross return \( R \) is the simple net rate of return \( r \) plus one, i.e. \( R = 1 + r \).

7The index number formula in equation 3.1 corresponds to the basic aggregation formula used by many statistical agencies in calculating consumer price indexes. A Laspeyres type CPI compares the cost of purchasing some fixed basket of goods and services over time, and can be thought of as a practical approximation to a ‘true’ cost-of-living index.

8Fisher (1927) and Marshall (1927) consider the construction of stock market price indexes with weights corresponding to the quantities of shares sold where the basic bilateral index is the Fisher ideal price index

\[
P_F(p^0, p^t, q^0, q^t) = \left( \frac{\sum_{i=1}^N p_i^0 q_i^0}{\sum_{i=1}^N p_i^t q_i^0} \right)^\frac{1}{2}.
\]  
Fisher argues that an index using quantities of shares sold as weights is a “speculator’s” index, while using quantities of shares outstanding is an “investor’s” index. Fisher maintains that the former captures speculative price movements.
3.2.1 Adjustments to the index

In reality, of course, constructing a stock index is not so simple. The class of stock market indexes in equation 3.1 compares the value (market capitalization) of the portfolio at the current date to its value at some base date for the same stocks—measuring the 'return' on the portfolio, where \( q^0 = (q^0_1, \ldots, q^0_N) \) are the quantities of shares in the portfolio over the evaluation period. The special character of the stock market follows from the fact that investors' returns are not only influenced by the stock price movements of the chosen portfolio of stocks, but also by stock/cash dividends and market capitalization changes among the firms. It will be important to clarify how such events should be incorporated into a stock index and how such actions affect the interpretation of the index. Clearly the index in equation 3.1 cannot be used indefinitely and without modification. Other reasons why adjustments need to be made to the index include some of the following:

1. firms may go bankrupt and newer dominant firms may emerge on the market;
2. firms may issue new shares from time to time or engage in large scale buybacks (re-purchases);
3. mergers and acquisitions may occur;
4. the base period may become too distant for useful and meaningful comparisons. This is why countries occasionally rebase their consumer price indexes.\(^9\)

The important distinction between a fixed base and a chained sequence of price indexes is explained below in the context of measuring price change for a fixed basket of commodities.

3.2.2 Fixed base and chaining

Consider the construction of an index number series extending over \( t \geq 3 \) time periods.\(^10\) Given price and quantity data \( p_i, q_i, i = 1, 2, \ldots, T \), and some bilateral price index

\(^9\)Statistics Canada is currently changing its official CPI time base from 1986=100 to 1992=100.
\(^10\)In the standard price index literature a price index is a positive function \( P : \mathbb{R}_{++}^{MN} \rightarrow \mathbb{R}_{++} : (p^0, p^1, q^0, q^1) \rightarrow P(p^0, p^1, q^0, q^1) \), where the comparison period is indexed by 1 and the base period is
for measuring the price movement over two periods $t$ and $t + 1$, the fixed base price sequence is given by

$$P^0 = 1 = P(p^0, p^0, q^0, q^0), P^1 = P(p^0, p^1, q^0, q^1), \quad P^2 = P(p^0, p^2, q^0, q^2), \quad \ldots, P^T = P(p^0, p^T, q^0, q^T).$$

Note that each period price level is compared to the base period (0) level. A chained price sequence, in contrast, is defined by

$$P^0 = 1 = P(p^0, p^0, q^0, q^0), P^1 = P(p^0, p^1, q^0, q^1), \quad P^2 = P(p^0, p^1, q^0, q^1)P(p^1, p^2, q^1, q^2), \quad \ldots, P^T = \Pi_{t=0}^{T-1} P(p^t, p^{t+1}, q^t, q^{t+1}).$$

The period $T$ price index is the cumulative product of all interperiod price indexes, so that each period price index is successively rebased in computing the index. Thus, the price level in period $T$ is compared with the price level in period $T - 1$, which in turn is compared with the price level in period $T - 2$, and so on. There is thus a continual update of baskets.

Marshall (1887) [page 373] is credited with first proposing the chain principle to overcome the difficulty of making price comparisons over distant periods due to the introduction of new goods. In the present context 'new goods' can be interpreted as the shares outstanding of new firms included in the index or as new stock issues (i.e. changing quantity weights) of firms currently included in the index. Fisher (1911) [page 204], who gave the chaining procedure its name writes

It may be said that the cardinal virtue of the successive base or chain system is the facility it affords for the introduction of new commodities, the dropping indexed by 0 and $p^t, q^t, t = 0, 1$ denote the corresponding vectors of prices and quantities consumed of the $N$ goods respectively. Thus $q^t, t = 0, 1$ refer to quantities of consumption for period $t$ which is of a certain duration. This differs from the current interpretation of $R(p^0, p^1, q^0)$ as measuring the gross return on a fixed portfolio $(q^0)$ over the evaluation period which begins at date 0 and ends at date 1.
out of obsolete commodities, and the continued readjustment of the system of weighting to new commodities.

The chain principle is invariant to changes in the base period and is one way of introducing (or retiring) goods into (or from) an index.

In light of the fixed base and chaining methods, and the allusions made to potential changes in the composition of a stock market portfolio over time, procedures need to be introduced to adequately deal with such changes. For example, when a firm \( r \leq N \) replaces a firm \( s \leq N \) in an \( N \)-stock index at time \( t \), the value of the portfolio (i.e. the current market capitalization) becomes \( \sum_{i \neq s}^{N} p_i^t q_i^0 \) (assuming once again that \( q_i^t = q_i^0 \forall t, i = 1, \ldots, N \)). This can no longer be meaningfully compared to \( \sum_{i \neq r}^{N} p_i^t q_i^0 \) (the base market capitalization) since comparisons are no longer being made over the same basket of stocks and firms.

To maintain comparability and continuity in the index, the index is basket-updated and the two indexes (i.e. the one in which \( s \) is excluded and the other in which \( r \) is excluded) are 'linked' together at the date when the basket changes. The assumption (or interpretation) underlying this chaining principle is the following: the initial portfolio (i.e. \( q^0\backslash\{q_r^0\} \)) is held until the composition of the index is changed.\(^{11}\) Prior to the change, the portfolio has a certain value (both in terms of index points and dollar equivalent). This portfolio is sold off and the proceeds are invested instantaneously and costlessly in the new portfolio in proportion to the share value of each stock in the new portfolio. This new basket thus forms the basis of comparison for all subsequent periods until the next component change occurs. The chaining procedure prevents a change in the composition of the index from directly affecting the index value and provides a convenient method for maintaining continuity in the index in the face of a change in the portfolio composition.

\(^{11}\)For the moment imagine that the investor initiates the change—i.e. he wants to change his investment strategy, because for example he has changed his expectations of future profitability of \( s \).
3.3 Three considerations for a stock index

Some other basic considerations that must be addressed in constructing any stock market gross return index are now considered. It will be seen that such decisions result in very different measured index levels, with corresponding implications for performance measurement and evaluation of different investment strategies. These considerations are examined by posing a number of questions related to the basic index in equation 3.1, and offering some plausible answers.

3.3.1 Choice of basket

Valuing a fixed basket of stocks over time seems the most natural candidate for a stock market gross return index where the quantities of stocks in the basket \( q^0 \) are fixed. The question which arises naturally in this instance is: how is \( q^0 \) chosen? From the individual investor's perspective \( q^0 \) is the basket of stocks which is purchased at some initial base date 0. For benchmarking purposes the portfolio \( q^0 \) is chosen to closely resemble the investment and business characteristics of the investor portfolio being evaluated. Clearly the number of potential \( q^0 \)'s is large—each corresponding to the particular quantities of shares purchased by each individual investor. Two reasonable (and feasible) choices for \( q^0 \) are

1. the market portfolio;
2. the equally-weighted portfolio.

In choice 1 \( q^0 \) is the aggregation over all investors of all the shares outstanding for a particular chosen basket of stocks. In choice 2 the 'quantity' weights are chosen in such a way that an equal dollar amount is invested in each stock in a given portfolio. For example, in the case of a $1 investment in \( N \) stocks \( q^0 \) is chosen so that $\frac{1}{N}$ is invested in each of the \( N \) stocks. If the base date prices are \( p^0 = (p_1^0, \ldots, p_N^0)^T \), then \( q^0 \) is chosen so that \( p_i^0 q_i^0 = \frac{1}{N} \), \( i = 1, \ldots, N \) and \( q^0^T q^0 = 1 \). It is easy to see that the resulting \( q^0 \), call it \( q^0 \) is
where $\tilde{p}^0$ is the diagonal matrix whose (diagonal) elements are $p^0_i, i = 1, \ldots, N$. In this case $p^0\tilde{p}^0 = 1$. This particular choice of $q^0$ corresponds perhaps to the investment strategy of an individual investor that invests randomly in the market.

### 3.3.2 Treatment of cash dividends

The *total return* to any investor is affected by changes in the share prices, cash from dividends, and sales of subscription privileges. Fisher (1966) distinguishes between an *investment performance index* and a *stock price index* since the former takes all cash dividends (as well as capital changes and changes in share prices) into account. Fisher (1966)[page 193] notes

I contend that a stock-market index based on price alone is intrinsically uninteresting . . . When indexes of price alone are used, they may give a fairly accurate picture of investment performance in the short run, because in the short run
dividends are only a small part of investment performance. However, even the comparison of two short-run periods is likely to be misleading if dividends are ignored.

This distinction between the two concepts is not made in this thesis—when reference is made to a stock market gross return index, it is the price performance (capital return) which is monitored. The terms total return or stock market total gross return index are reserved for indexes which take both capital gains and income gains fully into account. By tracking the value of such a total return index over time, the user can evaluate the total (nominal) return earned on a given portfolio. Calculation of a total return index requires an appropriate treatment of cash dividends. Some possible strategies are:

1. the cash payments are invested outside the portfolio at (say) some rate of interest such as that on Treasury bills;

2. the cash payments are used to purchase additional shares of all stocks in the portfolio in proportion to the value share of each stock in the portfolio on the ex-dividend date;

3. the cash payments are used to purchase additional shares of the dividend-generating stock.

It is apparent then, that one cannot speak of the total return index: in fact there are several such indexes—each corresponding to the particular portfolio chosen by the investor and the manner in which the cash dividends are reinvested (or not) in the portfolio. In fact Balk (1988) refers to option 2 as a Reinvestment Total (RP) policy, and to option 3 as a Reinvestment Partial (RP) policy. He notes that

... It is clear that the total return index depends on three factors: the portfolio composition at date 0, the policy followed, and the behaviour of stock prices and returns during the period \([0, t]\).

'Policy', in this context refers to the actions taken by the investor over the evaluation period such as the buying and selling of stocks, the reinvestment of dividends etc. The total return
concept of the stock index that is most often adopted corresponds to the following investment strategy (option 2 above): an investor has some initial monetary outlay $A^0$ which is to be invested in some portfolio of stocks. The proportions of $A^0$ to be invested in each of the stocks correspond to the weights in the index (the quantity of shares associated with each stock). Cash dividends are used to buy additional shares in the stocks in the portfolio—this is done in proportion to the value share of each stock in the portfolio on the ex-dividend date. For example, let $q_{i}^{t-1}$ be the number of shares of firm $i$ held by the investor at date $t - 1$ (the date before the ex-dividend date $t$). The ‘new’ number of shares $q_i^t$ (i.e. after using the cash dividends to purchase additional shares) is given by

$$q_i^t = q_{i}^{t-1} + \left( \frac{p_i^t q_{i}^{t-1}}{\sum_{j=1}^{N} p_j^t q_{j}^{t-1}} \right) \left( \frac{D_t^t}{p_i^t} \right),$$

(3.2)

where $D_t^t = \sum_{i=1}^{N} d_i^t q_{i}^{t-1}$ is the monetary value of the cash dividends on the ex-dividend date.

$\frac{D_t^t}{p_i^t}$ is the quantity of new shares of firm $i$ that could be purchased using $D_t^t$—this is scaled by the factor $\frac{p_i^t q_{i}^{t-1}}{\sum_{j=1}^{N} p_j^t q_{j}^{t-1}}$, the value share of stock $i$ in the portfolio at date $t$ (prior to reinvestment). Thus the new quantity of shares equals the ‘old’ quantity $q_{i}^{t-1}$ plus the additional shares purchased, given by the last expression on the right-hand-side of equation 3.2.\footnote{It is straightforward to show that if the ‘policy’ followed by the investor is to reinvest dividends in proportion to the share value of each stock, then the index at date $T$ (using equation 3.1) is given by $R_T^T = \prod_{t=1}^{T} \left( \frac{(p_i^t + d_i^t)^{q_i^t}}{p_i^{t-1} q_i^{t-1}} \right)$, where $d^t = (d_1^t, \ldots, d_N^t)^\top \geq (0, \ldots, 0)^\top$ is a $N$-vector of non-negative cash dividends paid out on date $t$. For option 3, the expression corresponding to 3.2 is

$$q_i^t = q_{i}^{t-1} + \frac{d_i^t q_{i}^{t-1}}{p_i^t}, \quad i = 1, \ldots, N,$$

(3.3)

and the corresponding index at $T$ is $R_T^T = \frac{\sum_{i=1}^{N} p_i^T q_i^T \left( \frac{\pi_i^T}{\sum_{i=1}^{N} q_i^T} \right)}{\sum_{i=1}^{N} p_i^T q_i^T}$. See Balk (1988) for more on these total return indexes.}

No cash flows in/out of the portfolio. Thus the total net return on the portfolio at date $T$ is $\frac{R_T^T - R_0^0}{R_0^0}$ where $R_0^0$ can be normalized to equal the initial investment $R_0^0 = A^0$.

The implicit assumption underlying options 2 and 3 is that the returns on alternative and feasible investment opportunities are irrelevant, or stated differently, that alternative
investment strategies yield identical rates of return to the current portfolio so that reinvest-
ment in the current portfolio is as reasonable an option as any other. It is unlikely that this
assumption about investment strategies will hold in practice—it is likely that the average
individual investor may withdraw cash dividends from his portfolio.

Consider the following simple example. An investor has $1 to invest. He purchases one
stock whose initial price is $1. At the end of the period the stock price has risen to
$p^1 = $1.50, in addition to which the firm pays a cash dividend of $0.50. The total
(gross) return from the dollar investment is thus

$$ R^1 = \left( \frac{p^1 + d}{p^0} \right) \times 100\% = 200\%. \quad (3.4) $$

If the share price $p^1 = $1.50 subsequently rises to $p^2 = $2, the second period gross return
is given by $p^2 / p^1 = \frac{4}{3}$. Thus the total return over the two periods is given by the (chained)
index:

$$ R^2 = \left( \frac{1 + \frac{d}{p^1}}{p^0} \right) \times 100\% = \left( \frac{p^1 + d}{p^0} \right) \left( \frac{p^2}{p^1} \right) \times 100\% = 266\frac{2}{3}\%. \quad (3.5) $$

Note that equation 3.5 is based on the assumption that the dividend is reinvested in the
stock.\(^{13}\) The $0.5 dividend received during the first period could have bought $\frac{0.5}{1.5} = 0.33
shares of stock at the end of the first period giving a total holding of 1.33 shares. The
share holdings then have a value of $1.33 \times $2 = $2.66 at the end of the second period,
which corresponds to equation 3.5. In the case where the cash dividend is used for con-
sumption purposes (i.e. it is not reinvested) the total gross return over this time period is

$$ \left( \left( \frac{p^2}{p^1} \right) \left( \frac{p^1}{p^0} \right) + \frac{d}{p^1} \right) \times 100\% = 250\%. $$

### 3.3.3 Chaining and Rebalancing

As has been mentioned the chain principle is used to account for making comparisons over
longer periods when the basket may change in different dimensions (such as component

\(^{13}\)Note that since this portfolio consists of only one stock, options 2 and 3 above are identical.
changes, reinvestment of cash dividends etc.). While chaining does provide a means of maintaining continuity and comparability in the index over time, what is less clear is how often chaining/rebalancing should be undertaken. Possible frequencies of rebalancing are daily, weekly or monthly. Indeed quarterly or even annual rebalancing are other possible options. At any given time then the options are either to rebalance or maintain a 'buy-and-hold' strategy. The choice of periodicity of rebalancing may well be determined by expense—rebalancing for the individual investor involves transactions costs. Thus maintaining a fixed basket (buy-and-hold strategy) for longer periods may be a compromise between expense of rebalancing and maintaining a constant basket of 'reasonable' performance.

3.3.4 Examples

Some examples are considered below which illustrate the degrees of freedom which can be exercised in constructing a stock market gross return index with respect to the choice of basket \( q^0 \), the treatment of cash dividends, and the periodicity (if any) of rebalancing. In each case considered the index number values do differ, with implications for measuring return, performance evaluation and benchmarking. The challenge that remains is to ask which method is most appropriate?

1 period examples

Consider first an example involving a two date (one period) comparison. Two points are worth explicit mention:

1. the quantities of shares \( q^0 \) in the basket of stocks is fixed over the evaluation period so that rebalancing is not an issue;

---

\[14\] The terms 'chaining' and 'rebalancing' are synonymous here and can be interpreted to mean 'basket updating' or 're-weighting'. However 'rebalancing' will typically be reserved for the special case where the basket composition changes or the number of shares held changes for reasons other than the use of cash dividends to purchase additional shares.

\[15\] To relate this to the consumer price situation, the Bureau of Labor Statistics conducts basket updates for the U.S. CPI every ten years, while Statistics Canada updates the Canadian CPI basket every four years.

46
2. it is assumed that cash dividends are paid out at the end of the investment period so that their reinvestment is not a consideration.

The initial vector of prices is \( p^0 = (p^0_1, \ldots, p^0_N)^T \) and the comparison vector of prices is \( p^1 = (p^1_1, \ldots, p^1_N)^T \). A vector of cash dividends \( d^1 = (d^1_1, \ldots, d^1_N)^T \geq 0_N \) is paid out at the end of the evaluation period where \( 0_N = (0, \ldots, 0)^T \) is a \( N \)-vector of zeros. Then the return on the portfolio is given by

\[
R^1 = R(p^0, p^1, d^1, q^0) = \frac{(p^1 + d^1)^T q^0}{p^0 q^0},
\]

where \( q^0 \) is the 'market' basket of stocks. If the basket is chosen so that \( \frac{1}{N} \) is invested in each of \( N \) stocks, then

\[
R^1 = R(p^0, p^1, d^1, q^0) = \frac{(p^1 + d^1)^T q^0}{p^0 q^0} = \frac{1}{N} \sum_{i=1}^{N} \left( \frac{p_i^1}{p_i^0} \right) + \frac{1}{N} \sum_{i=1}^{N} \left( \frac{d_i^1}{p_i^0} \right).
\]

Note that the index in equation 3.7 is a Carli price index, i.e. it is the arithmetic average of individual stock gross returns. While it appears that quantity weights are absent from the index in equation 3.7, one can say that the index contains implicit quantity weights—the 'quantity' weights correspond to inverse prices so that an equal dollar investment (\( \frac{1}{N} \)) is assigned to each of the \( N \) stocks in the basket.

2 period examples

When measuring the return performance over three dates two 'complications' arise. These concern the treatment of cash dividends (i.e. whether or not they should be reinvested in the portfolio as received) and whether or not the current portfolio needs to be rebalanced\(^{16} \) (basket updated). Thus several combinations and permutations of choices are possible. These are considered in turn.

\(^{16}\)Rebalancing' here refers to changes in the quantity of shares/portfolio composition for reasons other than the use of cash dividends for purchasing additional shares.
1. **Basket:** $q^0$ (market),

**Cash dividend:** invest in T-bills or a deposit account at interest rate $r^1$,

**Rebalance:** no.

In this case the two period return is computed as

$$R^2 = R(p^0, p^1, p^2, d^1, q^0) = \frac{p^{2^T} q^0}{p^{0^T} q^0} + \frac{d^{1^T} q^0}{p^{0^T} q^0} (1 + r^1),$$

(3.8)

and over $T$ periods this becomes

$$R^T = R(p^0, \ldots, p^T, d^1, q^0) = \frac{p^{T^T} q^0}{p^{0^T} q^0} + \frac{d^{1^T} q^0}{p^{0^T} q^0} (1 + r^1)(1 + r^2) \cdots (1 + r^T),$$

(3.9)

where the dividend return is invested at a compounded rate over the evaluation period and $r^t, t = 1, \ldots, T$ is the period $t$ interest rate. For benchmarking purposes $r^t$ can be thought of as the opportunity cost (in terms of returns that could have been realized) of holding the cash dividends as cash holdings.

