This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: The Measurement of Durable Goods Prices

Volume Author/Editor: Robert J. Gordon

Volume Publisher: University of Chicago Press, 1990

Volume ISBN: 0-226-30455-8

Volume URL: http://www.nber.org/books/gord90-1

Conference Date: n/a

Publication Date: January 1990

Chapter Title: I. INTRODUCTION AND METHODOLOGY 1. Introduction and Summary of Findings

Chapter Author: Robert J. Gordon

Chapter URL: http://www.nber.org/chapters/c8309

Chapter pages in book: (p. 3 - 39)

1 Introduction and Summary of Findings

There is much to be said in favor of the recent impressive advances in quantitative techniques but to benefit from them there must be a real commitment to measurement. For want of commitment, we abide the lack of adequate data which greatly limits the usefulness of these new techniques. This sad state is a consequence of our dependency on data that have been collected for purposes other than economic analysis. Very few university economists are engaged in "making data" that will stand up under strong empirical analysis.

Theodore W. Schultz (1974, 1)

1.1 Introduction: Purpose of the Project and Statement of Main Result

This book describes the creation of a new set of price indexes for producer and consumer durable goods that differ fundamentally in methodology of measurement, sources of data, and empirical implications from the official U.S. government price deflators. In contrast to the official national income and product accounts' (NIPA) deflator for producers' durable equipment (PDE), the new index for exactly the same commodities increases at an annual percentage rate 2.9 percent slower over the full period 1947–83. When compounded over time, this implies that the ratio of the new PDE deflator to the existing deflator on a base 1983 = 100 has a 1947 value of 286. When compared with this major difference for producers' durables, the difference between the new deflator for consumer expenditures on durable goods and the official consumer durable deflator is a smaller but still noteworthy 1.5 percent per annum. When compounded over time, the ratio of the new consumer durable deflator to the existing deflator on a base 1983 = 100 has a 1947 value of 174. Since all data on the output of durable goods in the United States depend directly on the validity of durable goods price deflators, the new indexes substantially alter, for the entire postwar

period, the official statistical picture of the behavior of investment, the capital stock, investment/output and capital/output ratios, the behavior of the prices of capital goods relative to consumer goods, and the growth rate of productivity in durable goods manufacturing as contrasted to other parts of the economy.

For instance, the annual rate of growth of real equipment investment during the period 1947–83, estimated to be 3.2 percent per annum in the official data, is a much higher 6.1 percent when the official deflator is replaced by the new index. The results overturn stylized facts that have reassured us about the relevance of standard economic growth models, particularly the rough constancy of the ratios of investment to GNP and the nonresidential capital stock to GNP. Even with no correction at all for possible errors in NIPA deflators for nonresidential structures, the new data indicate a continuous and rapid rise in the investment-GNP ratio over the entire postwar period. The ratio of equipment investment to GNP, as well as the ratio of equipment capital to GNP, almost triples in our new results. Another implication is a marked divergence of price changes for durable goods, as measured in this study, from price changes for nondurable goods and services.

1.1.1 Sources of Secular Drift of Alternative Relative to Official Price Indexes

The radical differences between this new PDE deflator and the official series reflect both theoretical and methodological innovations. I take seriously the old economic idea that capital goods are valued by their marginal products and extend it slightly to valuation by contribution to a firm's net revenue, that is, revenue less operating cost. Two capital goods are equivalent if they earn the same net revenue, in which case they will sell for the same age-adjusted price on the used asset market. This notion implies that an attempt must be made to value not just the change in performance of new models versus old models, but also changes in energy use and repair costs. Several of the product categories that exhibit the greatest differences between the alternative and the official deflators, particularly the new indexes for commercial aircraft and electric utility generating equipment, treat changes in operating characteristics explicitly. An important validation of the basic theoretical approach comes from the market for used aircraft; a price index that compares new and old models with quality relatives established by their subsequent value on the used aircraft market declines even more rapidly over the postwar period than an alternative index based on an explicit attempt to estimate net revenue, and the relative prices of individual models on the used aircraft market are highly correlated with our estimates of their ability to generate net revenue. Adjustments for energy use are applied as well for automobiles, railroad equipment, and almost all major

consumer appliances. Adjustments for both energy use and repair frequency are made for television sets.