2. **Basket:** equally-weighted, $q^0 = (\bar{p}^0)^{-1}$

**Cash dividends:** invest in T-bills,

**Rebalance:** no.

$$R^2 = R(p^0, p^1, p^2, d^1, q^0) = \frac{p^{2^T} q^0}{p^{0^T} q^0} + \frac{d^{1^T} q^0}{p^{0^T} q^0} (1 + r^1) = \frac{1}{N} \sum_{i=1}^{N} \left( \frac{p_i^2}{\bar{p}_i^0} \right) + \frac{1}{N} \sum_{i=1}^{N} \left( \frac{d_i^1}{\bar{p}_i^0} \right) (1 + r^1).$$

(3.10)

3. **Basket:** market,

**Cash dividends:** reinvest in the stocks in proportion to the value share of each stock on the ex-dividend date,

---

17It is assumed here that there are dividend payments at the end of period one only and none in other periods. Additional dividends can be handled in an entirely analogous manner.
Rebalance: no.

\[ R^2 = R(p^0, p^1, p^2, d^1, q^0) = \left( \frac{(p^1 + d^1)^\top q^0}{p^0 \cdot q^0} \right) \left( \frac{p^{2\top} q^0}{p^{1\top} q^0} \right). \] (3.11)

4. Basket: equally-weighted,

**Cash dividends:** reinvest in the stocks in proportion to the value share of each stock on the *ex-dividend* date,

Rebalance: no.

\[ R^2 = R(p^0, p^1, p^2, d^1, q^0) = \left( \frac{1}{N} \sum_{i=1}^{N} \left( \frac{p_i^1 + d_i^1}{p_i^0} \right) \right) \left( \frac{1}{N} \sum_{i=1}^{N} \left( \frac{p_i^{2\top}}{p_i^{1\top}} \right) \right). \] (3.12)

5. Basket: market,

**Dividends:** reinvest,

Rebalance: yes, at the beginning of every period at each basket change.

\[ R^2 = R(p^0, p^1, p^2, d^1, q^0) = R(p^0, p^1, d^1, q^0) R(p^1, p^2, q^0) = \left( \frac{(p^1 + d^1)^\top q^0}{p^0 \cdot q^0} \right) \left( \frac{p^{2\top} q^1}{p^{1\top} q^1} \right). \] (3.13)

6. Basket: equally-weighted,

**Dividends:** reinvest,

Rebalance: yes, to equal-dollar weighting in each period.

\[ R^2 = R(p^0, p^1, p^2, d^1, q^0, q^1) = R(p^0, p^1, d^1, q^0) R(p^1, p^2, q^1) = \left( \frac{(p^1 + d^1)^\top q^0}{p^0 \cdot q^0} \right) \left( \frac{p^{2\top} q^1}{p^{1\top} q^1} \right). \] (3.14)

\[ = \left( \frac{1}{N} \sum_{i=1}^{N} \left( \frac{p_i^1 + d_i^1}{p_i^0} \right) \right) \left( \frac{1}{N} \sum_{i=1}^{N} \left( \frac{p_i^2}{p_i^1} \right) \right). \] (3.15)
<table>
<thead>
<tr>
<th>Date</th>
<th>Share Price (dollars)</th>
<th>Shares outstanding</th>
<th>Cash dividend (dollars)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>( p_1^0 = 10 )</td>
<td>( q_1^0 = 1 )</td>
<td>( d_1^1 = 1^a )</td>
</tr>
<tr>
<td></td>
<td>( p_2^0 = 2 )</td>
<td>( q_2^0 = 2 )</td>
<td>( d_2^1 = 0 )</td>
</tr>
<tr>
<td></td>
<td>( p_3^0 = 1 )</td>
<td>( q_3^0 = 12 )</td>
<td>( d_3^1 = 1 )</td>
</tr>
<tr>
<td>1</td>
<td>( p_1^1 = 11 )</td>
<td>( q_1^1 = 2 )</td>
<td>( d_1^2 = 0 )</td>
</tr>
<tr>
<td></td>
<td>( p_2^1 = 3 )</td>
<td>( q_2^1 = 2 )</td>
<td>( d_2^1 = 0 )</td>
</tr>
<tr>
<td></td>
<td>( p_3^1 = \frac{1}{2} )</td>
<td>( q_3^1 = 12 )</td>
<td>( d_3^1 = 0 )</td>
</tr>
<tr>
<td>2</td>
<td>( p_1^2 = 11 )</td>
<td>( q_1^2 = 2 )</td>
<td>( d_1^3 = 0 )</td>
</tr>
<tr>
<td></td>
<td>( p_2^2 = 4 )</td>
<td>( q_2^2 = 2 )</td>
<td>( d_2^3 = 0 )</td>
</tr>
<tr>
<td></td>
<td>( p_3^2 = 1 )</td>
<td>( q_3^2 = 12 )</td>
<td>( d_3^3 = 0 )</td>
</tr>
</tbody>
</table>

*Notation: \( d_i^t \) means that the dividend for firm \( i \) is paid out at the end of evaluation period \((t-1, t)\), where \( t \) is the ex-dividend date. Thus \( d_i^t \) is paid out at the end of the \((0, 1)\) evaluation period.*

Table 3.1: Calculating a 3-stock index

As an illustration of these differences, consider the calculation of a 3-stock index using the information in table 3.1. The return index in table 3.2 ignores cash dividends whereas the total return index takes cash dividends in addition to price changes in stocks fully into account. In all cases \( R^0 \) has been normalized so that \( R^0 = 100 \). A change in the quantity of shares outstanding from one period to the next for an individual stock corresponds to a share issue by that particular company, or equivalently an increase in the share holdings of an individual investor. Note that unlike the case of the reinvestment of cash dividends, where the 'new' quantities are determined by the dividends received, the '\( q^1 \)' in cases 5 and 6 can be taken to be determined independently of the magnitude of the cash dividends. They are chosen by the investor by other means. Table 3.2 shows that the index levels are different in all cases.
<table>
<thead>
<tr>
<th>Scenario</th>
<th>Return Index</th>
<th>Total Return Index</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Base date ((R^0))</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All indexes</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td><strong>1 period</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Market basket</td>
<td>88.46</td>
<td>138.46</td>
</tr>
<tr>
<td>Equally-weighted basket</td>
<td>103.33</td>
<td>140.00</td>
</tr>
<tr>
<td><strong>2 period</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case 1(^a)</td>
<td>119.23</td>
<td>174.23</td>
</tr>
<tr>
<td>Case 2(^b)</td>
<td>136.66</td>
<td>177.00</td>
</tr>
<tr>
<td>Case 3</td>
<td>109.28</td>
<td>171.04</td>
</tr>
<tr>
<td>Case 4</td>
<td>149.26</td>
<td>202.22</td>
</tr>
<tr>
<td>Case 5</td>
<td>119.23</td>
<td>186.62</td>
</tr>
<tr>
<td>Case 6</td>
<td>136.66</td>
<td>185.16</td>
</tr>
</tbody>
</table>

\(^a\) \(r^1 = 10\% \) is assumed here.
\(^b\) \(r^1 = 10\% \) is assumed here.

Table 3.2: Index values based on a 3-stock index
Chapter 4

Stock Market Indexes in Practice

Chapter 3 outlined some of the main considerations that need to be addressed in constructing a stock market gross return index or a stock market total gross return index. These include the choice of basket, the treatment of cash dividends, and the frequency (if any) of chaining over the evaluation period. This chapter examines the extent to which these types of considerations are incorporated into the construction of actual stock indexes in use today. A brief discussion of the properties of these indexes is also included.

The starting point is an examination of the Dow Jones Industrial Average (DJIA), given its position as the best-known and most-cited index of financial market performance in economic and financial circles. In spite of this, it is known that there are some conceptual problems with the DJIA as a financial index, and that its calculation method is the exception rather than the rule. However, these criticisms have not formally identified the precise nature of the properties of the DJIA that give rise to its shortcomings. Indeed there seems to be confusion over the precise interpretation of the DJIA. Section 4.1 of this chapter addresses these two important issues.

Section 4.2 explains the calculation of three other types of indexes and shows that they all differ from the DJIA in the manner in which they are calculated. Furthermore, each is consistent with a particular investment strategy. The chapter concludes with an empirical comparison of the different index number formulae used in computing return indexes. The actual data employed in computing the DJIA are used. The different formulae can lead to
very different measured returns, particularly over shorter time periods.

4.1 The Dow Jones Industrial Average

4.1.1 A Brief History

The DJIA is the oldest and best-known indicator of stock market performance in the United States. Charles Henry Dow, co-founder of Dow Jones and Company first published his stock average in July 1884 in the company’s newsletter Customer’s Afternoon Letter—this newsletter ultimately evolved into the Wall Street Journal.

Dow’s original (unilateral) stock price index at time $t$, $R_{DJ}^t$, was computed as the arithmetic average of the prices of the eleven stocks that made up its composition, or

$$R_{DJ}^t = \frac{\sum_{i=1}^{11} p_i^t}{11}.$$  \hfill (4.1)

The components of the original Dow index are shown in table 4.1. Interestingly, nine of the eleven original components were railway companies, reflecting their economic importance at the turn of the nineteenth century. In 1896 the Wall Street Journal published an average based on twelve industrials, in addition to a railroad average based on twenty railroads. In 1916 the coverage of the DJIA was extended to twenty stocks and then to thirty in 1928, the total currently included in the average. The Dow Jones Utility Average was first published in 1929, and in 1970 the Railroad Average was renamed the Transportation Average and began to include companies other than railways in its coverage.

Dow originally created his indexes to give a crude indication of the average level of prices in the stock market and to gauge whether stock prices in the aggregate were rising or falling. The mechanics of computing the index were dictated by the necessity of having to compute the averages by pencil and paper. Indeed such simple price indexes (based on prices alone) can be traced back to the early indexes proposed by Carli (an arithmetic mean of price relatives), Dutot (a ratio of arithmetic mean prices), and Jevons (a geometric mean of price relatives).
Table 4.1: DJIA Components: 1884

<table>
<thead>
<tr>
<th>1884</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicago &amp; Northwestern</td>
</tr>
<tr>
<td>Chicago, Milwaukee &amp; St. Paul</td>
</tr>
<tr>
<td>Delaware, Lackawanna</td>
</tr>
<tr>
<td>Lake Shore</td>
</tr>
<tr>
<td>Louisville &amp; Nashville</td>
</tr>
<tr>
<td>Missouri Pacific</td>
</tr>
<tr>
<td>New York Central</td>
</tr>
<tr>
<td>Northern Pacific (pfd.)</td>
</tr>
<tr>
<td>Pacific Mail</td>
</tr>
<tr>
<td>Union Pacific</td>
</tr>
<tr>
<td>Western Union</td>
</tr>
</tbody>
</table>

The composition of the DJIA in 1928 and 1998 is shown in table 4.2. Interestingly, General Electric Company is the only stock included in the original (i.e. from 1896) industrial average that is still a component today.\(^1\) Since its inception as a thirty stock index in 1928, changes in the DJIA have reflected the rising importance of the services, entertainment and technology sectors. For example, IBM was added in 1979, American Express in 1982, McDonalds in 1985, Walt Disney in 1991, and Hewlett Packard and Wal-Mart in 1997.

Currently the DJIA tracks the price performance of thirty ‘blue-chip’ stocks listed on the New York Stock Exchange (NYSE). The index at time \(t\), \(R_{DJ}^t\), is calculated as

\[
R_{DJ}^t = \frac{\sum_{i=1}^{30} p_i^t}{d}, \tag{4.2}
\]

where \(d\) is the current divisor. Note that in order for \(R_{DJ}^t\) in equation 4.2 to be a proper arithmetic average, \(d\) would have to be thirty. In fact \(d\) is not thirty—its current value is

\(^1\)GE itself has changed drastically from a company concerned primarily with developing applications of electric power into a giant conglomerate with interests in broadcasting, finance and consumer goods as well as industrial and power equipment. If it had remained the company it originally was, GE probably would have been dropped from the index long ago.
<table>
<thead>
<tr>
<th>1928</th>
<th>1998</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allied Chemical &amp; Dye Co.</td>
<td>AlliedSignal Inc.</td>
</tr>
<tr>
<td>American Can Co.</td>
<td>Aluminum Co. of American</td>
</tr>
<tr>
<td>American Smelting &amp; Refining</td>
<td>American Express Co.</td>
</tr>
<tr>
<td>American Sugar Refining Co.</td>
<td>AT &amp; T Corp.</td>
</tr>
<tr>
<td>American Tobacco (Class B)</td>
<td>Boeing Co.</td>
</tr>
<tr>
<td>Atlantic Refining Co.</td>
<td>Caterpillar Inc.</td>
</tr>
<tr>
<td>B. F. Goodrich Co.</td>
<td>Chevron Corp.</td>
</tr>
<tr>
<td>Bethlehem Steel Corp.</td>
<td>Coca Cola Co.</td>
</tr>
<tr>
<td>Chrysler Corp.</td>
<td>Du Pont Co.</td>
</tr>
<tr>
<td>F. W. Woolworth Co.</td>
<td>Eastman Kodak Co.</td>
</tr>
<tr>
<td>General Electric Co.</td>
<td>Exxon Corp.</td>
</tr>
<tr>
<td>General Motors Corp.</td>
<td>General Electric Co.</td>
</tr>
<tr>
<td>General Railway Signal Corp.</td>
<td>General Motors Corp.</td>
</tr>
<tr>
<td>International Harvester Co.</td>
<td>Goodyear Tire &amp; Rubber Co.</td>
</tr>
<tr>
<td>International Nickel Co.</td>
<td>Hewlett-Packard Co.</td>
</tr>
<tr>
<td>Mack Trucks Inc.</td>
<td>International Business Machines Corp.</td>
</tr>
<tr>
<td>Nash Motors Co.</td>
<td>International Paper Co.</td>
</tr>
<tr>
<td>Paramount Famous Lasky Corp.</td>
<td>Johnson &amp; Johnson</td>
</tr>
<tr>
<td>Postum Inc.</td>
<td>McDonald’s Corp.</td>
</tr>
<tr>
<td>Radio Corp. of America</td>
<td>Merck &amp; Co.</td>
</tr>
<tr>
<td>Sears, Roebuck &amp; Co.</td>
<td>Minnesota Mining &amp; Manufacturing Co.</td>
</tr>
<tr>
<td>Standard Oil Co. (New Jersey)</td>
<td>Philip Morris Co.</td>
</tr>
<tr>
<td>Texaco Co.</td>
<td>Proctor &amp; Gamble</td>
</tr>
<tr>
<td>Texas Gulf Sulphur</td>
<td>Sears, Roebuck &amp; Co.</td>
</tr>
<tr>
<td>U. S. Steel Corp. (Common)</td>
<td>Travelers Group Inc.</td>
</tr>
<tr>
<td>Union Carbide &amp; Carbon Corp.</td>
<td>Union Carbide Corp.</td>
</tr>
<tr>
<td>Victor Talking Machine Co.</td>
<td>United Technologies Corp.</td>
</tr>
<tr>
<td>Westinghouse Electric &amp; Manufacturing Co.</td>
<td>Walt Disney Co.</td>
</tr>
<tr>
<td>Wright Aeronautical Corp.</td>
<td>Wal-Mart Stores Inc.</td>
</tr>
</tbody>
</table>

Table 4.2: DJIA Components: 1928 and 1998
approximately 0.25. Numerous adjustments have been made to the divisor over the years to preserve historical continuity and comparability over time. \( d \) is adjusted for stock splits, stock dividends (of 10% or greater), or changes in the component stocks which are used in computing the DJIA. The method for adjusting the divisor is examined in section 4.1.3.

4.1.2 Features of the Dow

The principal difference between the DJIA and other common stock indexes (such as the class of indexes introduced in equation 3.1 (section 3.2, chapter 3)) is the fact that the former is a price-weighted index while the latter are capitalization-weighted indexes. The DJIA implicitly weights each stock by its price (or equivalently each stock is assigned a constant quantity weight of one in the numerator) so that a dollar change in any stock produces the same change in the index. Explicit (quantity) weights are absent from its calculation. Another consequence of the index's construction is that higher priced stocks have a greater impact on the index than lower priced stocks (independent of the market capitalization of the respective firms). For example, a 10% increase in the price of a $500 stock has ten times the impact on the unilateral index as a 10% increase in the price of a $50 stock.²

An example will illustrate these features. Table 4.3 shows price and quantity (i.e. shares outstanding) data for two firms 1 and 2 at three dates \( t = 0, 1, 2 \). Using the information in table 4.3 the DJIA \( R_{DJ}^t, t = 0, 1, 2 \) for the three dates is given by

\[
R_{DJ}^0 = 51, R_{DJ}^1 = 51\frac{1}{2}, R_{DJ}^2 = 52, \quad (4.3)
\]
or alternatively if \( R_{DJ}^0 \) is assigned a base value of 100, then

\[
R_{DJ}^0 = 100, R_{DJ}^1 \sim 100.98, R_{DJ}^2 \sim 101.96. \quad (4.4)
\]

The same result obtains if \( p_2^t = 2, t = 0, 1, 2 \) and \( p_1^0 = 100, p_1^1 = 101, p_1^2 = 102 \). Compare this with a Laspeyres index \( (R_L) \) which prices out the market portfolio of stocks over the

²The impact of these changes on indexes such as the arithmetic or geometric mean of individual stock gross returns would be the same.
Table 4.3: DJIA versus Laspeyres index

<table>
<thead>
<tr>
<th>Date</th>
<th>Company</th>
<th>Share Price (cents)</th>
<th>Shares outstanding</th>
<th>Market value (cents)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>100</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>( \sum_{i=1}^{2} p_i^0 q_i^0 )</td>
<td>2,100</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>100</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>( \sum_{i=1}^{2} p_i^1 q_i^0 )</td>
<td>3,100</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>100</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>( \sum_{i=1}^{2} p_i^2 q_i^0 )</td>
<td>4,100</td>
</tr>
</tbody>
</table>

three dates. Setting \( R_L^0 = 100 \) results in the following series of index numbers

\[
R_L^0 = 100, \quad R_L^1 \sim 147.62, \quad R_L^2 \sim 195.24. \tag{4.5}
\]

Here the capital return on the portfolio is significantly higher than that indicated by the DJIA which ignores the quantities of shares of each of the companies in the portfolio. The example in table 4.3 is not entirely innocuous. Stock prices alone are *not* a true reflection of the economic importance of a company nor of the value contribution to an investor's portfolio with particular holdings of stocks. For example, in April 1988 Merck was the most expensive Dow stock (and thus had the greatest impact on the calculation of the Dow), Navistar was the cheapest and Exxon lay somewhere in between the two. However Exxon's market value was the greatest (at approximately $57.6 billion USD), whereas Merck had a smaller market value of $20.6 billion USD, and Navistar's market value was only $5.4 billion USD. In the example given, the DJIA underestates the performance of the market.

Interestingly Fisher (1927)[pages 419-420] notes that

It is remarkable that the very worst formulae have been used for stock market
indexes in spite of the fact that fuller data are available for such indexes than for commodity indexes or for any other sort of index, so that the very best formula could readily have been applied. In some cases, in fact, the indexes actually in use do not even constitute true indexes (an average of price relatives) at all.

It is conceivable that Fisher had the DJIA and other such indexes in mind when commenting on the class of stock market price indexes in existence in the late 1920's. Indeed the absence of share quantities from the calculation of the DJIA may well have sparked Fisher's criticisms of stock market indexes.

Another noteworthy feature of the Dow is its expression (in the business and finance literature) as a unilateral index, where the date \( t \) index is computed using only price information of the thirty component stocks at time \( t \), \( p^t = (p^t_1, \ldots, p^t_N)^T \) and the divisor \( d \). This contrasts with the Laspeyres bilateral index concept which compares the return performance of a portfolio over two dates (the base and comparison dates):

\[
R_L(p^0, p^1, q^0) = \frac{p^{1T}q^0}{p^{0T}q^0}.
\]  

The unilateral index \( R^t_{DJ} \), by itself, has no significance—it is only useful when compared to previous index values. Thus if

\[
R^t_{DJ} = \frac{\sum_{i=1}^{30} p^t_i}{d} \text{ and } R^{t-1}_{DJ} = \frac{\sum_{i=1}^{30} p^{t-1}_i}{d},
\]  

then their ratio\(^3\) becomes the DJIA expressed in bilateral form, or

\[
R^{t}_{DJ}(p^{t-1}, p^t) = \frac{\sum_{i=1}^{30} p^t_i}{\sum_{i=1}^{30} p^{t-1}_i} = \frac{1}{30} \frac{\sum_{i=1}^{30} p^t_i}{\sum_{i=1}^{30} p^{t-1}_i}.
\]  

Note that the index in equation 4.8 is a Dutot index—it is a ratio of arithmetic mean prices. Thus as a unilateral index the DJIA is a 'modified' arithmetic average, and as a bilateral index (assuming no divisor adjustments) it is a ratio of arithmetic mean prices.\(^4\)

\(^3\)It is assumed that there is no adjustment to the divisor here.

\(^4\)In fact the Dutot index is used in constructing elementary indexes (i.e. indexes at the lowest level of
4.1.3 Adjustments to the Dow Jones Industrial Average

The original method of calculation of the DJIA was elementary arithmetic (and still is today). When the DJIA first appeared in July 1884 it was an eleven-stock index, nine of which were railways, and was calculated by summing up the prices of the eleven stocks and dividing by eleven. Today, there are thirty stocks in the DJIA, but the divisor is not thirty. The reason for this is that the divisor has been changed over time to maintain historical continuity and comparability of the index. The principal reasons for changes in the divisor are stock splits, stock dividends, and changes in the constituent stocks which are used in computing the average. Consider the following example of a stock split, taken from Pierce (1994)[page 11]:

Such stock splits by companies included in the average would produce distortions if the new price was simply substituted for the old one. Here is an example: assume three stocks selling for $5, $10 and $15. Their average price is $10. Now the $15 stock is split three-for-one, which would make the new shares sell for $5. Nothing has happened to the value of an investment in these shares but the average price now is $6.67—an obvious distortion in comparison with the earlier average.

An adjustment is made to the index so that the 'average' remains at $10. This is done by adjusting the divisor so the index level remains the same before and after the stock split. Reducing the divisor from 3 to 2 achieves this and this new divisor is used in all subsequent calculations of the DJIA until the next adjustment is made. Most changes in the divisor are downward. The divisor of the DJIA fell below 1 in May 1986 when Merck & Company split two-for-one. This same divisor adjustment procedure is used for special dividends and aggregation) for CPI calculations. Prior to 1995 for example, the Dutot index was the bilateral formula used by Statistics Canada in computing its elementary indexes, after which time it was replaced by the geometric mean of price relatives. Since January 1995 the Dutot index has been used for the “owner's equivalent rent” portion of the housing index in the U.S. CPI (the “residential rent” portion of the index always used it).

The divisor in effect on December 17, 1998 was 0.24275214.
changes in the composition of the index. Cash dividend payments are not accounted for in the calculation of the DJIA. There are some exceptions to this rule—when a special dividend payment is made (as was made by Texaco on two occasions in 1989) or when a distribution of a share in another firm is made, the divisor is adjusted. Stock dividends of 10% or less are ignored. Furthermore, if a substitution or capital change does not impact the average by more than five index points, the divisor is not adjusted.6

Prior to 1928 adjustments for stock splits were made to the numerator: thus a stock which underwent a s-for-1 split was assigned a weight of s in subsequent index calculations. One feature of this technique (according to Dow Jones Indexes) was that the price movements of those stocks that had split tended to dominate the index. In light of this, the weighting technique was abandoned in favour of the current method of adjusting the divisor.

What rationale lies behind the current adjustment procedures, and what implicit assumptions underlie them? Furthermore, what properties do these procedures imply for the DJIA?

In the example shown in table 4.4 company 1 announces a two-for-one stock split (at date 1 to take effect at the beginning of the next trading period (next day))7 which means that

---

6Source: Dow Jones Indexes personnel.
7In table 4.4 and all subsequent tables a ‘→’ indicates the occurrence of a stock split. More precisely a ‘→’ at time t indicates the announcement of a stock split to take effect at the beginning of the next trading
shareholders receive an additional share for every share held. As a result of the stock split
the share prices are halved (ceteris paribus) and the number of shares outstanding doubles.
The value of the companies' market capitalizations and shareholders' investment holdings
have not changed. The divisor is adjusted to preserve comparability and continuity in the
average. In the example then

\[ R_{DJ}^0 = \frac{p_1^0 + p_2^0}{2}, \quad R_{DJ}^1 = \frac{p_1^1 + p_2^1}{2}, \]

and the new divisor \( \tilde{d} \) becomes

\[ \frac{p_1^1 + p_2^1}{\tilde{d}} = \frac{p_1^1 + p_2^1}{2}, \]

or

\[ \tilde{d} = \frac{p_1^1 + p_2^1}{2}. \]

\( \tilde{d} \) is used in all subsequent calculations so that \( R_{DJ}^2 \) becomes

\[ R_{DJ}^2 = \frac{p_1^2 + p_2^2}{\frac{p_1^1 + p_2^1}{2}} = R_{DJ}(p_1^0, p_1^1)R_{DJ}(p_1^1, p_2^2), \]

where \( R_{DJ}(p^t, p^{t+1}) = \frac{N}{N} \sum_{i=1}^{N} \frac{p_i^{t+1}}{p_i^t} \), \( \tilde{p}^1 \) is the post-split vector of prices corresponding to \( p^1 \),
and the series is normalized so that \( (p_1^0 + p_2^0) = 2 \). Thus the divisor adjustment is equivalent
to chaining the (Dutot) bilateral version of the DJIA. Recall from sections 3.2.2 and 3.3.3
in chapter 3 the interpretation of chaining: at a chain-link the portfolio is sold off and its
proceeds are reinvested instantaneously in the new portfolio. In the example in table 4.4 and
equations 4.9 to 4.12 the initial portfolio consists of one share of each of the two companies,
and the 'new' portfolio consists of one share of company two and one (not two) post-split
share(s) of company one. The DJIA can be regarded as a special case of the Laspeyres-type
period subsequent to \( t \). Adjustments to the divisor are done prior to trading on the following day.
Table 4.5: Stock splits in the DJIA

<table>
<thead>
<tr>
<th>Date</th>
<th>Company</th>
<th>Share Price (cents)</th>
<th>Shares outstanding</th>
<th>Market value (cents)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>$p_1^0 = 1$</td>
<td>$q_1^0 = 1$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>$p_2^0 = 3$</td>
<td>$q_2^0 = 1$</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>$p_1^1 = 2$</td>
<td>$q_1^1 = 1$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>$p_2^1 = 3$</td>
<td>$q_2^1 = 1$</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>$p_1^2 = 3$</td>
<td>$q_1^2 = 1$</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>$p_2^2 = 4$</td>
<td>$q_2^2 = 1$</td>
<td>4</td>
</tr>
</tbody>
</table>

stock index where $q^0$ is *always* maintained to be $q^0 = 1_{30}$, where $1_{30} = (1, \ldots, 1)^T$ is a vector of thirty ones.

Note that for stock splits in the case of the Laspeyres class of stock market indexes (equation 3.1, section 3.2, chapter 3), the increased number of shares times the lower price equals the old number of shares times the (previous) higher price. Thus the aggregate market capitalizations of the firms in the index and the *value* of the portfolio investment corresponding to the index have *not* changed. Individual investors whose portfolios match those of the market have not experienced a fall in their return due to the dilution in the share price. Indexes based on the Laspeyres concept do *not* require a divisor adjustment since the dilution in share price is offset by the corresponding increase in the quantity weights. The DJIA *does* need a divisor adjustment to prevent a drop in the average price. The implication then is that the prices of the stocks are not the *only* important input in calculating the DJIA, but rather that the value of the investment is important. As such then an *implicit* quantity adjustment is made—contrary to the absence from the DJIA calculations of explicit quantity weightings or market capitalization considerations.

Consider the example in table 4.5 which illustrates the calculation of a two stock index. An investor (or the market) purchases one share of each of two companies so that $q^0 = (q_1^0, q_2^0)^T = (1, 1)^T$. If this portfolio is held until the end of the second period then the
Laspeyres index series is given by $R_L^1 = 1, R_L^2 = \frac{5}{4}, R_L^3 = \frac{7}{4}$, indicating a net return of 75%.

This same result obtains (with appropriate normalizations) by calculating the DJIA (or equivalently the bilateral Dutot index). Suppose, however, that company two splits its stock 2-for-1 prior to trading in period 2. In this case the chained Laspeyres sequence is given by $R_L^0 = 1, R_L^1 = \frac{5}{4}, R_L^2 = \frac{7}{4}$ which is the same as before. This should not be surprising since the market values of the firms\(^8\) at each date are the same in both scenarios, so that the return performance is expected to be the same. The DJIA is given by $R_{DJ}^0 = 1, R_{DJ}^1 = \frac{5}{4}, R_{DJ}^2 = \frac{25}{14}$, indicating a net 'return' of approximately 79%. Note that given the method used in adjusting the divisor the DJIA cannot be interpreted as, nor is a measure of the rate of return on a portfolio initially consisting of $q^0 = (1, 1)^T$. The method of chaining results in a reinvestment in only one (not two) share(s) of the post-split company. The DJIA only coincides with the Laspeyres concept when $q^0 = k1_{30}, k > 0$, and when there are no stock splits or stock dividends—a highly unrealistic scenario.\(^9\) In the example in table 4.5 the DJIA 'overstates' the true return by approximately 4% over the evaluation period.

Suppose in table 4.4 the stock is split s-for-one (where previously $s = 2$). In addition suppose that there are $N$ firms whose share prices and quantities of shares in the portfolio at time $t = 0, 1, \ldots$ are $p_i^t$ and $q_i^t$ respectively. The divisor prior to any stock split is $d^t$ and the adjusted divisor is $\tilde{d}^t$.\(^10\) Using the same setup as in table 4.4 and equation 4.11 gives

$$
\tilde{d}^t = \frac{p_i^{t, s} + \sum_{i \neq 1}^N p_i^t}{p_i^{t, s} + \sum_{i \neq 1}^N p_i^t},
$$

which can be rewritten as

$$
\frac{1}{\tilde{d}^t} = \frac{1}{d^t} + \frac{1}{d^t} \left( \frac{1}{N} \left( \frac{p_1^t}{s} + \sum_{i \neq 1}^N p_i^t \right) - 1 \right).
$$

\(^8\)These are the market values of the firms if it is assumed initially that each firm has only one share outstanding.

\(^9\)The DJIA gives the 'correct' answer trivially in the case where $N = 1$.

\(^10\)Assume without loss of generality that firm one splits its stock $s$-for-one.
Note that $\frac{1}{d^t}$ is the new (i.e. adjusted) equal weight to be applied to each stock price in subsequent calculations of the DJIA in the sense that the impact of each stock $i$ on the next value of the unilateral DJIA is $\frac{p_{t+1}^i}{d^t}$. The interpretation of equation 4.14 is as follows: the new weight is equal to the previous weight plus some adjustment term. The term in parentheses on the right-hand-side of equation 4.14 compares the before-split average with the post-split average, i.e. the post-split stock price is simply substituted into the expression $\frac{1}{N} \sum_{i=1}^{N} p_i^t$ in the denominator. Since $\frac{p_{t}^s}{p_{t}^i} < p_{t}^i$, the expression in the parentheses is positive. This adjustment is scaled by the old ‘weight’ $\frac{1}{d^t}$ and is applied to all share prices. Essentially the asymmetric distortion in one price is distributed equally among all prices. This same procedure is used when component stocks are substituted in the index. Suppose firm $r$ replaces firm $s$ in this $N = 30$-stock index. Then as before

$$\frac{1}{d^t} = \frac{1}{d^t} + \frac{1}{d^t} \left( \frac{1}{N} \left( \frac{p_{t}^r + \sum_{i \neq r}^{30} p_i^t}{p_t^s + \sum_{i \neq s}^{30} p_i^t} - 1 \right) \right). \quad (4.15)$$

Whenever $p_{t}^r = p_{t}^s$, then no adjustment in the divisor is necessary. When $p_{t}^r > p_{t}^s$, the weights (i.e. inverse divisor) need to be adjusted downward and vice versa for the case when $p_{t}^r < p_{t}^s$. Once again quantity considerations have been completely sidestepped and price is considered the only important criterion in aggregating over firms.

Two further features of the DJIA which arise from its divisor adjustment procedure are demonstrated here.

**Timing of stock splits**

Consider now the case where two stock splits occur as in table 4.6. For this example it is seen that

$$R_{D,J}^0 = 2, R_{D,J}^1 = 2\frac{1}{3}, R_{D,J}^2 = 2\frac{1}{3}, \quad (4.16)$$

where the divisors $d^t$ for each date $t = 0, 1, 2$ are $d^0 = 2, d^1 = \frac{3}{2}, d^2 = \frac{15}{14}$. A Laspeyres fixed basket index is given by

64
<table>
<thead>
<tr>
<th>Date</th>
<th>Company</th>
<th>Share Price (cents)</th>
<th>Shares outstanding</th>
<th>Market value (cents)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2 → 1</td>
<td>1 → 2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1.5</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2 → 1</td>
<td>1 → 2</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1.5</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 4.6: Timing of stock splits in the DJIA

\[ R_L^0 = 2, R_L^1 = 2 \frac{1}{2}, R_L^2 = 2 \frac{1}{2}. \]  \hfill (4.17)

Suppose now that the magnitude of the stock splits is unchanged, and the price patterns are the same for both stocks. The only difference is the timing of the splits: both firms announce their splits during the first trading period (to take effect from date 1). Then the DJIA becomes

\[ R_{DJ}^0 = 2, R_{DJ}^1 = 2 \frac{1}{2}, R_{DJ}^2 = 2 \frac{1}{2}. \]  \hfill (4.18)

which differs from the example in table 4.6 where the stock splits occurred in successive periods, and now \( d^0 = 2, d^1 = d^2 = 1 \). For the Laspeyres index the same answer obtains as in equation 4.17. Thus the timing of the stock split does matter for the DJIA, since this will impact on the magnitude of the divisor used in the calculation of the DJIA in any given period. The DJIA is not invariant to the units in which stock prices are measured and thus fails an important commensurability property (i.e. sensitive to the units of measurement—see chapter 5 for more on this).\(^{11}\)

\(^{11}\)Stock splits represent a change in the units in which the shares of a company are traded and held.
A ‘reverse’ stock split

Consider now an example of a reverse stock split (consolidation). Using the information in table 4.7 yields the following:

\[ R_{DJ}^0 = 2, R_{DJ}^1 = 2 \cdot \frac{2}{3}, R_{DJ}^2 = 2 \cdot \frac{2}{15}, \]  
\( (4.19) \)

where \( d^0 = 2, d^1 = \frac{3}{2}, d^2 = \frac{15}{8} \). Despite the stock split reversal and the fact that all prices have returned to their initial levels it is not the case that \( R_{DJ}^0 = R_{DJ}^2 \). The change in the price of stock 2 during period one is given more weight in the index (due to its price-weighted feature) which then impacts upon any subsequent divisor values. In this case then the new divisor is not 2 which would be needed for \( R_{DJ}^2 = 2 \). A ‘stock split reversal property’ is failed. In the case of a Laspeyres index it is found that

\[ R_L^0 = 2, R_L^1 = 2 \cdot \frac{1}{2}, R_L^2 = 2. \]  
\( (4.20) \)

so that the index returns to its original value.\(^{12}\)

The DJIA divisor has declined over time, with the decline being checked only by the possible substitution of higher priced stocks for lower priced ones. The procedure also results in those stocks which split or issue stock dividends more frequently having a smaller impact on the performance of the unilateral average. If such stocks in fact outperform other stocks then the result is for the average to be lower than it might otherwise be.

\(^{12}\)This will generally not be the case when, for example the quantities of shares change for reasons other than stock splits (such as issuances of new shares). When the chaining principle is employed and prices and quantities return to their original values the index value may not return to its beginning value. When this does happen in standard price index theory the index is said to satisfy Walsh’s multiperiod identity test (see Diewert (1993a) and Walsh (1901)):

\[ P(p^0, p^1, q^0, q^1)P(p^1, p^2, q^1, q^2)P(p^2, p^0, q^2, q^0) = 1, \]  
\( (4.21) \)

where \( P(p^t, p^{t+1}, q^t, q^{t+1}), t = 0, 1 \) is a traditional bilateral price index. A frequent criticism of the chaining principle is its failure to satisfy equation 4.21.
Table 4.7: A consolidation in the DJIA

Stock splits entail a change in the size of the units in which ownership may be bought and sold—however the market capitalization of any firm is invariant under stock splits. Reasons often offered as plausible explanations for/characterizations of stock splits include the following:

- high priced equities are reduced in price to get the price down to a more favourable trading range;

- stock splits often reflect the growth of a company, where stock prices may rise again to their pre-split levels.

Under the second argument the proportionate influence of growth stocks on the DJIA will continue to be dampened, with the average failing to adequately reflect the importance of such growth stocks in the index. Equally, unsplit equities have an increasingly greater influence on the DJIA; in other words a statistical premium is placed on the failure to grow.

All stock splits in the DJIA result in a reduction in the divisor. If a particular stock splits when it is selling at an abnormally high level (due to favourable market news about its future potential profitability), the result is a divisor lower than would otherwise be needed to adjust for the stock split under current procedures. This inflates future values of the index relative to the case of a higher divisor. This observation merely reiterates the point raised
in table 4.6 that the timing of the split does affect the value of the DJIA by determining
the magnitude of the adjusted divisor. The downward pressure on the divisor caused by
splitting high priced stocks is partly offset by ignoring (small) stock dividends, which would
otherwise lower the value of the divisor.

The DJIA has been affected to an unknown and substantial (but unmeasurable) degree
by the procedure that has been adopted in adjusting for stock splits and stock dividends.
The preceding discussion indicates that the DJIA cannot be interpreted as a measure of the
rate of return on the ‘market’ portfolio composed exactly of the Dow constituents during
any given time horizon. Given the divisor adjustment process and the absence of a sensible
weighting system the DJIA is, at best, a crude indicator of whether stock prices are rising
or falling on average.

4.2 Other Indexes

As was mentioned at the outset of the chapter, the DJIA is an ‘atypical’ index, given its
unique and unusual calculation methodology. To put the DJIA in context and compare it
with other common stock market indexes in use today, three types of indexes are described
below. While several hundred stock indexes are currently published, differing primarily in
their portfolio composition, the three indexes described in this section are representative
of the computational formulae employed in a large majority of the world’s stock indexes.
Notwithstanding this, procedural guidelines for including/excluding certain companies from
indexes etc. vary widely from country to country and stock exchange to stock exchange—
these procedures are not considered in this thesis.

4.2.1 Standard and Poor’s Composite 500 (S & P 500)

Standard and Poor’s publishes several indexes, of which the best known is the Standard and
Poor’s 500 Composite (S & P 500) index. This index is computed using the stock prices of
500 companies listed on the New York Stock Exchange, the American Stock Exchange, and
the Over The Counter market. The basic bilateral formula used in computing the S & P 500 is a fixed basket Laspeyres index or

\[ R_{S&P}^1 = R_{S&P}(p^0, p^1, q^0) = \frac{p^1^T q^0}{p^0^T q^0} = \frac{\sum_{i=1}^{500} p_i^1 q_i^0}{\sum_{i=1}^{500} p_i^0 q_i^0}, \]  \hspace{1cm} (4.22)

where \( q^0 \) is the base date vector of quantities of shares outstanding of each of the 500 stocks in the portfolio. The base period value for the S & P 500 is the average of the weekly market capitalizations for the period 1941-1943, and this is normalized to \( R_{S&P}^0 = 10 \). The index over several periods is given by

\[ R_{S&P}^t = R_{S&P}(p^0, \ldots, p^t, q^0, \ldots, q^{t-1}) = \frac{\sum_{i=1}^{500} p_i^t q_i^{t-1}}{\sum_{i=1}^{500} p_i^0 q_i^0} \times 10, \]  \hspace{1cm} (4.23)

where \( \sum_{i=1}^{500} p_i^0 q_i^0 \) is the base period market capitalization adjusted\(^1\) for capitalization changes.\(^2\)

Until the first capital change in the index the quantity of shares outstanding remains fixed at \( q^0 \) and the index value at date \( t \) is given by

\[ R_{S&P}^t = \frac{p^t^T q^0}{p^0^T q^0} = \frac{\sum_{i=1}^{500} p_i^t q_i^0}{\sum_{i=1}^{500} p_i^0 q_i^0}. \]  \hspace{1cm} (4.24)

Note that this can be rewritten as

\[ R_{S&P}^t = \frac{\sum_{i=1}^{500} p_i^0 q_i^0 \left( \frac{p_i^t}{p_i^0} \right)}{\sum_{i=1}^{500} p_i^0 q_i^0} = \sum_{i=1}^{500} s_i^0 \left( \frac{p_i^t}{p_i^0} \right), \]

where \( s_i^0 = \frac{p_i^0 q_i^0}{\sum_{j=1}^{500} p_j^0 q_j^0} \), \( i = 1, \ldots, N \) is the value share of firm \( i \) in the total market capitalization at date 0, so that \( R_{S&P}^t \) can be interpreted as a weighted average of the individual stock gross returns for the components in the index. Let \( \Delta^t \) be the value of capital changes that occur

\(^{13}\)Notation: \( \tilde{x} \) denotes an adjusted (or modified) value of \( x \).

\(^{14}\)Examples of capitalization changes include consolidations and acquisitions, spin-offs and substitutions in the index.
at time $t$—these occur after the closing of trading on day $t - 1$ but prior to the calculation of the index on day $t$. At time $t - 1$ the index is given by

$$R_{S&;P}^{t-1} = \frac{p^{t-1^T}q^0}{p^{0^T}q^0} = \frac{p^{t-1^T}q^0}{d^{t-1}},$$  \hspace{1cm} (4.25)

where $d^{t-1}$ is the divisor for the index at time $t - 1$ (which, in this case is the same as $d^0$). Between any two dates the index should only change by movements in the prices of the component stocks. Capitalization changes should not change the level of the index—this is dealt with by adjusting the base period capitalization (the divisor in the index) so that the index level remains the same before and after any capitalization change. The adjusted divisor $\tilde{d}^{t-1} = \tilde{p}^{0^T}q^0$ is calculated as

$$\frac{p^{t-1^T}q^0}{p^{0^T}q^0} = \frac{p^{t-1^T}q^0 + \Delta^t}{p^{0^T}q^0},$$  \hspace{1cm} (4.26)

so that

$$\tilde{p}^{0^T}q^0 = p^{0^T}q^0 \left( \frac{p^{t-1^T}q^0 + \Delta^t}{p^{t-1^T}q^0} \right),$$  \hspace{1cm} (4.27)

or

$$\tilde{d}^{t-1} = d^{t-1} \left( \frac{m\text{cap}^{t-1}}{m\text{cap}^{t-1}} \right),$$  \hspace{1cm} (4.28)

so that the adjusted divisor is the product of the unadjusted divisor times the ratio of adjusted to regular (unadjusted) market capitalization ($m\text{cap}$) values. $d^t = \tilde{d}^{t-1}$ now becomes the relevant divisor for making date $t$ (and subsequent comparisons) until the next capitalization change takes place. Note that $p^{t-1^T}q^0 + \Delta^t$ can be rewritten as $\tilde{p}^{t-1^T}q^0 \equiv p^{t-1^T}q^{t-1}$ so that the new basket is now $q^{t-1}$ and $\tilde{p}^{t-1}$ are the adjusted prices for each of the stocks in the index.\textsuperscript{15} The bilateral index for making comparisons over period $t - 1$ and $t$ is given by

\textsuperscript{15}In the case of a component substitution $\tilde{p}^{t-1}$ is simply $p^{t-1}$ with the exiting stock price replaced by its substitute and $\tilde{q}^0$ is the vector $q^0$ with the corresponding quantity of shares outstanding of the exiting firm replaced by those of the replacement stock.
pricing out the 'new' basket $q^{t-1}$ and comparing it to $d^t$, the adjusted base period market capitalization so that

$$ R_{S&P}^t = \frac{p_\tau^T q^{t-1}}{d^t} $$

$$ = \frac{p_\tau^T q^{t-1}}{\bar{p}^{t-1} q^{t-1} \frac{q^0}{p_0^T q^0}} $$

$$ = \left( \frac{p_\tau^T q^0}{p_0^T q^0} \right) \left( \frac{\bar{p}^{t-1} q^{t-1}}{p^{t-1} q^{t-1}} \right) $$

$$ = R_{S&P}(p^0, p^{t-1}, q^0) R_{S&P}(\bar{p}^{t-1}, p^t, q^{t-1}) $$

so that the divisor adjustment process adopted by many financial houses who publish indexes is equivalent to the chaining procedure described earlier.\footnote{It is straightforward to show that before any capitalization change \( R(p^0, p^{t-1}, q^0) = \Pi_{i=1}^{t-1} R(p^{i-1}, p^i, q^0). \)} This is similar to the divisor adjustment technique used by Dow Jones and Company. Note however that, unlike for the DJIA, \emph{no} divisor adjustment is needed in the case of a stock split, since in this case $\Delta^t = 0$. In this sense then the basic bilateral index number formula is invariant to changes in the units of measurement. The published S & P 500 is \emph{not} a total return index, though Standard and Poor's do compute a companion total return index for the S & P 500, defined by option 2 in section 3.3.2, chapter 3.

### 4.2.2 Value Line Indexes

Value Line, Incorporated computes and publishes two major stock market price indexes—the Value Line Composite Geometric Index (VLG) and the Value Line Arithmetic Average (VLA). Both indexes are calculated using the closing stock prices of 1,600-1,700 firms—about 80% of these are listed on the New York Stock Exchange with the remainder traded on the American Stock Exchange or the Over The Counter market. \footnote{It is straightforward to show that before any capitalization change \( R(p^0, p^{t-1}, q^0) = \Pi_{i=1}^{t-1} R(p^{i-1}, p^i, q^0). \)}
Both Value Line indexes are computed using the same stocks—the only difference between them being the formula used in calculating the index. Let \( p^0 = (p^0_1, \ldots, p^0_N)^T \) be the \( N \)-vector of base date prices and \( p^1 = (p^1_1, \ldots, p^1_N)^T \) be the \( N \)-vector of current (comparison) date prices where \( N \) is the number of stocks. The basic bilateral indexes (i.e. for making price comparisons over two adjacent dates, a trading day apart) are as follows:

\[
VLG : R^1_{VLG} = R_{VLG}(p^0, p^1) = \prod_{i=1}^N \left( \frac{p^1_i}{p^0_i} \right)^{1/N},
\]

\[
VLA : R^1_{VLA} = R_{VLA}(p^0, p^1) = \frac{1}{N} \sum_{i=1}^N \left( \frac{p^1_i}{p^0_i} \right).
\]

The VLG is thus a geometric mean of individual stock gross returns (Jevons index) while the VLA is an arithmetic mean of individual stock gross returns (Carli index). Since an arithmetic average of \( N \) non-negative numbers \( (x_1, \ldots, x_N) \) is always at least as large as the corresponding geometric average

\[
\frac{1}{N} \sum_{i=1}^N x_i \geq \prod_{i=1}^N x_i^{1/N},
\]

it follows that the VLA is always at least as large as the VLG. Recall from section 3.3.1 in chapter 3 that the VLA corresponds to an equal-dollar weighting concept. The VLG, by contrast gauges the price performance of the median stock in the Value Line universe. The Value Line indexes over several periods are calculated by taking the cumulative product of the period-over-period bilateral indexes. This gives the chained sequences:

\[
R^t_x = R_x(p^0, p^1) \cdots R_x(p^{t-1}, p^t), x = VLG, VLA.
\]

Accounting for stock splits and stock dividends is straightforward—on the ex-dividend date the previous day’s closing price is set at the split/dividend adjusted level and the individual stock gross return is calculated accordingly. It is this adjusted individual stock gross return that is used in the computation of the bilateral index. Stock additions/deletions are handled in a similar manner—since the bilateral indexes are based on the concept of a
matched sample, additions/deletions are handled simply by adding or deleting the individual stock gross return from the calculation. Cash dividends are not accounted for in the indexes. No systematic periodicity of component changes is undertaken—it is done on an as needed basis. The interpretation of the VLA as an equal-dollar weighted index chained at the end of each trading day means that rebalancing to equal weights is done on a daily basis.