Most of the innovative empirical conclusions of this study stem, however, not from operating efficiency adjustments, but from a consistent implementation of current theoretical practice. The agency responsible for the NIPA data, the Bureau of Economic Analysis (BEA), already has introduced for the period extending back to 1969 a hedonic price index for computer systems; the BEA index is confirmed with new data and extended back fifteen years earlier. The widely accepted principles used by the BEA for computers are extended to communications equipment, the single most important category of PDE. The single largest body of data, however, comes from the Sears catalog and simply involves the application of the standard Bureau of Labor Statistics (BLS) specification technique, without explicit efficiency or reliability adjustments, carefully and consistently over a long period of time.

The results here go beyond current NIPA procedures, however, in order to repair the undesirable effects of the current official Paasche index number methodology for price deflators, in which the growth rate of an aggregate price index over a given time interval is highly sensitive to the choice of the base year, and in which products like computers with rapidly declining relative prices are severely underweighted in years prior to the base year and increasingly overweighted in years subsequent to the base year. Here, current NIPA practice is running well behind the widely accepted theoretical state of the art by failing to utilize the much superior Törnqvist (1936, 1937) index number formula, identified by Diewert (1976) as one of the class of what he calls "superlative" index numbers.[1] Throughout this study, the Törnqvist formula, in which nominal value share weights for aggregation are allowed to change each year, is used to aggregate price indexes across classes of products, for both producers durable equipment and consumer expenditures on durable goods. It should be emphasized that all statements contained in this chapter and in the rest of the book regarding "drift," that is, the growth rate of the ratio of the alternative index to the official price index for the same concept, always use identical weights for the numerator ("alternative index") and denominator ("official index") of any such ratio. Thus, the results on secular drift entirely reflect differences in the growth rates of the underlying price indexes and do not reflect differences in weighting schemes at all. Later in this chapter, results are exhibited on the drift of these new

1. What Diewert called "superlative" index numbers were those that provide a good approximation to a theoretical cost-of-living index for large classes of consumer demand and utility function specifications. In addition to the Törnqvist index, Diewert classified Irving Fisher's "Ideal" index as belonging to this class. Diewert's contribution is placed in the context of theoretical research on cost-of-living indexes by Triplett (1988) and is updated in Diewert (1989a).

indexes relative to the official indexes using three alternative weighting schemes, the Törnqvist formula, the NIPA method with 1982 weights, and the NIPA method with 1972 weights. The annual rate of drift of 2.86 percent for the PDE deflator refers to the Törnqvist version of the new index relative to a Törnqvist aggregation of the price indexes that underlie the NIPA deflator for PDE.

There is a need to stress not only what this study does demonstrate, but also what it does not. No claim is made that there is a consistent upward bias in the full range of government price indexes. In particular, there is no necessary conflict between these new results and the longstanding claim by Triplett (1975, 1988) that the overall bias in official price indexes is just as likely to be downward as upward. Triplett's claim has always been explicitly with reference to the CPI as a whole and does not conflict with the new results showing a major upward bias for durable goods in the PPI, and to the lesser extent for durable goods in the CPI. We cannot assess the direction of bias in the full CPI, and hence we cannot assess Triplett's claim, since we have no new data for the prices of consumer nondurables or services. This book makes no statement at all about the existence or direction of possible errors in price indexes for products other than durable goods. If the official price indexes for nondurable products are accurate, then these results imply major revisions in the relative prices and quantities of durables and nondurables, in the shares of durable goods in total output, and in the growth rate of capital input relative to total output. If, in contrast to the upward bias in the official price indexes for durable goods implied by these results, the official price indexes for products other than durable goods are biased downward—as has been suggested recently by Triplett (1988) for consumer nondurable goods and services—the suggested revisions in the relative prices and quantities of durables and nondurables would become even larger.[2] On the other hand, any finding in future research that the prices of consumer nondurables and services are on balance biased upward would reduce the

2. Triplett intends his assessment of downward bias to apply to the CPI as a whole and qualifies it to apply to the recent period: "However, because a number of large CPI components appear quite clearly downward biased, I suspect that the CPI has, if anything, understated inflation in the last several years" (1988, 67). Triplett's list of components claimed to be biased downward includes one type of consumer nondurables (clothing), two types of consumer services (housing and restaurant meals), and one type of durable good (new automobiles). He admits that there is upward bias in the CPI for used automobiles and states that evidence on other consumer services is inconclusive. His claim regarding new car prices does not necessarily conflict with my results since the major reason for his claim is that he disagrees with the BLS/NIPA treatment of safety and antipollution devices as quality improvements rather than price increases, whereas I accept the BLS/NIPA approach. Note in table 1.1 below that I find the PDE deflator for automobiles to be biased downward during 1973–83, owing to the erroneous CPI measure of the relative price of used to new cars, whereas that CPI error creates an upward bias in the consumer durable deflator. The opposite direction of these automobile biases results from the fact that on balance the business sector sells used cars to the household sector, resulting in a negative weight on used cars in PDE but a positive weight in consumer expenditures on durable goods.

magnitude of the suggested changes in the relative prices and quantities of durables and nondurables, but would correspondingly increase the magnitude of the estimated bias in absolute prices and quantities for GNP as a whole.

1.1.2 Sources of Data for the Alternative Deflators

Almost all the data on which the results are based have been newly collected for this study, and the final price indexes for producers' and consumer durables are weighted averages of 25,650 separate price quotations selected to satisfy the twin criteria of careful adjustment for quality change and measurement of transaction rather than list prices.[3] There is no overlap between the sources of the official deflators, based almost entirely on components of the consumer and producer price indexes, and the current project, which assembles and analyzes data from numerous sources.[4] Chief among these the following:

1. Mail-order catalog prices;
2. Price observations from *Consumer Reports* (*CR*) articles on household appliances, typewriters, outboard motors, and home power tools, as well as data on operating characteristics, repair records, and reliability for some products;
3. Unit value data from the census *Current Industrial Reports;*
4. Data from the Department of Transportation on prices paid by airlines for commercial aircraft, combined with data from the same source on the subsequent operating cost characteristics of specific aircraft models and from the Avmark Company on price quotations for used commercial aircraft;
5. Department of Energy data on equipment costs and operating characteristics of almost all the nonnuclear electric utility generating plants in the United States;
6. Data on prices and characteristics of mini- and mainframe computers from Phister (1979) and *Computerworld* magazine and also, for the last few years, from *PC* magazine on mail-order prices for personal computers and peripherals;
7. Data on the prices of new and used automobiles from the *Red Book* and the *NADA Official Used Car Guide;*

3. To use the technical terminology of chap. 12, this count includes only the data used in the sixteen "primary" categories of PDE. It is slightly overstated because it includes price observations for eleven Sears catalog indexes that are used twice, but it is also understated by excluding all the price observations collected for individual models in *Consumer Reports.* In addition, 8,000 additional observations were collected but were rejected for use in the final index. This total consists of about 4,500 unit values developed for an earlier unpublished study but not included in the analysis of chap. 11, and about 3,500 automobile list price quotations.

4. The one exception to the statement about overlap is that the CPI for new automobiles receives a weight of half in our automobile index. The new index for computer systems is entirely independent of data collected for the BEA computer deflator.

8. Data on the prices of new and used wheel and crawler tractors from the leading price guide for used tractors;
9. Data recently collected from AT&T sources by Kenneth Flamm (1989) on the prices of telephone transmission and switching equipment; and
10. Data collected from railroad industry magazines and other sources on the prices and quality characteristics of railroad locomotives and freight cars.

The primary focus of the study is on the detailed analysis of price data for individual products. This daunting task has in the end become a fascinating one, because the chapters of this book make up a detailed history of technological change and quality improvement for many types of durable goods. For each major product, in addition to incorporating adjustments for quality change to the maximum extent that the data allow, two additional exercises are central ingredients in establishing the credibility of the results. First, price comparisons are developed of "closely similar" models over long periods of time. If the new index says that there has been no price change over a particular interval for a specific product, while the PPI says that the price has doubled, it is reassuring to be able to provide a specific example of a model of roughly the same quality characteristics and size that has remained roughly unchanged in price. This technique is carried out mainly in chapter 10, where many such comparisons are developed for the catalog data. The methods of data collection used by the BLS to compile the PPI not only prevent outsiders like myself from carrying out any such comparisons of similar models on BLS data, but even prevent BLS specialists from accomplishing this task, particularly for the earlier part of the postwar period. Another important ingredient in the credibility of the results is the consistent attention throughout the book to unmeasured aspects of quality change. Below are summarized the numerous aspects of unmeasured quality change that, while difficult to quantify, imply that this study, however radical its results, still fails to capture the full range of quality improvements.