4.3 What return is it anyway?

Sections 4.1 and 4.2 taken together, make it clear that an investor with particular holdings in \( N \) stocks has several options for evaluating his/her portfolio in terms of the different index number formulae that can be used in addition to the different treatment of cash dividends. There are numerous ways of aggregating individual stock gross returns \( \frac{p_i^{t+1}}{p_i^t}, i = 1, \ldots, N \), both cross-sectionally and over time. However, there appears to be some confusion as to what each method of aggregation actually means. It should be apparent that different index number formulae cannot all be measuring the same thing.

The example from table 4.5 in section 4.1.3 indicates that in the presence of stock splits, the DJIA is not a rate of return index on a portfolio composed initially of \( q^0 = (1, 1)^T \). Given its bilateral index number formula and associated divisor adjustment, it is natural to ask what return the DJIA actually does measure. Alternatively, if the DJIA is to be interpreted as a return index what portfolio investment strategy does the DJIA calculation correspond to? The following attempts to elucidate the investment strategies behind the four types of indexes examined—this is done by identifying the particular portfolio in question, the basic aggregation formula and its interpretation, and the periodicity of rebalancing.

4.3.1 Investment strategies

Dow Jones Industrial Average

The bilateral index number formula used in calculating the DJIA is the Dutot index—i.e. the DJIA is a ratio of arithmetic mean stock prices or

\[
R_{DJ}^1 = R_{DJ}(p^0, p^1) = \frac{1}{30} \sum_{i=1}^{30} p_i^1.\]
Furthermore, the technique for adjusting the divisor results in a chained sequence of Dutot bilateral index numbers. The DJIA thus tracks the performance of the following portfolio investment strategy: an investor\(^{17}\) purchases one share of each of \(N\) (\(N = 30\) for the DJIA) firms in the index. The index tracks the performance of this portfolio over time.\(^{18}\) In the event of a stock split/stock dividend or component change the initial investment has a certain value—this is sold off and is invested instantaneously in the new portfolio which is then followed over time. The ‘new’ portfolio (i.e. after the split or substitution) consists of one share of the post-split firm or substitution firm respectively, in addition to one share of each of the other firms that were previously in the index. This new portfolio is then tracked over time until the next adjustment occurs. The DJIA does not take account of cash dividends and is thus not a total return index, in the sense discussed in section 3.3.2 in chapter 3.

Identifying the investment strategy associated with the DJIA permits a precise interpretation of what the DJIA measures. When journalists and financial analysts report that the “Dow is up” such a sweeping statement has a precise interpretation—what is actually meant is that the return on a portfolio consisting of one share of each of thirty firms has increased over time. In a less formal sense though, the index indicates whether stock prices in the Dow portfolio are rising or falling in general.

**Standard and Poor’s Composite 500**

The basic bilateral index of the S & P 500 is given by the Laspeyres index \(R_{S&P}(p^0, p^1, q^0) = \frac{p^1 \cdot q^0}{p^0 \cdot q^0} \). Thus an investor (the ‘market’) purchases (and holds) quantities \(q^0 = (q^0_1, \ldots, q^0_{500})^\top\)\(^{19}\) of each of the 500 stocks in the index. The performance of these stocks is tracked over time. Capital changes or component changes are accounted for by chaining the index (described in section 3.2.2, chapter 3 and again in section 4.2.1 of this chapter). Thus at a chain link,\(^{17}\)The terms ‘investor’ and ‘market’ are used interchangeably in what follows.\(^{18}\)It also tracks the performance of a portfolio composed \(q^0 = k1_{30}, k > 0.\)\(^{19}\)This \(q^0\) is the aggregation over all investors who have holdings in these particular 500 stocks that constitute the index.
the portfolio is sold off and the proceeds are invested in the new basket in proportion to the value share of each stock in the portfolio at the chain-link.

**Value Line Arithmetic**

The VLA index is calculated as an arithmetic mean of individual stock gross returns or

\[ R_{VLA}^1 = R_{VLA}(p^0, p^1) = \frac{1}{N} \sum_{i=1}^{N} \left( \frac{p_i^1}{p_i^0} \right) \]

where all share prices have been adjusted for stock splits. The index over \( T \) periods is calculated as

\[ R_{VLA}^T = \prod_{t=1}^{T} \left( \frac{1}{N} \sum_{i=1}^{N} \left( \frac{p_i^t}{p_i^{t-1}} \right) \right), \]

so that the period-over-period indexes are chained together. As discussed in section 3.3.1 and 3.3.4 in chapter 3 this index corresponds to an initial equal dollar investment in each of the \( N \) stocks in the portfolio. Thus if an initial total investment of $1 is made then \( \frac{1}{N} \) is invested in each of the \( N \) stocks in the portfolio—i.e. the implicit ‘quantity’ weights\(^{20}\) \( q_i^0, i = 1, \ldots, N \) are set so that \( p_i^0 q_i^0 = \frac{1}{N}, i = 1, \ldots, N \) so that \( q_i^0 = \frac{1}{N} p_i^0 \). It is seen that the VLA corresponds to the S & P 500 (a Laspeyres index) where \( q^0 = \frac{1}{N} (p^0)^{-1} \). At the end of the \((0,1)\) evaluation period the portfolio is sold off. The proceeds are now reinvested in the same portfolio, but the ‘quantity’ weights are reset so as to achieve an equal investment in each of the \( N \) stocks in the portfolio at the end of the \((0,1)\) evaluation period—i.e. the ‘new’ weights are \( q_i^1 = \frac{1}{N} (p_i^1)^{-1} \).\(^{21}\) This procedure is repeated on a daily basis. Thus at the end of each trading day the portfolio proceeds are sold off and reinvested to maintain an equal (i.e. equal at the beginning of each trading day) value representation in each of the \( N \) stocks in the portfolio. If stock \( i \) has performed well in the first period (so that \( p_i^1 \) is ‘high’), its ‘quantity weight’ (which is inverse price) is reduced in subsequent periods, and vice versa for stocks whose performance has been poor. The continual buying and selling of the portfolio may be called an ‘active’ strategy, whereas those underlying the S & P 500 and the DJIA are ‘passive’ or ‘buy-and-hold’ strategies. In other words, the portfolios underlying the S &

\(^{20}\)Note that explicit quantities of shares outstanding do not enter the calculation of the VLA formula.

\(^{21}\)The concept of rebalancing to equal weights is referred to as ‘fixed investment proportion maintenance’ by Evans (1968).
P 500 and the DJIA are held (without rebalancing) until a substitution, stock split, or other capital change occurs. This typically happens on a monthly (or longer) basis.

**Value Line Geometric**

The VLG index is calculated as a geometric mean of individual stock gross returns or

$$R_{VLG}^1 = R_{VLG}(p^0, p^1) = \prod_{t=1}^{N} \left( \frac{p_t^1}{p_t^0} \right)^{\frac{1}{N}}$$

where all stock prices have been adjusted for stock splits/dividends. The index over several periods is calculated as

$$R^T = \prod_{t=1}^{T} \left( \prod_{t=1}^{N} \left( \frac{p_t^T}{p_t^{t-1}} \right) \right)^{\frac{1}{T}}$$

so that the period-over-period indexes are chained together. Interestingly, this index does *not* correspond to any investment strategy at all! The VLG is thus *not* a return index on any investment strategy. However the VLG might be interpreted as some average of stock price inflation. One result relevant here derives from the relationship between geometric and arithmetic means—since arithmetic means are *at least* as large as geometric means, then the VLG is a *lower bound* on the cumulative return on a portfolio whose investment strategy corresponds to that associated with the VLA.

These results are summarized in table 4.8. As an empirical illustration of the differences in returns that may arise from using different index number formulae some calculations are shown below using data on the DJIA components.\(^{22}\)

### 4.3.2 DJIA recalculation

The DJIA has been criticized for its price-weighted nature and its method of adjusting the divisor (for examples see Lerner (1963), Stillman (1986)[chapter 5] and McCallum & van Zijl (1987)), yet numerical estimates of the magnitude of its 'error' (or *bias*) are not available in the finance literature. It is worth emphasizing that the term *bias* is a relative concept, i.e.

\(^{22}\)This data has been provided by Dow Jones Indexes.
<table>
<thead>
<tr>
<th>Index</th>
<th>Bilateral formula</th>
<th>Basket</th>
<th>Return (investment strategy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DJIA</td>
<td>$R_{DJIA} = \frac{1}{N} \sum_{i=1}^{N} \frac{p_i}{p_i^0} q_i^0$</td>
<td>$q^0 = k_{130}, k &gt; 0$</td>
<td>buy-and-hold $q^0 = k_{130}$</td>
</tr>
<tr>
<td>S &amp; P 500</td>
<td>$R_{S&amp;P500} = \frac{1}{N} \sum_{i=1}^{N} \frac{p_i}{p_i^0} q_i^0$</td>
<td>$q^0 = (q_1^0, \ldots, q_{500}^0)$</td>
<td>buy-and-hold $q^0 = (q_1^0, \ldots, q_{500}^0)$</td>
</tr>
<tr>
<td>VLA</td>
<td>$R_{VLA} = \frac{1}{N} \sum_{i=1}^{N} \frac{p_i}{p_i^0} \left(\frac{q_i}{p_i}\right)^{t}$</td>
<td>$q^0(p^0) = \frac{1}{N}(\bar{p}^0)^{-1}$</td>
<td>rebalance daily to $\bar{q}'(\bar{p}') = \frac{1}{N}(\bar{p}')^{-1}$</td>
</tr>
<tr>
<td>VLG</td>
<td>$R_{VLG} = \Pi_{i=1}^{N} \left(\frac{p_i}{p_i^0}\right)^{t}$</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Table 4.8: Index formulae and investment strategies

It is expressed relative to some underlying benchmark. The benchmark adopted here for evaluating the performance of the DJIA is the chained Laspeyres concept outlined previously. This is tantamount to evaluating the DJIA as a gross rate of return on a portfolio held by investors in the aggregate that is composed exactly of the thirty Dow stocks. This seems to be the most natural candidate for evaluating the performance of any stock market gross return index.

Given the dominant influence of large companies (such as those in the DJIA) in determining the index values of value-weighted indicators (such as the Laspeyres), the tendency of all stocks to move together and the degree of diversification implicit in the set of stocks comprising the DJIA, it is not surprising that, despite price-weighting and small sample size, movements in the average have been found to mirror closely the movements in other more broadly-based value-weighted indicators such as the S & P 500. Notwithstanding this, the observation that the DJIA and the S & P 500 move together is not a necessarily useful

---

23 For example, a statistical estimate $\hat{\beta}$ is said to be an unbiased estimate of $\beta$ when $E(\hat{\beta}) = \beta$, where $E$ denotes the expectations operator, and $\beta$ is the true population parameter. The difference between the two, $\beta - E(\hat{\beta})$ is the bias.

24 An early discussion of the problem of an adequate sample is found in Fisher (1927). He advocates a “total value criterion” where samples are chosen so that the product of their price and quantity indexes gives the true value index for the universe represented by the samples. Fisher reports startling results of poorly designed samples. The total value criterion is the analogue of the factor reversal test discussed in Fisher (1922).
comparison since each is based on a different basket of stocks and a different computation method. Furthermore, short term deviations do occur, with important implications for measuring return performance and benchmarking. For this reason a chained Laspeyres index using the Dow components is computed, facilitating a direct and meaningful comparison with the present computation of the DJIA.

Closing share prices and quantities of shares outstanding for the DJIA components from 1/1/80 to 3/31/98 are used in computing the Laspeyres version of the DJIA (hereafter $DJIA_L$). Its composition on a given day is matched exactly with that of the DJIA, and the basket is updated whenever a substitution is made in the index components. These basket changes are described in table 4.9. The Laspeyres index takes account of splits, dividends, rights offerings, issuances and repurchases. A total return index $DJIA_{L(TR)}$ is also computed where it is assumed that cash dividends are reinvested in proportion to the value share of each stock in the index on the ex-dividend date.$^{25}$

Figures 4.1 and 4.2 plot DJIA, $DJIA_L$ and $DJIA_{L(TR)}$, where all three indexes have been normalized to equal 100 on January 1, 1980. Both DJIA and $DJIA_L$ generally track one another over time. Table 4.10 lists the annual opening index values for each of the years 1980 through 1998 for each of the three indexes. The table indicates that the average annual growth rates for the DJIA, $DJIA_L$ and $DJIA_{L(TR)}$ are (respectively) 13.3%, 12.7% and 17.4%, so that on an annual basis, the DJIA overstates the gross return on its portfolio (comprising the shares outstanding of investors in the aggregate) by approximately 0.6% per annum between 1980 and 1998, indicating that higher priced stocks have risen faster than lower priced ones. Figure 4.3 shows an example where the DJIA and $DJIA_L$ can deviate significantly over a shorter time horizon. Note that all these indexes report nominal returns on the portfolio and no inflation adjustments have been made to allow for the eroding purchasing power of nominal gross returns in inflationary environments.

One other frequent criticism of the DJIA, other than its computation, is its small sample

---

$^{25}$Recall again that the DJIA does not take cash dividends into account and thus a return counterpart to the DJIA does not exist.
<table>
<thead>
<tr>
<th>Date</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>8/30/82</td>
<td>American Express Co. replaces Manville Corp.</td>
</tr>
<tr>
<td>1/4/84</td>
<td>“New” AT &amp; T Co. replaces “old” AT &amp; T Co. (the substitution was required because of the divestiture of the local operating companies of the Bell system)</td>
</tr>
<tr>
<td>10/30/85</td>
<td>Philip Morris Cos. and McDonald’s Corp. replace General Foods Corp. and American Brands Inc.</td>
</tr>
<tr>
<td>3/12/87</td>
<td>Coca Cola Co. and Boeing Co. replace Owens-Illinois Inc. and Inco Ltd.</td>
</tr>
<tr>
<td>12/16/88</td>
<td>Primerica (new) replaces Primerica Corp. (Primerica Corp. (old) merged into Commercial Credit Group Inc., which adopted Primerica Corp’s name)</td>
</tr>
</tbody>
</table>

Table 4.9: DJIA basket updates: 1980-1998

Figure 4.1: DJIA and $DJIA_L$, 1980-1998
Figure 4.2: DJIA and $DJIA_L$, 1980-1998

Figure 4.3: DJIA and $DJIA_L$, 1983-1987
<table>
<thead>
<tr>
<th>yr</th>
<th>S&amp; P</th>
<th>DJIA</th>
<th>DJIA_L</th>
<th>DJIA_L(TR)</th>
<th>DJIA_EW</th>
<th>DJIA_EW(TR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>81</td>
<td>125.77</td>
<td>114.93</td>
<td>115.27</td>
<td>123.14</td>
<td>114.79</td>
<td>121.62</td>
</tr>
<tr>
<td>82</td>
<td>113.54</td>
<td>104.32</td>
<td>106.27</td>
<td>121.49</td>
<td>102.77</td>
<td>115.11</td>
</tr>
<tr>
<td>83</td>
<td>128.16</td>
<td>122.45</td>
<td>128.31</td>
<td>157.28</td>
<td>119.09</td>
<td>141.62</td>
</tr>
<tr>
<td>84</td>
<td>152.79</td>
<td>150.06</td>
<td>152.3</td>
<td>196.78</td>
<td>154.46</td>
<td>191.17</td>
</tr>
<tr>
<td>85</td>
<td>154.94</td>
<td>144.45</td>
<td>156.66</td>
<td>213.34</td>
<td>150</td>
<td>194.31</td>
</tr>
<tr>
<td>86</td>
<td>195.74</td>
<td>184.4</td>
<td>193.47</td>
<td>276.32</td>
<td>189.48</td>
<td>256.50</td>
</tr>
<tr>
<td>87</td>
<td>224.36</td>
<td>226.05</td>
<td>209.18</td>
<td>311.13</td>
<td>217.43</td>
<td>306.67</td>
</tr>
<tr>
<td>88</td>
<td>228.91</td>
<td>231.16</td>
<td>212.19</td>
<td>326.86</td>
<td>230.84</td>
<td>339.67</td>
</tr>
<tr>
<td>89</td>
<td>257.29</td>
<td>258.55</td>
<td>235.94</td>
<td>378.85</td>
<td>271.44</td>
<td>414.93</td>
</tr>
<tr>
<td>90</td>
<td>327.40</td>
<td>328.25</td>
<td>291.47</td>
<td>486.25</td>
<td>331.75</td>
<td>529.56</td>
</tr>
<tr>
<td>91</td>
<td>305.93</td>
<td>314</td>
<td>286.62</td>
<td>496.84</td>
<td>299.27</td>
<td>498.61</td>
</tr>
<tr>
<td>92</td>
<td>386.41</td>
<td>377.8</td>
<td>350.22</td>
<td>628.36</td>
<td>375.06</td>
<td>639.90</td>
</tr>
<tr>
<td>93</td>
<td>403.66</td>
<td>393.58</td>
<td>354.81</td>
<td>657.4</td>
<td>412.24</td>
<td>713.48</td>
</tr>
<tr>
<td>94</td>
<td>431.20</td>
<td>447.89</td>
<td>384.33</td>
<td>733.46</td>
<td>487.76</td>
<td>843.86</td>
</tr>
<tr>
<td>95</td>
<td>425.49</td>
<td>457.17</td>
<td>388.47</td>
<td>763.25</td>
<td>491.4</td>
<td>877.53</td>
</tr>
<tr>
<td>96</td>
<td>570.62</td>
<td>610.1</td>
<td>533.91</td>
<td>1076.99</td>
<td>665.25</td>
<td>1207.89</td>
</tr>
<tr>
<td>97</td>
<td>686.25</td>
<td>768.8</td>
<td>667.28</td>
<td>1377.11</td>
<td>835.84</td>
<td>1532.41</td>
</tr>
<tr>
<td>98</td>
<td>899.05</td>
<td>942.87</td>
<td>859.38</td>
<td>1805.96</td>
<td>1034.28</td>
<td>1924.85</td>
</tr>
</tbody>
</table>

av growth(%) | 12.9 | 13.3 | 12.7 | 17.4 | 13.8 | 17.9 |

Table 4.10: Index values: S& P 500, DJIA, DJIA_L, DJIA_L(TR), DJIA_EW, DJIA_EW(TR)
size and the fact that its representation of American industry is not as broad-based as that of the S & P 500, the Wilshire Equity Indexes or the Russell Equity Indexes. The DJIA is computed using only a sample of thirty industrial stocks listed on the NYSE and tends to reflect ‘blue-chip’ stocks rather than small-cap stocks. To address this issue of representativity and sample selection figure 4.4 shows the evolution of $DJIA_L$ and S & P 500 over the 1980-1998 period. Note that this comparison is meaningful since both are based on precisely the same computation method—the only difference is the broader portfolio of 500 stocks tracked by the S & P 500 (incidentally the 30 DJIA stocks are a subset of this 500). Table 4.10 lists the opening index values for each of the years 1980 through 1998 for both the $DJIA_L$ and S & P 500. The S & P 500 grows at an annual rate of 12.9 % per annum compared to 12.7 % per annum for $DJIA_L$. This similarity in returns is not too surprising when one considers that a high percentage of the S & P 500’s market capitalization is dominated by the DJIA components.

Table 4.10 shows one further calculation based on the concept of the equally-weighted index whose basic bilateral formula is an arithmetic average of individual stock gross returns.
This index $DJIA_{EW}$ is rebalanced to equal weights at the end of each trading day, and its corresponding total return index is $DJIA_{EW(TR)}$. $DJIA_{EW}$ grows at an annual rate of 13.8% and $DJIA_{EW(TR)}$ grows at an annual average rate of 17.9% per annum. Figure 4.5 and table 4.10 indicate that of the three investment strategies (i.e. DJIA, $DJIA_L$ and $DJIA_{EW}$) daily rebalancing appears to give the highest return, since its index value is highest at the end of the 1980-1998 evaluation period.

Clearly the calculation of the returns is sensitive to the choice of index number formula, though each computation is not directly comparable since each is based on a different investment strategy. While the $DJIA_{EW}$ yields the highest return over the 1980-98 evaluation period, it differs from the S & P 500 and the DJIA since the former is based on a very active strategy based on daily rebalancing to equal weights.

---

26Transactions costs associated with this continual rebalancing have not been accounted for.
Chapter 5

An Axiomatic Approach to Stock Market Gross Return Indexes

Chapter 3 outlined some of the main considerations that arise in the construction of stock market gross return indexes, while chapter 4 demonstrated that different index number formulae are used in constructing the stock indexes in use today. Given the numerous ways of constructing stock market return indexes one is left to determine which to use, and to ascertain which one is 'best' in some sense. This chapter addresses the choice of appropriate formula for a stock market index, whose primary purpose is to measure the gross return on a given portfolio of stocks. With few exceptions (notably Balk (1988)) the theoretical underpinnings of stock indexes is not well developed. Nowhere in the literature is there any strong justification for using one formula over another, nor is there any systematic framework for differentiating among the various indexes in terms of the properties that they possess.

Section 5.1 describes the principal features of, and differences between the axiomatic and economic approaches to standard bilateral index number theory. Section 5.2 develops a system of axioms for stock market gross return indexes—i.e. a number of a priori desirable axioms (properties) are proposed for stock market gross return indexes. The relationships among these various axioms are explored. A minimal set of 'reasonable' and independent axioms is proposed, resulting in a specific functional form for \( R \).
5.1 Axiomatic and Economic Approaches to Index Numbers

Since the middle of the nineteenth century several price indexes have been proposed, the best known of which include the Laspeyres, Paasche and Fisher price indexes. However, it is not sufficient to propose and use a new index number formula without evidence or reason that such an index 'improves' on existing ones. The test or axiomatic approach to index numbers, initiated by Walsh (1901), Fisher (1911) and Fisher (1922) looks at an index number formula from the perspective of its mathematical properties with a view to developing a set of criteria for distinguishing among them, and addressing the question: which is the 'best' index?

The axiomatic approach to index numbers is well developed in the case of the measurement of consumer price inflation (see for example Eichhorn & Voeller (1983), Diewert (1992) and Balk (1995) for a discussion). This approach looks at the mathematical characteristics of the index. For example, if the current date prices increase, does the index increase? If current date prices increase by the same factor of proportionality, does the price index increase by that same factor of proportionality? If the units of measurement for the commodities are changed (for example, from litres to gallons for milk), is the price index affected? The axiomatic framework offers two general avenues for exploring the index number problem. First, different index number formulae can be evaluated and compared by determining whether or not they satisfy certain axioms. Second, and more common, it is natural to ask what properties a price index should possess, or in other words, on a priori grounds what are reasonable properties to expect of a price index? The objective of the axiomatic approach then is to specify a number of 'reasonable' tests or axioms which are sufficient to determine a unique functional form for a price index \( P \). Imposition of many (independent) axioms restricts the class of admissible index numbers for measurement purposes. Characterization results arise when the solution set to a given group of axioms is unique. Such a group of axioms is then said to characterize that particular index. On the other hand, specification of too many desirable properties may lead to inconsistency results, i.e. the solution set to a given group
of axioms is empty so that there does not exist any index number satisfying a given group of axioms.

The economic theory of index numbers, by way of contrast, is based on the definition of constrained optimization problems—the axiomatic approach makes no such assumptions about optimizing behaviour. As an example of the economic approach, let $e(p, \bar{u}) = \min_x \{ p^\top x | u(x) \geq \bar{u} \}$ be the expenditure function of a representative consumer which is dual to the utility function $u(x)$—it is the minimum expenditure required to achieve a reference level of utility $\bar{u}$ at prices $p$. The Konüs cost-of-living index (or price index) $P_K$ is defined as

$$P_K(p^0, p^1, \bar{u}) = \frac{e(p^1, \bar{u})}{e(p^0, \bar{u})},$$

that is, it is the ratio of minimum expenditure required to achieve a reference level of utility $\bar{u}$ at current prices $p^1$ relative to the minimum expenditure to achieve the same level of utility $\bar{u}$ at prices $p^0$. Clearly $P_K$ is based on the assumption of expenditure minimizing behaviour. Under a variety of assumptions on consumer preferences, the complicated and unobservable theoretical expression defined by equation 5.1 reduces to something simpler and observable. For example, Diewert (1976) shows that equation 5.1 reduces to the Fisher price index $P_F = \left( \left( \begin{array}{cc} p^1 & q^0 \\ p^0 & q^1 \end{array} \right) \right)^{\frac{1}{2}}$, when $e(p, \bar{u})$ is dual to the quadratic utility function $u(x) = (x^\top A x)^{\frac{1}{2}}$ where $A$ is a symmetric matrix of constants. $P_F$ is said to be exact for the quadratic utility function.$^1$ If $u(x)$ is linearly homogeneous then $P_K$ simplifies to $P_K = \frac{c(p^1)}{c(p^0)}$, a ratio of unit cost functions. Thus certain index number formulae are ‘consistent’ with particular representations of consumer preferences.

The economic approach to index numbers is not considered in this chapter—no assumptions are made concerning an investor’s risk/return preferences. Rather the focus is on the measurement of the return performance for a chosen and observable (ex post) portfolio of stocks $q^0 = (q^0_1, \ldots, q^0_N)^\top$. This chapter considers the properties of a bilateral stock market gross return index,$^2$ i.e. an index that measures the gross return (where the gross return $R$ is

---

$^1$Note that $P_F$ depends on observables and does not require estimation of the parameters of the $A$ matrix.

$^2$Since returns are typically generated by price movements of stocks a return index could be called a stock
the simple net rate of return \( r \) plus one, or \( R = 1 + r \) on a portfolio over one evaluation period, comprised of a base date (beginning of period, indexed by 0) and current (comparison) date (end of period, indexed by 1). Over the evaluation period prices and quantities may change. A price index (in this case a gross return index) summaries the price changes over the period in a single number, as does a quantity index for the quantity changes. Typically however the manner in which the price information is summarized or aggregated will depend on the quantity information and vice versa.

### 5.2 An axiomatic framework for stock indexes

Let \( p^t_i, q^t_i \) denote respectively the stock price and quantity of shares of each stock \( i = 1, \ldots, N \) in the portfolio at date \( t = 0, 1 \). Let \( p^t = (p^t_1, \ldots, p^t_N)^T \in \mathbb{R}^N_+, t = 0, 1 \) denote the corresponding \( N \)-vectors of prices and \( q^0 = (q^0_1, \ldots, q^0_N)^T \in \mathbb{R}^N_+ \) denote the fixed \( N \)-vector of quantities of shares, where \( N \) is the number of distinct stocks held in the portfolio. A gross return index is considered to be a positive function

\[
R : \mathbb{R}^N_+ \to \mathbb{R}_+
\]

satisfying certain axioms. Such an index measures the gross return on a portfolio of stocks over the \((0, 1)\) evaluation period. This departs slightly from the standard index number literature in which a price index is considered a function of \( 4N \) arguments, i.e. the two price vectors \((p^0, p^1)\) and the two quantity vectors \((q^0, q^1)\) so that \( P : \mathbb{R}^{4N}_+ \to \mathbb{R}_+, (p^0, p^1, q^0, q^1) \to P(p^0, p^1, q^0, q^1) \). The reason for this difference is as follows: \( q^0 \) represents some initial purchase of certain quantities of shares of firms at date 0. It is assumed that these share price index (consistent with the standard index number literature). However since the value of the portfolio over time is being considered here, the term gross return seems more appropriate. In what follows, when reference is made to a stock index, it is understood that a stock market gross return index is implied.
quantities are held (without rebalancing) until the end of the evaluation period (in this case date 1). If this is the case then there is no role for \( q^1 \) as a variable in \( R \). Since \( R(p^0, p^1, q^0) \) attempts to measure the gross return generated by the basket \( q^0 \) (which is held fixed by investors over the evaluation period) \( q^1 \) is not relevant to measuring this return. In fact the present indexes \( R \) can be regarded as special cases of those in standard bilateral index number theory where \( R(p^0, p^1, q^0) = \tilde{P}(p^0, p^1, q^0, q^0) \), where \( \tilde{P} \) is a traditional bilateral price index with \( q^0 = q^1 \). In the case of \( P : \mathbb{R}_{++}^{4N} \rightarrow \mathbb{R}_{++} \), the \( q^t, t = 0, 1 \) refer to quantities of consumption for period \( t \) which is of a certain duration. \( q^t \) is the consumption flow for a given period. This differs from the current interpretation of \( R(p^0, p^1, q^0) \) as measuring the gross return on a fixed portfolio \( (q^0) \) over the evaluation period, beginning at date 0 and ending at date 1. \( q^0 \) are the initial purchases or holdings of stocks, and these remain in the portfolio for the entire period.

Another important distinction between a standard price index (such as a cost-of-living index) and a stock market gross return index is the importance of value (price times quantity) in the latter. When measuring gross return it is value that is important (and the ‘gross return’ index is really a value index)—its decomposition into a price and quantity component is of no immediate interest or relevance. It is doubtful then if Fisher’s product test, which decomposes a value ratio into the product of a price aggregate (index) and a quantity aggregate (index) has an analogue for stock market indexes (see Diewert (1992)[page 214] for a discussion of the product test)—since quantities are held fixed over the evaluation period, a quantity index is trivially unity. As a further contrast, Eichhorn & Voeller (1976)[chapter 2], Dalén (1992) and Diewert (1995) consider price indexes of the form

\[
P : \mathbb{R}_{++}^{2N} \rightarrow \mathbb{R}_{++}
\]

\[
(p^0, p^1) \rightarrow P(p^0, p^1),
\]

where quantity information is often unavailable.

The axiomatic approach to index number theory is well developed in the \( P(p^0, p^1, q^0, q^1) \)
situation. These standard axioms are adapted to the present context, and are given new interpretations. In addition, some new axioms are proposed. A particular realization of $R$ is called a *gross return index number* or *value*. The axioms concern structural properties on the function $R$ which are assumed to hold for any particular realization of the arguments. The axioms proposed are generally multidimensional analogues to the one-stock index formula $R(p_0, p_1, q_0) = \frac{p_1^q}{p_0}$.

**A1: Monotonicity in current date prices:** $R(p_0, p_1^1, q_0^0) \geq R(p_0, p_1^0, q_0^0), \forall p_1^1 \geq p_1^0$. This axiom says that if no stock price is lower at the end of the evaluation period and at least one of the stock prices is higher then the index cannot be lower. Note that this is a *comparative static* result, i.e. it compares returns on two identical portfolios, differing only in at least one component of the current date stock price vector.

**A2: Linear homogeneity in current date prices:** $R(p_0, \lambda p_1^1, q_0^0) = \lambda R(p_0, p_1^1, q_0^0), \forall \lambda > 0$. This means that if all current date stock prices are multiplied by the same factor $\lambda$, then the new index should equal $\lambda$ times the old index. For example, if all current date stock prices are doubled, then the value of the portfolio is doubled. The index is thus homogeneous of degree one in current date prices.

**A3: Identity:** $R(p, p, q_0^0) = 1$. This axiom says that if the share prices remain constant over the evaluation period ($p_0 = p_1^1 = p$) the index should be unity, indicating that the gross return on the portfolio is unchanged (or equivalently that the *net* return between dates 0 and 1 is zero, $r = 0$).

**A4: Commensurability (stock split invariance):** $R(Dp_0^0, Dp_1^1, D^{-1}q_0^0) = R(p_0^0, p_1^1, q_0^0)$, where $D$ is a $N \times N$ diagonal matrix whose elements $\delta_{ii}, i = 1, \ldots, N \in \mathbb{R}_{++}$. This means that the index is invariant to changes in the units of measurement. From the point of view of

---

3Notation: $p_1^1 \geq p_1^0$ means that $p_1^i \geq p_1^0, i = 1, \ldots, N$ but that $p_1^i \neq p_1^0, i = 1, \ldots, N$.  

89
measuring investment return, it is irrelevant (for example) whether prices are expressed in
terms of 'prices per share' or 'prices per batch lot', and the corresponding quantity units
are 'shares' or 'batch lots' respectively. Stock splits, which are common occurrences in the
corporate world correspond to a change in the units in which shares are held, and do not affect the value of an investor's portfolio. For example, a stock which splits two-for-one results in the share price being halved but investors now receive two shares for each share previously held. Their portfolio values are unchanged by such an action.

A5: Dimensionality: \( R(\lambda p^0, \lambda p^1, q^0) = R(p^0, p^1, q^0), \forall \lambda > 0 \). The rationale for this axiom is that a dimensional change in the unit of currency in which each stock price is measured will not change the value of the index \( R \). The \( \lambda \) factor above is a currency conversion factor. Thus, if two otherwise identical stock exchanges differ only in the definition of their monetary unit, the value of their indexes should be the same. The index is then homogeneous of degree 0 in base and current date prices.

A6: Symmetric treatment of stocks: \( R(p^0, p^1, q^0) = R(p^0, p^1, q^{<>}), \) where \( p^*, t = 0, 1 \) and \( q^0 \) represent the same permutations of the elements of the price and quantity vectors. This axiom means that no stock can be singled out in the index in playing an asymmetric role in calculating the return. An example of an index which violates this axiom is \( R(p^0, p^1, q^0) = \sum_{i=1}^{N} (-1)^i \left( \frac{p^1_i}{p^0_i} \right) \).

Axioms A1 to A6 are independent—i.e. it is possible for any five of them to be satisfied without the sixth being satisfied. This is demonstrated in the appendix.

The following five properties A7 to A11 are immediate consequences of axioms A1 to A6. Since A7-A11 could legitimately have been proposed as desirable properties they are labeled A7 to A11, and are called axioms. Notwithstanding this, axioms A7 to A11 are dependent axioms and in this sense are redundant properties for a stock market gross return
index. Proofs of the dependency of A7 to A11 are found in the appendix.\(^4\)

**A7: Proportionality:** \(R(p^0, \lambda p^0, q^0) = \lambda, \forall \lambda > 0.\)

**A8: Mean value property:**

\[
\min_{i=1,\ldots,N} \left( \frac{p^1_i}{p^0_i} \right) \leq R(p^0, p^1, q^0) \leq \max_{i=1,\ldots,N} \left( \frac{p^1_i}{p^0_i} \right), i = 1, \ldots, N.
\]

This means that \(R\) is bounded above and below respectively by the largest and smallest of the \(N\) individual stock gross returns \(\frac{p^1_i}{p^0_i}, i = 1, \ldots, N.\)

**A9: Positivity:** \(R(p^0, p^1, q^0) > 0, \forall p^0, p^1 \gg 0_N.\) This means that the index is never non-positive or zero, so that net returns are bounded below by \(-1\) (a 100% loss).

**A10: Invariance to scaling of quantities:** \(R(p^0, p^1, \lambda q^0) = R(p^0, p^1, q^0), \forall \lambda > 0.\)

This axiom means that the index is homogeneous of degree zero in the components of the base date quantity vector. In other words, it is only the relative quantity holdings of shares in the portfolio that are relevant to determining the return.

**A11: Homogeneity of degree -1 in base date prices:** \(R(\lambda p^0, p^1, q^0) = \lambda^{-1} R(p^0, p^1, q^0), \lambda > 0.\)

This means that if all base date prices are multiplied by the same factor \(\lambda,\) then the new index should equal the old index divided by \(\lambda.\) For example, if all base date prices are doubled, then the gross return on the portfolio of stocks would only be half as large as before.

The above axioms (A1 to A11) are considered to be minimal tests or properties that

\(^4\)If A7 to A11 were proposed independently of A1 to A6, it would be possible to show that some of A1 to A6 were implied by A7 to A11, or that combinations from both groups imply some of the others. For example A5 and A11 imply A2; A4 and A10 imply A5. Readers may find some axioms more appealing or intuitive than others—for this reason, all axioms/properties (independent or otherwise) are listed.
A number of price indexes proposed in the literature satisfy the above axioms and these axioms are most often associated with the measurement of price changes for a basket of commodities. Further axioms are proposed below which relate to specific features of holding stocks.

**A12:** *Cash dividend:* \( R(p^0, p^1 + d, q^0) = R(p^0, p^1, q^0) + R(p^0, d, q^0) \), where \( d = (d_1, \ldots, d_N)^T \) is a vector of (non-negative) *cash* dividends. This axiom says that if some non-negative constant \( d_i \) is added (the cash dividend for each stock \( i \)) to each date 1 price \( p^1_i, i = 1, \ldots, N \), thus augmenting the return to that stock, then the resulting index is equal to the original index (which measures the *price* (capital) dimension of return) plus the index that results by replacing the vector of current date prices with the vector of cash dividends (which measures the *income* dimension of return).\(^6\) This axiom implies that the stock market index is weakly additive, so that the overall return (total return) can be decomposed into a ‘pure’ capital gains and ‘pure’ income gains component.\(^7\) Note that if \( N = 1 \), the double inequality in \( A8 \) implies that \( R(p^0, p^1, q^0) = \frac{p^1}{p^0} \), and

\[
R(p^0, p^1 + d, q^0) = \frac{p^1 + d_1}{p^0} = \frac{p^1}{p^0} + \frac{d_1}{p^0},
\]

which satisfies the cash dividend axiom.

Consider an index number formula \( R_N(p^0, p^1, q^0) \) that uses positive price and quantity information on \( N \) stocks. Consider now the *same* functional form calculated using only the first \( N - 1 \) stocks \( R_{N-1}(p^0, p^1, q^0) \).

\(^5\)Strictly speaking, \( A1 \) to \( A6 \) are considered minimal requirements since \( A7 \) to \( A11 \) are direct consequences of \( A1\)–\( A6 \).

\(^6\)It is assumed that date 1 is the *ex-dividend* date for all \( N \) stocks—i.e. each dividend payment corresponds to the income gains for the entire \((0, 1)\) evaluation period.

\(^7\)In general, a total return index (as in section 3.3.2 in chapter 3) will be difficult to decompose into ‘pure’ capital and income gains components. The reason for this is that dividends, when reinvested, generate their own returns, depending on the price movements of stocks and the receipt of further dividends. Because of this continual interaction between prices and dividends isolating a pure dividend effect is almost impossible.
A13: *Irrelevance of tiny holdings:*

\[
\lim_{q_N^0 \to 0} R_N(p^0, p_1^0, q^0) = R_{N-1}(p_1^0, \ldots, p_{N-1}^0, p_1^1, \ldots, p_{N-1}^1, q_1^0, \ldots, q_{N-1}^0).
\]

This axiom says that the limiting gross return on a portfolio of \(N\) stocks as the holding of stock \(N\) in the portfolio is diminished is equal to the gross return obtained if stock \(N\) is deleted from the calculation and the gross return is computed for the basket consisting of the remaining \(N-1\) stocks. Roughly speaking then, this axiom means that the existence of a diminishing holding should not materially change the gross return (to an investor) on the \(N\)-stock portfolio as measured by the other \(N-1\) stocks in the portfolio. Note that in order for this axiom to make sense the index formula must satisfy axiom A6. This axiom is the analogue to Dievert’s *irrelevance of tiny countries test* (Dievert 1993b)[page 314].

A14: *Merging of firms:*

Suppose that between the base and current dates firms \(N-1\) and \(N\) merge to become firm \(N-1\) so that the number of firms in the index changes from \(N\) to \(N-1\). The merger takes the form of a one share of firm \(N\) equals \(k\) shares of the newly merged firm \((N-1)\) and firm \(N-1\) gets one share of the ‘new’ \(N-1\) firm. Then an appropriate axiom for this merger is:

\[
R_N(p^0, p_1^1, \ldots, p_{N-1}^1, \underbrace{kp_{N-1}^1}_{\text{imputed value}}, q^0) =
\]

\[
R_{N-1}(p_1^0, \ldots, p_{N-2}^0, \underbrace{p_{N-1}^0 q_{N-1}^0 + p_N^0 q_N^0}_{\text{imputed value}} / q_{N-1}^0 + kq_N^0),
\]

where \(p_{N-1}^1\) is the date 1 stock price for the merged firm and \(q_{N-1}^0 + kq_N^0\) is the quantity of shares associated with the newly merged firm. Note that the \(N\)-stock index function \(R_N\) is a theoretical construct (i.e. the index if a merger did not take place), since there are no longer \(N\) stocks in the portfolio. However, when making bilateral comparisons \(2N\) prices and \(N\) quantities are needed (for \(R_N\)). The ‘imputed’ current date price pertaining to firm
$N$ (the 'disappearing' firm) is given by $p^1_N = kp^1_{N-1}$, i.e. $\frac{p^1_N}{p^1_{N-1}} = k$, where $k$ is the share conversion factor. What are the corresponding price and quantities at the base date for the newly merged firm (in $R_{N-1}$)? It seems appropriate that the relevant price for comparative purposes is the ‘average’ or ‘unit value’ price over the two merging firms or $p^0_{N-1}q^0_{N-1} + p^0_Nq^0_N$.

Given this setup then the base date market capitalization for the merged firm $p^0_m q^0_m = p^0_{N-1}q^0_{N-1} + p^0_Nq^0_N$ is the sum of the base date market capitalizations of the firms forming the merger, so that the merger is value preserving.

To illustrate the essence of this axiom consider the following example. Two firms have post-merger prices $p^0_1 = 2, p^0_2 = 1$ and quantities of shares outstanding $q^0_1 = 200, q^0_2 = 100$. The market capitalizations of the firms are (respectively) 400 and 100 so that firm one has a market value four times as large as that of firm two. The share price of firm one is twice that of firm two so that the share conversion factor is 2 ($k = 2 = \frac{p^0_1}{p^0_2}$ above). In other words, firm one will have a 400 stock stake in the newly merged firm and firm two will have only a 100 stock stake. Since the two firms now share the 'same' stock price (i.e. the stock price of the newly merged firm), this stock division ensures that the market capitalization contribution of each ‘component’ of the merger is preserved at the pre-merger ratios. At the instance of the merger the stock price is 1 (since this equates the pre and post-merger market capitalizations) though typically the price will rise subsequent to the merger.

This axiom formalizes the notion that the percentage value contribution to market capitalization and portfolio value of each merging firm is preserved before and after the merger. In other words, if a merged firm is arbitrarily split into its two components in such a way that the percentage value contribution of each were preserved then the returns computed using both methods (i.e. $R_N$ and $R_{N-1}$) should be identical.\footnote{More complex mergers involving both share and cash swaps are not considered here.}
any analogue in the standard price index literature.

**A15: Transitivity:** The transitivity test relates the gross return index for some given period to the gross return indexes for certain subperiods contained within the period, or

\[
R(p^0, p^1, q^0)R(p^1, p^2, q^0) = R(p^0, p^2, q^0).
\]  

(5.3)

Note that equation 5.3 corresponds to a constant holding strategy (of \(q^0\)) over the three dates 0, 1, 2 comprising the two periods. Then the axiom says that the overall gross return for the two periods is equal to the product of the two single period returns. The equally-weighted index (an arithmetic average of individual stock gross returns) does not satisfy transitivity. Another version of the test is strong transitivity:

**A15': Strong transitivity:**

\[
R(p^0, p^1, q^0)R(p^1, p^2, q^1) = R(p^0, p^2, q^0),
\]  

(5.4)

where the basket of stocks for the second period index is now \(q^1 \neq q^0\). Note that equation 5.4 implies the following

\[
R(p^0, p^1, q^0) = \frac{R(p^0, p^1, q)}{R(p, p^0, q)},
\]  

(5.5)

so that \(R\) must be independent of quantities (since the RHS of 5.5 is independent of \(q^0\) and the LHS of 5.5 is independent of \(q\)), implying the absence of a sensible weighting scheme (other than by prices). This implied functional form for \(R\) seems undesirable.

This observation resembles similar results in standard price index theory. In fact Eichhorn & Voeller (1983) [pages 439-440] note that “... there simply does not exist a function \(P\) which satisfies our four basic axioms for price indexes/purchasing power parities and the Circular Test, simultaneously, and which enables adequately changing weights.” The four basic axioms that Eichhorn and Voeller propose are monotonicity, proportionality, dimensionality and commensurability. The “changing weights” refers to the fact that the quantities may
change over the periods, i.e. \( q^1 \neq q^0 \). Theorem 5.13 in Eichhorn & Voeller (1983)\[pages 441-442\] shows that the monotonicity, proportionality, circularity (transitivity), and another condition akin to proportionality characterize the Cobb-Douglas index—an index that does not depend on quantities. They argue that this result follows from the imposition of circularity, analogous to strong transitivity in equation 5.4 implying independence of quantities. Eichhorn and Voeller’s result is similar to Theorem 11 in Balk (1995)\[pages 81-82\] which states that monotonicity, linear homogeneity, identity, commensurability and circularity characterize the Cobb-Douglas index. The original characterization result is due to Funke, Hacker & Voeller (1979).

If the share quantities in the portfolio are unchanged (as in equation 5.3) the imposition of transitivity seems natural. Strong transitivity implies independence of quantities—this is one of the reasons why Funke et al. (1979) proposed to discard the circular test which cannot be met with changing quantity weights.

Typically, in the finance literature, when measuring returns it is always assumed that some version (whether explicit or implicit) of transitivity holds (as in the calculation of time-weighted rates of return). Thus if \( r^t \) is the net return on some portfolio for period \( t = 1, \ldots, T \) then the net return for the entire holding period \( r^{hp} \) (comprised of the \( T \) periods) is calculated as:

\[
(1 + r^1)(1 + r^2) \cdots (1 + r^T) = (1 + r^{hp}). \tag{5.6}
\]

Axiom \textbf{A15} imposes restrictions on the functional form that the stock market gross return index can assume. Define \( p^0, q^0 \) to be the \( N \times N \) diagonal matrices whose diagonal elements are the components of the vectors \( p^0 \) and \( q^0 \) respectively.

**Proposition 5.1** If \( R \) satisfies axioms \textbf{A4} and \textbf{A15} then there exists a function \( f \) such that \( R(p^0, p^1, q^0) = \frac{f(q^0 p^1)}{f(q^0 p^1)} \). Conversely any function \( R(p^0, p^1, q^0) = \frac{f(q^0 p^1)}{f(q^0 p^1)} \) satisfies axioms \textbf{A4} and \textbf{A15}.

**Proof:** see appendix.
The imposition of further axioms implies a specific functional form type for $f$ in proposition 5.1.

**Proposition 5.2** If $R$ satisfies axioms A1, A2, A3, A4, A6 and A15 then there exists a function $f$ such that $R(p^0, p^1, q^0) = \frac{f(q^0 p)}{f(q^0 p)}$. Furthermore, $f$ is a homogeneous symmetric mean, i.e. it is an increasing, linearly homogeneous and symmetric function of the individual market capitalizations $q_i p_i, i = 1, \ldots, N$ that satisfies the following mean value property $f(\lambda x) = \lambda$.

**Proof:** see appendix.

Thus if $R$ satisfies axioms A1, A2, A3, A4, A6 and A15 then the stock market gross return index $R(p^0, p^1, q^0)$ is a ratio of homogeneous symmetric means where the same functional form appears in the numerator and the denominator. While the result of proposition 5.2 restricts the class of stock market gross return indexes to ratios of homogeneous symmetric means, it is not a useful one in the sense that the class of homogeneous symmetric means is large. For example the mean of order $r$, $M^r(x)$ of $N$ positive numbers $x = (x_1, \ldots, x_N)^T$, is a homogeneous symmetric mean defined as

$$M^r(x_1, \ldots, x_N) = \begin{cases} \left(\sum_{i=1}^{N} \frac{1}{N} x_i^r\right)^{\frac{1}{r}} & \text{if } r \neq 0 \\ \prod_{i=1}^{N} x_i^\frac{1}{N} & \text{otherwise.} \end{cases} \tag{5.7}$$

Note that $M^1(x)$ is the arithmetic mean of $x$ and $M^0(x)$ is the geometric mean of $x$, and there are numerous other symmetric means, each corresponding to a different value of $r$.

In order to further restrict the class of admissible $f$ for the purpose of measuring the return on some portfolio, the imposition of additional axioms is required. It turns out that adding axiom A12 (the cash dividend axiom) results in a characterization for stock market gross return indexes.

**Proposition 5.3** Suppose that $R(p^0, p^1, q^0)$ satisfies axioms A1, A2, A3, A4, A6, A12 and A15. Then $R(p^0, p^1, q^0) = \frac{p^1 T p^0}{p^0 T q^0}$, so that the stock market index is value-weighted. Conversely, $R(p^0, p^1, q^0) = \frac{p^1 T p^0}{p^0 T q^0}$ satisfies axioms A1, A2, A3, A4, A6, A12 and A15.
Proof: see appendix.

Thus satisfaction of the independent axioms A1, A2, A3, A4, A6, A12 and A15 imply that the functional form for the index must be Laspeyres. Conversely the Laspeyres index satisfies axioms A1, A2, A3, A4, A6, A12 and A15. It is said that this subset of axioms characterizes the Laspeyres index.

Corollary 5.1 \( R(p^0, p^1, q^0) = \frac{p^0 + q^0}{p^0 + q^0} \) satisfies all the axioms A1 to A15.

If A15 is replaced by A15' a characterization of the geometric mean of individual stock gross returns is obtained.

Proposition 5.4 Suppose that \( R(p^0, p^1, q^0) \) satisfies axioms A1, A2, A3, A4, A6, A12 and A15'. Then \( R(p^0, p^1, q^0) = \Pi_{i=1}^{N} \left( \frac{p_i^{1}}{p_i^{0}} \right) \). Conversely, \( R(p^0, p^1, q^0) = \Pi_{i=1}^{N} \left( \frac{p_i^{1}}{p_i^{0}} \right) \) satisfies axioms A1, A2, A3, A4, A6, A12 and A15'.

Proof: see appendix.

Note that proposition 5.4 is the stock index analogue to the results in Eichhorn & Voeller (1983)[pages 441-442] and Balk (1995)[pages 81-82].

A direct consequence of weak transitivity is time reversal, a property that has been proposed in the axiomatic approach to standard price indexes.

A16: Time reversal:

\[
R(p^0, p^1, q^0) = \frac{1}{R(p^1, p^0, q^0)}.  \tag{5.8}
\]

This axiom states that it does not matter in which direction the return on a portfolio between two dates is measured. Thus the index number for date 1 relative to date 0 is the reciprocal of the index number for date 0 relative to date 1. Note that equation 5.8 can be rewritten as
\[ R(p^0, p^1, q^0)R(p^1, p^0, q^0) = R(p^0, p^0, q^0) = 1, \] (5.9)

which follows from transitivity and axiom A3. Thus time reversal is a direct consequence of transitivity and the identity axiom. It says that the product of the index number going from 0 to 1 and the index number going from 1 back to 0 should be unity, indicating a net return of zero over the overall evaluation period (since the prices at the end points of the overall evaluation period (two periods long) are all equal). The time reversal test thus treats time in a symmetric manner so that if the roles of the date 0 and 1 variables are interchanged in \( R \) then the new index is simply the reciprocal of the old index. Kravis, Kenessey, Heston & Summers (1975)[page 47], in commenting on the usefulness of time reversal in the context of making inter-country comparisons note that

In many applications the utility of the results would be greatly diminished if, for example, it were necessary to provide two estimates of the comparison between the countries depending on which was taken as the denominator country in the ratio between them.

In this context the ‘time reversal’ is in fact ‘country reversal’.

**A17:** Limiting dominant stock:

\[ \lim_{q^0_i \to \infty} R(p^0, p^1, q^0) = \frac{p^1_i}{p^0_i}. \] (5.10)

The intuition behind this axiom is as follows. The ratio \( \frac{p^1_i}{p^0_i} \) is the gross return on holding a portfolio consisting of the single stock \( i \). The \( q^0_i, i = 1, \ldots, N \) are the holdings of the stocks in the portfolio. As \( q^0_i \) increases, so that the holding of stock \( i \) dominates the portfolio, then the gross return on the portfolio should tend to the gross return on the individual stock. This is a kind of ‘dual’ to axiom 13 where \( q^0_i \to 0 \); now \( q^0_i \to \infty \). However they are not equivalent: for example \( R(p^0, p^1, q^0) = \Pi_{i=1}^{N} \left( \frac{p^1_i}{p^0_i} \right)^{(1+q^0_i)} \) satisfies A17 but not A13.
Other important index number axioms relate to \textit{consistency in aggregation}. These axioms are important when an economic aggregate is, or can be partitioned into a number of sub-aggregates. For example the Dow Jones Composite (65) Index is subdivided into the Dow Jones Industrial, Transportation and Utilities Averages. In Canada, the Consumer Price Index \textit{All-Items} aggregate can be partitioned into several sub-aggregates including \textit{Food}, \textit{Clothing and Footwear}, and \textit{Health and Personal Care}.

The preceding axioms have concerned aggregate portfolios consisting of \( N \) stocks on some domestic stock exchange. Typically this aggregate can be divided into sub-aggregates such as for example, ones based on industrial classification or size of market capitalization. Suppose \( A \) is some aggregate set that can be partitioned into \( M \) mutually exclusive subaggregate sets \( A_m \) such that \( A = \bigcup_{m=1}^{M} A_m \) with \( A_i \cap A_j = \emptyset, i, j = 1, \ldots, M, i \neq j \). Excluding \( A_{13}, A_{14} \) and \( A_{17} \) which consider the likes of mergers and the behaviour of 'outlier' stocks, the indexes considered to this point have been of the form

\[
R_N : \mathbb{R}_{++}^N \rightarrow \mathbb{R}_{++}
\]

\[
(p^0, p^1, q^0) \rightarrow R_N(p^0, p^1, q^0),
\]

defined over the \textit{entire} portfolio of \( N \) stocks (the universe of stocks). Consider now an index \( R_{N_m} \) defined on the subaggregate \( A_m \) where \( |A_m| = N_m \), so that each \( A_m \) contains \( N_m \) stocks, \( m = 1, \ldots, M \). Note then that \( N = \sum_{m=1}^{M} N_m \). An investor thus holds shares of only \textit{some} of the firms that make up some broader market portfolio. This results in a series of \( M \) gross return indexes (one for each subaggregate) of the form

\[
R_{N_m} : \mathbb{R}_{++}^{N_m} \rightarrow \mathbb{R}_{++}
\]

\[
(p^{0N_m}, p^{1N_m}, q^{0N_m}) \rightarrow R_{N_m}(p^{0N_m}, p^{1N_m}, q^{0N_m}),
\]

where \( p^{tN_m}, t = 0, 1 \) and \( q^{0N_m} \) are the subvectors of \( p^t, t = 0, 1 \) and \( q^0 \) of dimension \( N_m \) corresponding to the subaggregates \( A_m, m = 1, \ldots, M \).
Consistency in aggregation results pertain to uncovering relationships between the aggregate index and its constituent subaggregate indexes. In particular it is of interest to know whether the index calculated in two steps (i.e. the aggregate index calculated from its subaggregate indexes) gives the same result as that from the direct (one-stage) computation. Of further interest is whether the functional form of the one-stage computation (i.e. the aggregate index) is the same as that of the subaggregate indexes (which is the first stage in the two-stage computation). Indeed the functional form of the second stage of the two-stage computation may have the same functional form as the other two stages. More generally then an index is said to be consistent in aggregation if the value of the index as calculated in two stages necessarily coincides with the value of the index when calculated directly.

Recall axiom A4 which states that \( R(Dp^0, Dp^1, D^{-1} q^0) = R(p^0, p^1, q^0) \) where \( D \) is a \( N \times N \) diagonal matrix. Setting \( D = (p^0)^{-1} \) yields the following reparameterization of the stock market gross return index \( R(p^0, p^1, q^0) = R(1_N, \frac{p_1^1}{p_1^0}, \ldots, \frac{p_N^1}{p_N^0}, q_1^0, \ldots, q_N^0) = R(1_N, \frac{p_1^0}{p_1^0}, p^0 q^0) \), where \( \frac{p_i^1}{p_i^0} = (\frac{p_1^1}{p_1^0}, \ldots, \frac{p_N^1}{p_N^0})^T \) and \( p^0 q^0 = (p_1^0 q_1^0, \ldots, p_N^0 q_N^0)^T \), so that the index can be written as a function depending on the \( N \) individual stock gross returns and the \( N \) base date market capitalizations. Below is a formal statement of the consistency in aggregation axiom.

**A18: Consistency in aggregation:** The following relationship holds between the aggregate index \( R_N(p^0, p^1, q^0) \) and its subaggregate indexes \( R_{N_m}(p_{N_m}^0, p_{N_m}^1, q_{N_m}^0), m = 1, \ldots, M, \)

\[
R_N(1_N, \frac{p_1^1}{p_1^0}, \ldots, \frac{p_N^1}{p_N^0}, q_1^0, \ldots, q_N^0) = R_M(1_M, R_{N_1}, \ldots, R_{N_M}, \sum_{i \in A_1} p_i^0 q_i^0, \ldots, \sum_{i \in A_M} p_i^0 q_i^0) \tag{5.11}
\]

where \( R_N, R_M, R_{N_m}, m = 1, \ldots, M \) all have the same formula. This axiom states formally what has been said in the previous paragraph about the equivalence of two-stage and one-stage aggregation. Note that each subindex \( R_{N_m} \) in equation 5.11 is treated like a gross return for a single stock in the second stage of the two-stage aggregation. This aggregation property is important when subaggregates are combined into larger aggregates so that no inconsistencies arise between the index numbers for the parts (such as the Dow Jones Indus-
trial, Transportation, and Utilities Averages) and those for the whole (such as the Dow Jones Composite Average). It is perhaps for this reason that the Fisher price and quantity indexes (which fail this test) have not been used in national income accounting. An earlier informal statement of this test is given by Montgomery (1937) [page 65] who notes that consistency in aggregation means that an index "may be calculated in a series of stages without affecting the ultimate result."

The requirement of consistency in aggregation makes sense in the context of measuring the return on a portfolio over time. Section 4.3 of chapter 4 noted that different index number formulae are consistent with different types of investment strategies. It follows that if different formulae are used for one-stage and two-stage aggregation this means that each represents a potentially different investment strategy. This has two implications. First, it means that an investor is not consistently following a particular investment strategy. Second, it opens up an arbitrage opportunity. Thus if subindexes result in different measured returns they may be combined (i.e. aggregated) using a different formula in a manner to yield a return higher than that of the one-stage index. In essence, the imposition of consistency in aggregation disallows arbitrage opportunities. Indeed returns on subportfolios can be regarded as returns on different investment vehicles, and may be interpreted as returns on individual stocks, where these 'individual' stocks may be thought of as composite stocks. It seems natural that such composite stock returns should be aggregated in the same way as the individual stock gross returns in the overall portfolio.

Note that the Laspeyres index is consistent in aggregation. To see this note the following

\[ R_N(p^0, p^1, q^0) = \frac{\sum_{i=1}^{N} p_i^1 q_i^0}{\sum_{i=1}^{N} p_i^0 q_i^0} \]

\[ = \sum_{i=1}^{N} s_i^0 \left( \frac{p_i^1}{p_i^0} \right), \]

where \( s_i^0 = \frac{p_i^0 q_i^0}{\sum_{j=1}^{N} p_j^0 q_j^0} \), \( i = 1, \ldots, N \) is the share of market capitalization of firm \( i \) at date 0 in the total market capitalization. Note that this can be rewritten as
where \( S^0_m = \frac{\sum_{i \in A_m} p^0_i q^0_i}{\sum_{i} p^0_i q^0_i} \) is the share of market capitalization (at the base date) of subaggregate \( A_m \) in the total market capitalization, and \( s^0_m = \frac{p^0_i q^0_i}{\sum_{j} p^0_j q^0_j} \) is the share of market capitalization of firm \( i \) in the subaggregate \( A_m, m = 1, \ldots, M \). Note then that \( R_N(p^0, p^1, q^0) \) can be rewritten as

\[
R_N(p^0, p^1, q^0) = \sum_{m=1}^{M} S^0_m R_{N_m}(p^{0N_m}, p^{1N_m}, q^{0N_m}),
\]

so that the direct index \( R_N \) can be calculated as a weighted average of the subindexes \( R_{N_m}, m = 1, \ldots, M \) (where the subindexes are treated like gross returns for individual stocks) and the functional form of \( R_N, R_{N_m} \) and that for aggregating \( R_{N_m} \) are all the same.

**A19: Equal impact of subaggregates (mean value property for the final stage of aggregation):** Suppose \( R_{N_m}(p^{0N_m}, p^{1N_m}, q^{0N_m}) = \rho, m = 1, \ldots, M \) then \( R_N(p^0, p^1, q^0) = \rho \). The rationale for this axiom is as follows. If each investor were able to achieve the same gross return of \( \rho \) on each mutually exclusive partition of \( N \), then the gross return from holding the aggregate portfolio (consisting of all subaggregates) should also be equal to \( \rho \). This may be considered as a type of no arbitrage condition analogous to the discussion for consistency in aggregation.

van IJzeren (1958) appears to have been the first to apply this axiom to index numbers and notes that this so-called "equality test" is "... a natural consequence of the aggregation property" (van IJzeren 1958)[page 434]. Consider however the Vartia-I index (Vartia 1976) \( R_{VI} \) defined by:

\[
\ln R_{VI}(p^0, p^1, q^0) = \frac{\sum_{i=1}^{N} L(p^1_i q^0_i, p^0_i q^0_i) \ln \left( \frac{p^1_i}{p^0_i} \right)}{L(p^1 q^0, p^0 q^0)},
\]

(5.13)
where $L : \mathbb{R}_{++}^2 \to \mathbb{R}_{++}$ is the logarithmic mean of two positive numbers defined by:

$$L(\alpha, \beta) = \begin{cases} 
\frac{\alpha - \beta}{\ln(\frac{\alpha}{\beta})}, & \alpha \neq \beta \\
\alpha, & \alpha = \beta.
\end{cases} \quad (5.14)$$

$R_{VI}$ satisfies A18 and is thus consistent in aggregation. If each subaggregate index has the same numerical value, i.e. $R_{VI(N_m)} = \rho$, $\forall m$, then it is not the case that $R_{VI(N)} = \rho$, since in general

$$L(p_1^T q^0, p_0^T q^0) \neq \sum_{m=1}^{M} L(\sum_{i \in A_m} p_1^i q_i^0, \sum_{i \in A_m} p_0^i q_i^0). \quad (5.15)$$

Satisfaction of axiom A18 does not guarantee satisfaction of A19. The geometric mean of individual stock gross returns satisfies A19 but fails A18. This means that A18 and A19 are independent.

Indeed A19 appears similar in spirit to the proportionality test (A7) which says that $R(p_0^0, \lambda p_0^0, q_0^0) = \lambda$, i.e. if $\frac{p_0^i}{q_0^i} = \lambda, i = 1, \ldots, N$ then the overall return is also equal to $\lambda$. In other words, the proportionality test resembles axiom A19, where each subaggregate consists of an individual stock with the same gross return—however these axioms are not equivalent. The index given by

$$R(p_0^0, p_1^0, q_0^0) = \frac{(p_0^0 + p_1^0)^T q_0^0 c^T p_1^0}{p_0^0 q_0^0 c^T (p_0^0 + p_1^0)},$$

satisfies A7 but not A18.

A20: Invariance to arbitrary subdivision of firms:

$$R_N(p_0^0, p_1^0, q_0^0) = R_{N+1}(p_0^0, p_0^N, p_1^1, p_1^N, q_1^0, \ldots, q_1^N, (1 - \lambda)q_N^0). \quad (5.17)$$

This axiom states that if firm $N$ is arbitrarily partitioned into two trading units of $\lambda q_N^0$ and $(1 - \lambda)q_N^0$ shares outstanding then the index number calculated using the $N$-stock index is the same as that computed by the $(N + 1)$-stock index where the two blocks of $N$ have the same base and current date prices and the shares outstanding of the two blocks are $\lambda q_N^0$ and
(1 - \lambda)q_N^0. This axiom can be generalized to more than two partitions of a given firm. This axiom only makes sense if axiom A6 is satisfied. This axiom specializes axiom A14, the merging axiom. In \( R_{N+1} \), the firm is arbitrarily split up into two trading blocks, and these two trading blocks merge with one another with \( k = 1 \) (the share conversion factor) in \( R_N \), i.e. the two separate components are put back together again, where \( p'_N = p'_{N+1}, t = 0,1 \) and '\( q^0_{N+1} + q_N^0 \)' (in \( R_{N+1} \)) = \( q_N^0 \) (in \( R_N \)). This axiom is the analogue to Diewert’s country partitioning test (Diewert 1993b)[page 314].

These axioms are used in chapter 6 in evaluating the various stock indexes in use today.
Chapter 6

Stock Market Indexes: An Evaluation

Chapter 5 proposed a set of axioms which are considered 'reasonable' for a stock market gross return index to possess. The dependencies among the various axioms were explored. Furthermore, it was seen that the imposition of certain axioms restricted the class of admissible index numbers and yielded a characterization result for stock market gross return indexes.

Axioms A1 to A20 are now used to evaluate some of the major stock market indexes in use today in order to determine which is 'best' in the sense that it is a 'reasonable' measure of gross return on a portfolio of stocks and satisfies a number of the axioms proposed in chapter 5. That index which satisfies most (or alternatively the most important) axioms will be deemed best according to the axiomatic evaluation criterion.

Recall again that the axiomatic approach to index numbers that has been adopted here looks at an index number formula from the point of view of its mathematical properties. Reasonable properties have been proposed that are considered desirable for an index $R(p^0, p^1, q^0)$ to possess, where the index depends (possibly) on the quantities of shares $q^0$ in a particular portfolio and the base and comparison date stock prices $p^0$ and $p^1$. No mention has been made of the appropriate choice of the basket $q^0$. For common stock market indexes such as the Standard and Poor's 500 Composite Index, the Wilshire Equity Indexes, the Russell Indexes, and the Toronto Stock Exchange Composite 300 Index the basket consists of the aggregate holdings of shares outstanding of investors (the 'market') for a given group
of firms listed on various stock exchanges. The potential choices for \( q^0 \) in terms of composition are limitless. Section 3.1 in chapter 3 discussed the uses of stock market indexes as benchmarks for investment performance—an appropriate benchmark involves choosing a basket which closely matches the investment and business characteristics of the investor and portfolio under evaluation. For example, \( q^0 \) may contain lower-capitalization stocks, higher-capitalization stocks, high technology stocks, or financial stocks. Furthermore the axioms consider the functional form for a stock market gross return index for a fixed basket of stocks over the evaluation period—changes in the portfolio, reinvestment of cash dividends etc. require adjustments to, and procedures involving the basic bilateral stock market gross return indexes—these issues have been discussed in chapter 3.

No attempt has been made to 'axiomatize' the choice of \( q^0 \) nor the manner in which cash dividends should be reinvested as received. Axiom A12, the cash dividend axiom proposed that an investor’s total return could be decomposed (additively) into a capital gains and income gains component. The manner in which these income gains are subsequently reinvested has not been formalized as a desirable property of an index—three possible choices for reinvesting cash proceeds were discussed in section 3.3.2 in chapter 3. Furthermore the bilateral axioms allow for no variation in the quantities of shares in the portfolio—the focus is on the measurement of return performance for a fixed portfolio of stocks over an evaluation period.

For the DJIA, a ratio of arithmetic means of stock prices, \( R_{DJ}(p^0, p^1, q^0) = \frac{\sqrt[N]{\sum_{i=1}^{N} p_i^0}}{\sqrt[N]{\sum_{i=1}^{N} p_i^1}} \) and the Value Line Geometric Average (VLG), a geometric mean of the gross returns on the individual stocks in the portfolio or \( R_{VLG}(p^0, p^1, q^0) = \prod_{i=1}^{N} \left(\frac{p_i^1}{p_i^0}\right)^{q_i^0} \) there is an apparent lack of quantity weightings. The same can be said of the Value Line Arithmetic Average (VLA), though as has been demonstrated in section 4.3.1 in chapter 4 the ‘basket’ \( q^0 \) for VLA has been chosen so that \( q^0(p^0) = \frac{1}{N}(p^0)^{-1} \) so \( p_i^0 q_i^0 = \frac{1}{N}, i = 1, \ldots, N \) or an equal dollar amount is invested in each stock (where the initial investment outlay is one dollar). Thus inverse prices act as implicit quantity weightings in a Laspeyres-type formula (although explicit quantities of shares outstanding are absent from the index calculation) and \( R_{VLA}(p^0, p^1, q^0) = \frac{\prod_{i=1}^{N} p_i^1 q_i^0}{p_0^1 q_0^0} = \)
\[ \frac{p_{i}^{1}}{p_{i}^{0}} \frac{1}{(p_{i}^{0})^{-1}} = \frac{1}{N} \sum_{i=1}^{N} \left( \frac{p_{i}^{1}}{p_{i}^{0}} \right). \]

The Standard and Poor's Composite 500 Index (S & P 500) explicitly considers the quantities of shares in the portfolio over the evaluation period by comparing the total market capitalizations of firms in the portfolio at the base and current dates so \( R_{S & P}(p_{0}^{0}, p_{1}^{0}, q_{0}^{0}) = \frac{p_{1}^{0} q_{0}^{0}}{p_{0}^{0} q_{0}^{0}}. \)

### 6.1 Axiomatic Evaluation of Indexes

Each index is now evaluated in terms of its satisfaction or failure of the proposed set of axioms.

#### 6.1.1 Standard and Poor's 500 Composite

**Proposition 6.1** The S & P 500 (whose basic bilateral index number formula is Laspeyres) satisfies axioms A1 to A20.

**Proof:** This follows from the Laspeyres functional form for this index and corollary 1, chapter 5. Verification of axioms A16 to A20 is straightforward to check.

#### 6.1.2 Dow Jones Industrial Average

Some of the more frequent criticisms of the DJIA can be formalized by noting that of the twenty axioms proposed, it fails six. It fails axiom A4 and thus is not invariant to the units of measurement. It fails axioms A13 and A14, the tiny holdings and merging axioms respectively. Its value does not tend to the gross return on the dominant firm as it gains market capitalization among the portfolio of stocks constituting the index, and thus it fails axiom A17. It is not consistent in aggregation since an arbitrary set of weights needs to be assigned to each of the subindexes so that they aggregate up to the overall direct index value. Axiom A20, the invariance to arbitrary partitioning axiom is only satisfied when the
gross return to the partitioned firm \(N\) equals the gross return calculated on the \(N\)-stock portfolio (which includes firm \(N\)), i.e. \(\frac{p_N^t}{p_N^0} = R_{DJ(N)}(p^0, p^1, q^0)\).

Note that the DJIA, \(R(p^0, p^1, q^0)_{DJ} = \frac{1}{N} \sum_{i=1}^{N} \frac{p_i^1}{p_i^0}\) corresponds to the Laspeyres fixed-basket index in the special case where \(q^0 = k(1, \ldots, 1)^T = k_{130}, k > 0\), since \(\frac{p_i^1}{p_i^0} q^0 = \frac{p_i^1}{p_i^0} k_{130} = \frac{1}{N} \sum_{i=1}^{N} \frac{p_i^1}{p_i^0}\). If in fact the basket is \(q^0 = k_{130}\) then the DJIA does satisfy all of the axioms since it is a Laspeyres index with a very particular basket. However this basket is very unrealistic, and unlikely to correspond to the actual investment strategy of any particular investor (or investors in the aggregate).

### 6.1.3 Value Line Geometric Average

The VLG, being multiplicative in individual gross returns fails axiom A12, the cash dividend test. Firms that go bankrupt (and whose share prices tend to zero) drive the index value down to zero—a highly undesirable feature of a stock market gross return index. It fails axioms A14 and A17, the merging and dominance axioms, respectively. It is consistent in aggregation only when each subindex assumes the same value. It satisfies the arbitrary subdivision axiom A20 only when \(\frac{p_N^t}{p_N^0} = R_{VLG(N)}(p^0, p^1, q^0)\), so that the gross return on the dividing firm is equal to the gross return on the portfolio consisting of the \(N\) original firms.

### 6.1.4 Value Line Arithmetic Average

The VLA satisfies axioms A1 to A12 and fails A13 and A14, the tiny holdings and merging axioms respectively. Interestingly, of the four common index types in use today it violates time reversibility so that the product of the index going from 0 to 1 times the index going from 1 back to 0 does not necessarily equal one. In fact

\[
R_{VLG}(p^0, p^1, q^0) R_{VLG}(p^1, p^0, q^0) = \left( \frac{1}{N} \sum_{i=1}^{N} \left( \frac{p_i^1}{p_i^0} \right) \right) \left( \frac{1}{N} \sum_{i=1}^{N} \left( \frac{p_i^0}{p_i^1} \right) \right) \geq 1, \tag{6.1}
\]

so that if all stock prices revert back to their original level \((p^0)\) at the end of the second period the overall net return is non-negative. Equation 6.1 holds with equality when \(p^1 = 1\).
<table>
<thead>
<tr>
<th>Test</th>
<th>S &amp; P 500</th>
<th>DJIA</th>
<th>VLG</th>
<th>VLA</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1 monotonicity</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>A2 homogeneity</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>A3 identity</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>A4 commensurability</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>A5 dimensionality</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>A6 symmetry</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>A7 proportionality</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>A8 mean-value</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>A9 positivity</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>A10 scaling of quantities</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>A11 homogeneity of degree minus one</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>A12 cash dividend</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>A13 tiny holdings</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>A14 merging</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>A15 transitive</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>A16 time reversal</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>A17 dominance of firms</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>A18 consistency in aggregation</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>A19 equal impact</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>A20 subdivision</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
</tbody>
</table>

S & P 500 (Standard and Poor’s 500 index): \[ R_{S&P}(p^0, p^1, q^0) = \frac{p^1}{p^0} \]
DJIA (Dow Jones Industrial Average): \[ R_{DJ}(p^0, p^1, q^0) = \frac{1}{N} \sum_{i=1}^{N} \frac{p_i^1}{p_i^0} \]
VLG (Value Line Geometric Average): \[ R_{VLG}(p^0, p^1, q^0) = \prod_{i=1}^{N} \left( \frac{p_i^1}{p_i^0} \right)^{\frac{1}{q}} \]
VLA (Value Line Arithmetic Average): \[ R_{VLA}(p^0, p^1, q^0) = \frac{1}{N} \sum_{i=1}^{N} \left( \frac{p_i^1}{p_i^0} \right) \]

Table 6.1: Evaluation of indexes based on axiomatic approach
Interestingly this failure of time reversal has been one of the reasons why some statistical agencies around the world have changed their elementary indexes (i.e. indexes at the lowest level of aggregation) from arithmetic means of price relatives to geometric means of price relatives.

Given the implied investment strategy underlying the arithmetic mean of price relatives (equal proportionate investment in each stock in a N-stock portfolio) the index \( R_{VLA}(p^1, p^0, q^0) \) does not correspond to any return on the initial basket \( q^0(\hat{p}^0) = \frac{1}{N}(\hat{p}^0)^{-1} \).

In fact, in calculating \( R_{VLA}(p^1, p^0, q^0) \) the initial weightings (in terms of its Laspeyres formulation) on the stocks have been changed to 'new' weights \( q^1(\hat{p}^1) = \frac{1}{N}(\hat{p}^1)^{-1} \) relative to \( R_{VLA}(p^0, p^1, q^0) \). In calculating the overall return \( R_{VLA}(p^0, p^1, q^0) R_{VLA}(p^1, p^0, q^0) \) the basket has been rebalanced to equal weights at the end of the \((0, 1)\) evaluation period, so that the actual calculation corresponds to \( R_L(p^0, p^1, q^0(\hat{p}^0)) R_L(p^1, p^0, q^1(\hat{p}^1)) \) when the functional form is the Laspeyres index \( (R_L) \). In other words there is a basket change at \( t = 1 \). Thus axiom A16 is not applicable here since the basket has not remained fixed over the evaluation period. While the VLA fails time reversal it must be remembered that the basket (i.e. 'weights') has varied, contrary to the non-allowance of variations in quantities. If the weightings (initial proportionate investments) are fixed at \( q^0(\hat{p}^0) = \frac{1}{N}(\hat{p}^0)^{-1} \) over the entire (two period) evaluation period and the index is written in Laspeyres form \( R_L(p^0, p^1, q^0(\hat{p}^0)) \), then time reversal is satisfied since \( R_L(p^0, p^1, q^0(\hat{p}^0)) R_L(p^1, p^0, q^0(\hat{p}^0)) = 1 \).

The VLA is not consistent in aggregation since an arbitrary set of weights needs to be assigned to each of the subindexes to make them aggregate to the overall index. It satisfies axiom A20 only when \( R_{VLA(N)}(p^0, p^1, q^0) = \frac{p^1_p}{p^0_p} \).

\(^1\)The relationship in equation 6.1 follows from the Theorem of the Arithmetic, Geometric and Harmonic Means which states that for all \( x_i > 0, i = 1, \ldots, N \) that

\[
\sum_{i=1}^{N} \frac{1}{N} x_i \geq \prod_{i=1}^{N} x_i^{\frac{1}{N}} \geq \left( \sum_{i=1}^{N} \frac{1}{N} x_i^{-1} \right)^{-1}.
\]

The result follows by setting \( x_i = \frac{p^1_p}{p^0_p}, i = 1, \ldots, N \).
Table 6.1 summarizes each index's satisfaction or failure of each of the axioms. While the table lists all twenty axioms it should be recalled from chapter 5 that some axioms are redundant in the sense that they are implied by other axioms. In particular axioms A7-A11 follow from axioms A1-A6, and A16 is a special case of A15. However all axioms have been listed for completeness—as was mentioned some readers may consider some axioms more appealing or intuitive than others. Table 6.1 illustrates whether or not these independent or dependent axioms are satisfied by the DJIA, S & P 500, VLA and VLG. A '✓' indicates satisfaction of the test, while a '×' denotes failure of the test.

6.2 Concluding Remarks

As was mentioned at the outset, the 'best' index from the axiomatic point of view is that index which satisfies the most axioms, or alternatively that index number formula which satisfies the most important axioms. Axioms A1-A6 are considered minimal requirements (and A7-A11 follow from requiring A1-A6). The DJIA violates axiom A4 and thus is sensitive to the units of measurement in which the index is measured. This means that when the DJIA is used as a return index it is not invariant to stock splits. Since stock splits do not change the value of an investor's portfolio, failure of this axiom is a serious shortcoming of the DJIA.

Imposing A1-A6 however means that the three index types S & P 500, VLA and VLG are admissible return indexes. Adding the requirement of transitivity implies that the return index can be written as the ratio of the same function of current and base date market capitalizations, where the function is a homogeneous symmetric mean. However, the class of homogeneous symmetric means is large. Adding the requirement of additivity (A12) implies that the functional form is Laspeyres. Thus a specific functional form for a stock market gross return index is suggested. The axioms leading to this conclusion appear self-evident and have intuitive appeal. While it appears that the Laspeyres index dominates all the others in terms of its satisfaction of axioms, it is possible for an index to violate an axiom mathematically, though satisfy it approximately numerically. Diewert (1978) for example
has shown that superlative indexes are approximately consistent in aggregation.

In practice, investors hold only specific stocks in the universe of stocks listed on stock exchanges. In this sense ‘market’ portfolios of stocks are partitioned into subportfolios. Indeed investors may wish to gauge the performance of certain units of their portfolio holdings such as small-cap versus large-cap, financial versus non-financial etc. Quite often the need for publishing certain subindexes arises from differences in risk characteristics—certain subsectors are associated with different levels of risk and therefore require separate benchmark indexes. When this is done axioms \(A_{18}\) and \(A_{19}\) are important for computing such indexes at these more disaggregated levels. The requirement of consistency in aggregation is indispensable for such subindexes. Of the four indexes, only the S & P 500 satisfies the consistency in aggregation axiom.

The axioms proposed make an overwhelming case for the adoption of the Laspeyres index as the basic bilateral index for measuring the return performance on some portfolio of stocks. However, choosing the ‘ultimate’ index cannot and should not be made with respect to satisfaction of axioms alone. Section 3.1 in chapter 3 noted that one of the uses of stocks indexes is as benchmarks for portfolio evaluation. Choosing a benchmark requires identifying that index which most closely matches the investment and business characteristics of the investor and portfolio in question. For example, the VLA corresponds to an investment strategy consistent with daily rebalancing to equal dollar holdings in each of the stocks in a given portfolio—such an index may be useful for evaluating a fund manager with an ‘active’ strategy of continual buying and selling of stocks. The VLA would be unsuited to evaluating a fund manager with a passive ‘buy-and-hold’ type of strategy.

Each index is consistent with a particular type of investment strategy—an important feature to bear in mind when using such indexes as return measures. Quite apart from its failure of commensurability and consistency in aggregation the DJIA corresponds to an investment in one share of each of its thirty component stocks—this is unlikely to correspond to the investment strategy of any one individual investor.

Section 3.3.2 in chapter 4 discussed the concept of a total return index, i.e. one that mea-
sures both capital and income gains of investment returns. Such an interpretation requires an appropriate treatment of cash dividends. Three possible treatments of cash dividends were discussed—each resulting in a different total return index. Thus the total return index chosen depends on what seems to be the most plausible use of the cash dividends earned—this may vary from investor to investor and is thus not 'axiomatized'.

The analyses in chapters 3-6 strongly suggest that the S & P 500 is the 'best' index on several counts. It satisfies the most axioms and is consistent with a sensible investment strategy. The work highlights in a formal sense the principal drawbacks of the DJIA: namely that it is not invariant to stock splits, it is not consistent in aggregation and corresponds to an unusual investment strategy. This examination of stock indexes breaks new ground—nowhere in the literature is there a clear and coherent discussion of the uses, meaning and construction of such indexes. Indeed the theoretical underpinnings of such indexes is not well developed. Other interesting questions remained unexamined in this thesis. For example, given the globalization of the world's equity markets, calculation of returns on international portfolios is necessary. When one enters the realm of international equity markets one is faced with the difficult problem of dealing with currency fluctuations in measuring multilateral returns.
Bibliography


Fisher, I. (1922), The making of index numbers, Houghton Mifflin, Boston.


Statistics Canada (1989), Retail commodity survey, Cat. no. 63-541-XPB.
Statistics Canada (1997), Retail chain and department store survey, 1994, Cat. no. 63-210-XPB.


Appendix A

Appendix to Chapter 2

A.1 Outlet coverage

Tables A.1 to A.3 document some examples of the distribution of the sample size by outlet type over the 1990-1996 period, where the distribution is based on the average sample size over the year in question.

The distributions are those used by Statistics Canada in its data capture process for Ontario (strata level compositions may vary from those of the province,\(^1\)) and figures in parentheses refer to the outlet sample size. As an example, for the basic class *photographic services and supplies* (see table A.2), the distribution of the sample has remained relatively stable over the period. For *audio equipment* there has been some shift in distribution: in 1990 department stores and independents represented 31% and 27% of the sample respectively—in 1996 these had fallen to 20% and 10% respectively.

\(^1\)There are currently four pricing strata in Ontario. With all basic classes in this study a qualifier is requisite: in 1995 the sample size was cut by over 40%, since which time the majority of price movements have been imputed from the Toronto strata. Thus the 1996 distribution in the tables is a strata-specific one, though is supposed to represent that of the province. Indeed given the small sample sizes, the addition or removal of just one outlet can drastically alter the distribution.
### Table A.1: Percentage distribution of outlet sample by outlet type for non-prescribed medicines, 1990-1996.

<table>
<thead>
<tr>
<th>non-prescribed medicines</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>0%(0)</td>
<td>36%(16)</td>
<td>0%(0)</td>
<td>38%(17)</td>
<td>27%(12)</td>
<td>100%(45)</td>
</tr>
<tr>
<td>1992</td>
<td>2%(1)</td>
<td>34%(15)</td>
<td>0%(0)</td>
<td>43%(19)</td>
<td>20%(9)</td>
<td>100%(44)</td>
</tr>
<tr>
<td>1994</td>
<td>2%(1)</td>
<td>38%(18)</td>
<td>0%(0)</td>
<td>40%(19)</td>
<td>19%(9)</td>
<td>100%(47)</td>
</tr>
<tr>
<td>1996</td>
<td>0%(0)</td>
<td>5%(1)</td>
<td>0%(0)</td>
<td>67%(14)</td>
<td>29%(6)</td>
<td>100%(21)</td>
</tr>
</tbody>
</table>

Figures in parentheses refer to the number of outlets.

### Table A.2: Percentage distribution of outlet sample by outlet type for photographic services and supplies, 1990-1996.

<table>
<thead>
<tr>
<th>photo services and supplies</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>24%(5)</td>
<td>14%(3)</td>
<td>38%(8)</td>
<td>0%(0)</td>
<td>24%(5)</td>
<td>100%(21)</td>
</tr>
<tr>
<td>1992</td>
<td>17%(4)</td>
<td>13%(3)</td>
<td>39%(9)</td>
<td>0%(0)</td>
<td>30%(7)</td>
<td>100%(23)</td>
</tr>
<tr>
<td>1994</td>
<td>22%(5)</td>
<td>13%(3)</td>
<td>35%(8)</td>
<td>0%(0)</td>
<td>30%(7)</td>
<td>100%(23)</td>
</tr>
<tr>
<td>1996</td>
<td>29%(2)</td>
<td>14%(1)</td>
<td>43%(3)</td>
<td>0%(0)</td>
<td>14%(1)</td>
<td>100%(7)</td>
</tr>
</tbody>
</table>

Figures in parentheses refer to the number of outlets.

### Table A.3: Percentage distribution of outlet sample by outlet type for audio equipment and other household equipment, 1990-1996.

<table>
<thead>
<tr>
<th>audio equipment</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>31%(8)</td>
<td>4%(1)</td>
<td>8%(2)</td>
<td>31%(8)</td>
<td>27%(7)</td>
<td>100%(26)</td>
</tr>
<tr>
<td>1992</td>
<td>24%(6)</td>
<td>4%(1)</td>
<td>4%(1)</td>
<td>40%(10)</td>
<td>28%(7)</td>
<td>100%(25)</td>
</tr>
<tr>
<td>1994</td>
<td>22%(5)</td>
<td>0%(0)</td>
<td>13%(3)</td>
<td>39%(9)</td>
<td>26%(6)</td>
<td>100%(23)</td>
</tr>
<tr>
<td>1996</td>
<td>20%(2)</td>
<td>10%(1)</td>
<td>10%(1)</td>
<td>50%(5)</td>
<td>10%(1)</td>
<td>100%(10)</td>
</tr>
</tbody>
</table>

Figures in parentheses refer to the number of outlets.
A.2 Sample turnover

In this section approximate measures of the rate of sample turnover are described, and these are meant to capture the extent to which new outlets are being added to the outlet sample, and other less representative ones are being dropped. Outlets may exit from the outlet sample for a number of reasons, including closure or non-availability of a previously priced item specification. In Canada, unlike the U.S., no systematic sample rotation is undertaken, and when an outlet is dropped from the sample it is quite often the case that it is not replaced for several months (if at all). Sample inertia is defined here as the fraction of the sample that is priced consistently throughout a calendar year. The annual turnover rate ($ATR$) is defined as the percentage of the outlet sample that undergoes 'rotation' (i.e. does not remain in the sample) in any given year, and is

$$ATR = \frac{1}{T} \left( \sum \frac{\text{no. of (rotations, non-matches)}}{\text{no. of (rotations, non-matches, non-movements)}} \right) \times 100\% . \quad (A.1)$$

where $T$ is the number of years, a rotation is an entry paired with an exit, non-matches are unpaired outlets (e.g. an entry with no corresponding exit), and non-movements are outlets that neither enter nor exit over the course of a year. Note that this measure is a hypothetical one and is not related to any procedure used by Statistics Canada. Extraneous years are excluded from this calculation—these include years in which the sample size was cut (due to budget constraints) or when new outlets were priced for an entirely new item specification. The $ATR$'s are generally low.

A.3 Market shares of outlet types

The lowest level of commodity detail on which reliable market share information is available is sales by department (in department stores)—these departments were chosen to overlap most closely with the basic classes under examination (see table A.4). The breakdown of sales by department between major and discount department stores for Ontario was obtained
<table>
<thead>
<tr>
<th>Basic classes</th>
<th>Departments used for comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-prescribed medicines</td>
<td>Toiletries, cosmetics and pharmaceutical products</td>
</tr>
<tr>
<td>Photographic services and supplies</td>
<td>Photographic equipment and supplies</td>
</tr>
<tr>
<td>Audio equipment</td>
<td>Home entertainment equipment</td>
</tr>
<tr>
<td>Other household equipment</td>
<td>Housewares and small electrical appliances</td>
</tr>
</tbody>
</table>

Sources: Statistics Canada *Consumer Price Index Reference Paper* and *Annual Retail Chain and Department Store Survey*, 1995.

Table A.4: Department store categories and associated basic classes.

from unpublished data collected by Statistics Canada for the ARC. Potential discrepancies may arise between the two. The sales by retail department include all goods/varieties sold during the survey period. The corresponding CPI basic class, by contrast, usually covers only two or three narrowly defined items. Nonetheless these items have been chosen to be representative of the basic class to which they belong, and the correspondence between the two categories is satisfactory for present purposes. Furthermore it is unlikely that the items included in the CPI basic class are more or less subject to outlet substitution biases than the corresponding items comprising the departments in the ARC.

For the remaining three outlet classifications, the principal SICs (*excluding* outlets covered by the ARC) where each of the commodities were thought to be sold were documented. As an example, *non-prescribed medicines* are sold at discount department stores, pharmacy chains, supermarkets and general merchandisers. For each SIC code the corresponding outlets in the RTS sample for Ontario were classified into one of *T3*- *T5*, resulting in a tabulation of total sales by SIC code for outlet types *T3* to *T5*.

Recall that each SIC sells a range of commodities—for example SIC 6221 (appliance, television, radio and stereo stores) can sell major appliances, audio equipment, televisions, computers, and small electrical appliances. Commodity breakdowns for SIC codes are not currently available. However approximate commodity breakdown estimates for each SIC/outlet type classification were obtained using information from Statistics Canada (1989), preliminary data from the 1997 *Retail Commodity Survey*, *Discount Store News* (U.S.) industry surveys (various issues) and *Retail Chain Guide* (U.S.) industry surveys (various issues).
### Other household equipment

<table>
<thead>
<tr>
<th>Year</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>23.2</td>
<td>42.3</td>
<td>0.1</td>
<td>19.6</td>
<td>14.8</td>
</tr>
<tr>
<td>1992</td>
<td>23.2</td>
<td>42.8</td>
<td>0.1</td>
<td>19.6</td>
<td>14.2</td>
</tr>
<tr>
<td>1993</td>
<td>22.0</td>
<td>41.8</td>
<td>0.2</td>
<td>21.3</td>
<td>14.7</td>
</tr>
<tr>
<td>1994</td>
<td>19.2</td>
<td>46.1</td>
<td>0.2</td>
<td>21.3</td>
<td>13.2</td>
</tr>
<tr>
<td>1995</td>
<td>22.2</td>
<td>45.4</td>
<td>0.2</td>
<td>19.3</td>
<td>12.9</td>
</tr>
<tr>
<td>1996</td>
<td>16.1</td>
<td>50.4</td>
<td>0.2</td>
<td>17.5</td>
<td>15.8</td>
</tr>
</tbody>
</table>

### Non-prescribed medicines

<table>
<thead>
<tr>
<th>Year</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>2.2</td>
<td>16.7</td>
<td>30.8</td>
<td>32.4</td>
<td>17.9</td>
</tr>
<tr>
<td>1992</td>
<td>2.2</td>
<td>15.4</td>
<td>29.4</td>
<td>35.6</td>
<td>17.4</td>
</tr>
<tr>
<td>1993</td>
<td>2.1</td>
<td>14.2</td>
<td>29.0</td>
<td>38.3</td>
<td>16.4</td>
</tr>
<tr>
<td>1994</td>
<td>2.1</td>
<td>17.7</td>
<td>28.4</td>
<td>36.8</td>
<td>15.0</td>
</tr>
<tr>
<td>1995</td>
<td>1.3</td>
<td>18.9</td>
<td>28.0</td>
<td>36.2</td>
<td>15.5</td>
</tr>
<tr>
<td>1996</td>
<td>1.1</td>
<td>19.6</td>
<td>26.9</td>
<td>37.7</td>
<td>14.7</td>
</tr>
</tbody>
</table>

### Audio equipment

<table>
<thead>
<tr>
<th>Year</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>20.4</td>
<td>22.1</td>
<td>5.4</td>
<td>24.7</td>
<td>27.4</td>
</tr>
<tr>
<td>1992</td>
<td>18.7</td>
<td>24.0</td>
<td>5.9</td>
<td>25.8</td>
<td>25.6</td>
</tr>
<tr>
<td>1993</td>
<td>15.2</td>
<td>24.9</td>
<td>6.9</td>
<td>28.0</td>
<td>25.0</td>
</tr>
<tr>
<td>1994</td>
<td>13.3</td>
<td>28.5</td>
<td>7.6</td>
<td>29.9</td>
<td>20.8</td>
</tr>
<tr>
<td>1995</td>
<td>11.4</td>
<td>35.5</td>
<td>8.0</td>
<td>29.8</td>
<td>15.3</td>
</tr>
<tr>
<td>1996</td>
<td>12.2</td>
<td>36.5</td>
<td>7.9</td>
<td>28.1</td>
<td>15.4</td>
</tr>
</tbody>
</table>


Table A.5: Market shares of outlet types
The outlets for which commodity breakdown information was available were assumed to be representative of the SIC/outlet type category to which it belonged. Each SIC/outlet type classification’s sales were pro-rated by the approximate commodity breakdowns. The resulting market shares are shown in table A.5.

**A.4 Statistical significance of average price differences**

Since the sample sizes for discount department stores are typically ‘small’ some t-statistics were computed to test for the statistical significance of differences in average prices across the outlet types. The t-statistics reported in table A.6 are those for four CPI basic classes—these reported t-statistics test for the statistical significance of differences in average prices between discount department stores ($T_2$) and all other types combined ($T_1, T_3 - T_5$) for the entire seven year period, since the biases reported in table 2.7 in chapter 2 estimate the bias arising from $T_2$ compared to all other outlet types. Figures in parentheses in table A.6 show the degrees of freedom.

All average price differences are significant at the 10% level of significance. Pairwise comparisons of average prices for the five outlet types (not shown here) show that the differences in average prices between types $T_3$ and $T_5$ are not statistically significant. However the differences in price levels between discount department stores and all other outlet types are

<table>
<thead>
<tr>
<th>Basic class</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>other household equipment</td>
<td>1.86</td>
</tr>
<tr>
<td></td>
<td>(1628)</td>
</tr>
<tr>
<td>non-prescribed medicines</td>
<td>3.51</td>
</tr>
<tr>
<td></td>
<td>(3230)</td>
</tr>
<tr>
<td>audio equipment</td>
<td>392.56</td>
</tr>
<tr>
<td></td>
<td>(1715)</td>
</tr>
<tr>
<td>photographic services and supplies</td>
<td>618.59</td>
</tr>
<tr>
<td></td>
<td>(1558)</td>
</tr>
</tbody>
</table>

Table A.6: t-statistics for differences in average prices across outlet types.
statistically significant though less significant for T4 (discount chains).
Appendix B

Appendix to Chapter 5

B.1 Independence of axioms A1-A6

Axioms A1 to A6 are independent—i.e. it is possible for any five of them to be satisfied without the sixth being satisfied:

1. 
   \[ R(p^0, p^1, q^0) = \prod_{i=1}^{N} \left( \frac{p_i^1}{p_i^0} \right)^{-\frac{1}{N}} \left( \min_{i=1, \ldots, N} \left( \frac{p_i^1}{p_i^0} \right) \right)^2 \]
   violates axiom A1 but satisfies A2 to A6.

2. 
   \[ R(p^0, p^1, q^0) = \frac{c^\top p^1}{c^\top p^0}, c = k N, k > 0, \]
   satisfies A1, A2, A3, A5, A6, but not A4.

3. 
   \[ R(p^0, p^1, q^0) = \frac{p_i^0 q^0}{p_i^0 q^0 + 1} \left( \frac{1}{N} \sum_{i=1}^{N} \frac{p_i^1}{p_i^0} \right) + \frac{1}{p_i^0 q^0 + 1} \max \{ p_i^1, \ldots, \frac{p_i^1}{p_i^0} \}, \]
   satisfies A1, A2, A3, A4, A6 but not A5.

4. 
   \[ R(p^0, p^1, q^0) = \sqrt{\frac{p_i^1 q^0}{p_i^0 q^0}}, \]
   satisfies A1, A3, A4, A5, A6 but not A2.
5. \[ R(p^0, p^1, q^0) = \frac{p^1 \tau(kq^0)}{p^0 \tau q^0}, \quad k > 0, \]
satisfies A1, A2, A4, A5, A6 but violates A3 unless \( k = 1 \).

6. \[ R(p^0, p^1, q^0) = \frac{p^1}{p^0}, \]
satisfies A1 to A5, but violates A6 since stock 1 plays an asymmetric role in determining the return on the portfolio.

B.2 Dependency of axioms A7-A11

A7: Proportionality: \( R(p^0, \lambda p^0, q^0) = \lambda, \forall \lambda > 0 \).

Proof:
\[
R(p^0, \lambda p^0, q^0) = \lambda R(p^0, p^0, q^0) \quad \text{by axiom A2}
\]
\[
= \lambda \quad \text{by axiom A3}.
\]
Thus linear homogeneity and identity imply proportionality. The converse is not true. Clearly if \( R \) satisfies proportionality it satisfies the identity axioms which is a weaker form of proportionality got by setting \( \lambda = 1 \). However, the index given by
\[
R(p^0, p^1, q^0) = \frac{(p^0 + p^1)^\tau q^0 c^\tau p^1}{p^0 \tau q^0 c^\tau (p^0 + p^1)},
\]
(B.1)
satisfies proportionality but fails linear homogeneity.

A8: Mean value property:
\[
\min_{i=1,...,N} \left( \frac{p^1_i}{p^0_i} \right) \leq R(p^0, p^1, q^0) \leq \max_{i=1,...,N} \left( \frac{p^1_i}{p^0_i} \right), \quad i = 1, \ldots, N.
\]

Proof: Define \( \alpha = \min_{i=1,...,N} \left( \frac{p^1_i}{p^0_i} \right) \) and \( \beta = \max_{i=1,...,N} \left( \frac{p^1_i}{p^0_i} \right) \). By definition of \( \alpha \) and \( \beta \),
\[
\alpha p^0 \leq p^1 \leq \beta p^0.
\]

\[1\]This method of proof is due to Eichhorn (1978)[page 155]. Eichhorn attributes the method of proof to Helmut Funke.
\[ R(p^0, p^1, q^0) \geq R(p^0, \alpha p^0, q^0) \text{ by axiom A1} \]
\[ = \alpha R(p^0, p^0, q^0) \text{ by axiom A2} \]
\[ = \alpha \text{ by axiom A3}. \]

Analogously
\[ R(p^0, p^1, q^0) \leq R(p^0, \beta p^0, q^0) \text{ by axiom A1} \]
\[ = \beta R(p^0, p^0, q^0) \text{ by axiom A2} \]
\[ = \beta \text{ by axiom A3}. \]

The result follows.

**A9: Positivity:** \( R(p^0, p^1, q^0) > 0, \forall p^0, p^1 \gg 0_N \). This follows from the *mean value property*—if all stock prices are positive then \( R(p^0, p^1, q^0) \geq \alpha > 0 \).

**A10: Invariance to scaling of quantities:** \( R(p^0, p^1, \lambda q^0) = R(p^0, p^1, q^0), \forall \lambda > 0. \)

**Proof:**
\[ R(p^0, p^1, \lambda q^0) = R(\mu p^0, \mu p^1, \lambda q^0), \mu > 0 \text{ by axiom A5} \]
\[ = R(\lambda^{-1} p^0, \lambda^{-1} p^1, \lambda q^0) \text{ by setting } \mu = \lambda^{-1} \]
\[ = R(p^0, p^1, q^0) \text{ by axiom A4}. \]

**A11: Homogeneity of degree -1 in base date prices:** \( R(\lambda p^0, p^1, q^0) = \lambda^{-1} R(p^0, p^1, q^0), \lambda > 0. \)

**Proof:**
\[ R(p^0, p^1, q^0) = R(\lambda p^0, \lambda p^1, q^0) \text{ by axiom A5} \]
\[ = \lambda R(\lambda p^0, p^1, q^0) \text{ by axiom A2}. \]

### B.3 Proofs of Propositions

**Proposition 5.1** If \( R \) satisfies axioms A4 and A15 then there exists a function \( f \) such that \( R(p^0, p^1, q^0) = \frac{f(q^0, p^1)}{f(q^0, p^0)}. \) Conversely any function \( R(p^0, p^1, q^0) = \frac{f(q^0, p^1)}{f(q^0, p^0)} \) satisfies axioms A4 and A15.
Proof: By transitivity

\[ R(p, p^0, q^0)R(p^0, p^1, q^0) = R(p, p^1, q^0), \]  
(B.2)

or alternatively

\[ R(p^0, p^1, q^0) = \frac{R(p, p^1, q^0)}{R(p, p^0, q^0)}. \]  
(B.3)

Since the LHS of equation B.3 is independent of \( p \), then it follows that the RHS must also be independent of \( p \). Setting \( p = 1_N \) on the RHS of B.3 yields

\[ R(p^0, p^1, q^0) = \frac{R(1_N, p^1, q^0)}{R(1_N, p^0, q^0)} \]  
(B.4)

\[ = \frac{g(p^1, q^0)}{g(p^0, q^0)}, \]  
(B.5)

where \( g(p, q) \) is defined as \( g(p, q) \equiv R(1_N, p, q) \). Now by axiom A4, setting \( D = q^0 \) yields

\[ R(p^0, p^1, q^0) = \frac{R(1_N, q^0p^1, (q^0)^{-1}q^0)}{R(1_N, q^0p^0, (q^0)^{-1}q^0)} \]  
(B.6)

\[ = \frac{g(q^0p^1, 1_N)}{g(q^0p^0, 1_N)} \]  
(B.7)

\[ = \frac{f(q^{0}p^{1})}{f(q^{0}p^{0})}, \]  
(B.8)

where \( f(qp) = g(qp, 1_N) \).\(^3\) The converse is straightforward to verify. \( \blacklozenge \)

**Proposition 5.2** If \( R \) satisfies axioms A1, A2, A3, A4, A6 and A15 then there exists a function \( f \) such that \( R(p^0, p^1, q^0) = \frac{f(q^{0}p^{1})}{f(q^{0}p^{0})} \). Furthermore, \( f \) is a homogeneous symmetric mean, i.e. it is an increasing, linearly homogeneous and symmetric function of the individual

\(^2\)This method of proof is due to Eichhorn (1978)[pages 156-157].

\(^3\)Note that \( q_{11}p_{1}, \ldots, q_{NN}p_{N}^{T} \).
market capitalizations \( q_i p_i, i = 1, \ldots, N \) that satisfies the following mean value property
\[
f(\lambda 1_N) = \lambda .
\]

**Proof:** By proposition 5.1, if \( R(p^0, p^1, q^0) \) satisfies axioms A4 and A15 then
\[
R(p^0, p^1, q^0) = \frac{f(\hat{q}^0 p^1)}{f(\hat{q}^0 p^0)}.
\]
Recall that \( f(\hat{q} p) = g(\hat{q} p, 1_N) = R(1_N, \hat{q} p, 1_N) \). Now
\[
f(\hat{q} p) = R(1_N, \hat{q} p, 1_N)
= R(\hat{1}_N, \hat{q} p, \hat{1}_N) \text{ by axiom A6}
= R(1_N, \hat{q} p, 1_N) \text{ since } \hat{1}_N = 1_N
= f(\hat{q} p),
\]
so that \( f \) is a symmetric function.\(^4\) Also
\[
f(\lambda \hat{q} p) = R(1_N, \lambda \hat{q} p, 1_N)
= \lambda R(1_N, \hat{q} p, 1_N) \text{ by axiom A2}
= \lambda f(\hat{q} p),
\]
which shows that \( f \) is a (positively) linearly homogeneous function.

Suppose that \( \hat{q}^1 p^1 \geq q^0 p^0 \) then
\[
f(\hat{q}^1 p^1) = R(1_N, \hat{q}^1 p^1, 1_N)
\geq R(1_N, \hat{q}^0 p^0, 1_N) \text{ by axiom A1}
= f(\hat{q}^0 p^0),
\]
so that \( f \) is an increasing function. Also note that \( f(1_N) = R(1_N, 1_N, 1_N) = 1 \) by axiom A3
and \( f(\lambda 1_N) = P(1_N, \lambda 1_N, 1_N) = \lambda \) by axioms A2 and A3. \( \blacklozenge \)

**Proposition 5.3** Suppose that \( R(p^0, p^1, q^0) \) satisfies axioms A1, A2, A3, A4, A6, A12 and A15. Then \( R(p^0, p^1, q^0) = \frac{p^1}{p^0} \frac{q^0}{q^1} \), so that the stock market index is value-weighted. Conversely, \( R(p^0, p^1, q^0) = \frac{p^1}{p^0} \frac{q^0}{q^1} \) satisfies axioms A1, A2, A3, A4, A6, A12 and A15.

\(^4\) Recall that \( \hat{x} \) denotes a permutation of the elements of \( x \).
Proof: If $R$ satisfies axioms A1, A2, A3, A4, A6 and A15 then by proposition 5.2

$$R(p^0, p^1, q^0) = \frac{f(q^0 p^1)}{f(q^0 p^0)}, \quad (B.10)$$

where $f$ is an increasing, linearly homogeneous and symmetric function. By axiom A12

$$R(p^0, p^1 + \tilde{p}, q^0) = R(p^0, p^1, q^0) + R(p^0, \tilde{p}, q^0), \quad (B.11)$$

or

$$\frac{f(q^0 (p^1 + \tilde{p}))}{f(q^0 p^0)} = \frac{f(q^0 p^1)}{f(q^0 p^0)} + \frac{f(q^0 \tilde{p})}{f(q^0 p^0)} \text{ using } B.10, \quad (B.12)$$

from which it follows that

$$f(q^0 (p^1 + \tilde{p})) = f(q^0 p^1) + f(q^0 \tilde{p}), \quad (B.13)$$

so that $f$ is a weakly additive function. Using the additivity of $f$, the following holds

$$\frac{f(q^0 (p^0 + \tilde{p}))}{f(q^0 p^1)} = \frac{f(q^0 p^0)}{f(q^0 p^1)} + \frac{f(q^0 \tilde{p})}{f(q^0 p^1)}, \quad (B.14)$$

which implies from equation B.10 that

$$\frac{1}{R(p^0 + \tilde{p}, p^1, q^0)} = \frac{1}{R(p^0, p^1, q^0)} + \frac{1}{R(\tilde{p}, p^1, q^0)}. \quad (B.15)$$

$R(p^0, p^1, q^0)$ satisfies equation B.15, axioms A3 and A12, then by a theorem in Eichhorn (1978)[page 157]^[5]

---

^[5] The theorem states that every function $F: \mathbb{R}^N_{++} \rightarrow \mathbb{R}_{++}$ satisfying $F(p, p) = 1$ (the identity axiom) and the properties

$$F(p^0, p^1 + \tilde{p}) = F(p^0, p^1) + F(p^0, \tilde{p}), \quad (B.16)$$

and

$$\frac{1}{F(p^0 + \tilde{p}, p^1)} = \frac{1}{F(p^0, p^1)} + \frac{1}{F(\tilde{p}, p^1)}, \quad (B.17)$$

can be written in the form
\[ R(p^0, p_i^1, q^0) = \frac{c_i^T p_i^1}{c_i^T p_0^1}, c_i > 0, i = 1, \ldots, N. \] (B.19)

Using equations B.10 and B.19 gives

\[ R(p^0, p_i^1, q^0) = \frac{f(q^0 p_i^1)}{f(q^0 p_0^1)} = \frac{c_i^T p_i^1}{c_i^T p_0^1}, \] (B.20)

and the only way of reconciling B.10 and B.19 is by setting \( c = q^0 \) so that \( R(p^0, p_i^1, q^0) = \frac{p_i^T q^0}{p_0^T q_0^0} \).

The proof in the other direction is straightforward. ♦

Remark: Propositions 5.2 and 5.3 do not make use of A5, one of the initial six independent axioms introduced. However, it can be shown that linear homogeneity (A2) and transitivity (A15) imply dimensionality (A5).

Proof:

\[ R(p^0, p_i^2, q^0) = R(p^0, \lambda p_i^1, q^0) R(\lambda p_i^1, p_i^2, q^0) \] by axiom A15
\[ = \lambda R(p^0, p_i^1, q^0) R(\lambda p_i^1, p_i^2, q^0) \] by axiom A2
\[ = R(p^0, p_i^1, q^0) R(\lambda p_i^1, \lambda p_i^2, q^0) \] by axiom A2.

Also by transitivity

\[ R(p^0, p_i^2, q^0) = R(p^0, p_i^1, q^0) R(p_i^1, p_i^2, q^0). \]

Combining the above two expressions yields \( R(p^0, p_i^2, q^0) = R(\lambda p^0, \lambda p^2, q^0) \). ♦

Proposition 5.4 Suppose that \( R(p^0, p_i^1, q^0) \) satisfies axioms A1, A2, A3, A4, A6, A12 and A15'. Then \( R(p^0, p_i^1, q^0) = \Pi_{i=1}^N \left( \frac{p_i^1}{p_0^1} \right)^{\frac{1}{N}} \). Conversely, \( R(p^0, p_i^1, q^0) = \Pi_{i=1}^N \left( \frac{p_i^1}{p_0^1} \right)^{\frac{1}{N}} \) satisfies axioms A1, A2, A3, A4, A6, A12 and A15'.

Proof.\(^6\) By strong transitivity

\[ F(p^0, p_i^1) = \frac{\sum_{i=1}^N c_i p_i^1}{\sum_{i=1}^N c_i p_0^1}, \] (B.18)

where \( c_i > 0, i = 1, \ldots, N \). Conversely, every function \( F \) given by B.18 satisfies B.16, B.17 and \( F(p, p) = 1. \)

\(^6\)This method of proof is due to Balk (1995)
\[ R(p^0, p^1, q^0) = \frac{R(p, p^1, q)}{R(p, p^0, q)} \]
\[ = \frac{g(p^1)}{g(p^0)}, \quad (B.22) \]

where \( g(p) \) is defined as \( g(\bar{p}) = R(p, \bar{p}, q) \). By axiom \textbf{A4}

\[ \frac{g(Dp^1)}{g(Dp^0)} = \frac{g(p^1)}{g(p^0)}, \quad (B.23) \]

where \( D \) is a \( N \times N \) diagonal matrix with diagonal elements \( \delta_i \in \mathbb{R}_+, i = 1, \ldots, N \). Setting \( D = (p^0)^{-1} \) gives

\[ \frac{g(p^1)}{g(1_N)} = \frac{g(p^1)}{g(p^0)}, \quad (B.24) \]

where \( \frac{p^1}{p^0} = (\frac{p_{11}^1}{p_{11}^0}, \ldots, \frac{p_{NN}^1}{p_{NN}^0})^T \), and \( 1_N = (1, \ldots, 1)^T \). This means that \( h(p) = \frac{g(p)}{g(1_N)} \) is a multiplicative function. Furthermore \( h \) is an increasing function by \textbf{A1}. It follows by a theorem in Aczél (1966)[pages 39-41] that \( h(p) = \prod_{i=1}^{N} p_{ii}^{\alpha_i}, \alpha_i > 0, i = 1, \ldots, N \). Imposing \textbf{A2} gives \( \sum_{i=1}^{N} \alpha_i = 1 \), and \textbf{A6} yields \( \alpha_i = \frac{1}{N} \forall i \). Thus the return index is given by

\[ R(p^0, p^1, q^0) = \frac{g(p^1)}{g(1_N)} = h(p^1) = \prod_{i=1}^{N} \left( \frac{p_{ii}^1}{p_{ii}^0} \right)^{\frac{1}{N}}. \quad (B.25) \]

The proof in the other direction is straightforward to verify. ✶