1.2 The Importance of Accurate Price Measures

Among the most important goals of aggregate economic policy are the achievement of full employment, price stability, and a high rate of economic growth. However, the evaluation of a nation's progress in achieving two of the goals, price stability and economic growth, depends on the accuracy with which prices are measured. Estimates of the overall rate of inflation and of changes in relative prices depend, of course, on accurate measurement of the prices of individual commodities. And estimates of the level and rate of growth of aggregate output and its components depend on the accurate deflation of current-dollar sales and production data. Because of their crucial role in output measurement, correct price quotations are also

absolutely necessary for data on labor productivity (output per manhour) and the contribution of capital input and technical change to economic growth.

The scope of this book is limited to the prices of durable goods to the exclusion of nondurable goods and services, both to keep the scale of the research project at a manageable size and to concentrate on the segment of the economy where official price data are most vulnerable to inaccuracy, owing to the heterogeneity and changing specifications of durable goods. Why is a new set of price indexes necessary to replace the official government deflators for durable goods? The official indexes compiled by the BLS and other government agencies have been subjected to a steady barrage of criticism during the three decades that have passed since the publication of the "Stigler report" (NBER 1961). But, despite this fact, official deflators for durable goods today are based on essentially the same methodology as in the 1950s, with the new hedonic index for computers representing a rare exception to this statement.

One reason for the lack of progress is that most economic research on price measurement has been either too general or too specific to convince government officials that existing official deflators are very seriously in error. One research approach has been an exploration of the implications of casual empiricism, for example, the substitution by Jorgenson and Griliches (1967) of the CPI durable deflator for the producers durable equipment deflator (which has a faster rate of growth) on the grounds that, since "expenditures on the wholesale price index are less than those on the consumers' price index, adjustments for quality change are less frequent and less detailed (263).[5] This is obviously too general an approach to be convincing, since most types of producers' durables are not covered by the CPI, while the CPI itself has been subject to criticism regarding many of the same issues in the correction of quality change that raise problems for the PPI.

Most price research since the Stigler Committee report (NBER 1961), however, is at the opposite extreme, consisting of detailed regression studies of quality change in individual products (perhaps a dozen products in all), with relatively little attempt to evaluate the general implications of the studies for price, output, and input measurement. The concentration on one or two specific products is not surprising, since most tests have been conducted by professional econometricians more interested in refining their methodological techniques than in applying their conclusions to broader issues.[6] One indication that the authors of the first round of hedonic studies in the 1960s were mainly interested in econometric methodology rather than substantive implications is the way price research has essentially "died out"

5. The wholesale price index (WPI) was renamed the producers price index (PPI) in 1978. I have tried to purge the old WPI nomenclature from the book everywhere except in quoted material.

6. The most notable exception is the series of books and papers by Irving B. Kravis and Robert E. Lipsey on foreign trade prices, especially Kravis and Lipsey (1971).

since then. With the exception of research on computer prices, there have been remarkably few studies since 1971 requiring citation for the development of new product-level price indexes.

There are, of course, a few exceptions to this characterization. Since his earliest work on agriculture, Griliches has always been interested in extracting general implications from specific product studies of quality change using the hedonic and other methodologies. In addition to his work with Ohta on automobiles in the 1970s and 1980s, Griliches has written several follow-up pieces (including Griliches 1979) that supplement his well-known (1971) retrospective on hedonic research. Triplett is another leading figure who has written numerous surveys and evaluations of research on price measurement and quality change, most recently in his (1989) comprehensive survey of research on computer prices. Griliches and Triplett have generally taken opposing viewpoints about the importance and direction of price measurement bias due to problems in measuring quality change. In his first overall evaluation (1964) and again in his work with Jorgenson (1967, 1972), Griliches indicated the likelihood of a major upward bias in official measures of price change, while Triplett has consistently asserted that, "if individual components show both upward and downward errors, the overall error may go either way."[7] More recently, Triplett (1988) has argued that the reasons to believe that the CPI is downward biased outweigh those that work in the opposite direction.

The decisiveness of the results in this book stand in contrast to Triplett's (1971a, 1975) demonstration, on the basis of empirical studies predating the research reported here, that some hedonic regression studies indicate a downward rather than an upward bias in the CPI. As indicated above, these new results do not necessarily imply an upward bias in the official deflator for GNP as a whole, because I do not study prices of consumer nondurables and services. But most of the research reviewed by Triplett in the early 1970s, and indeed the great majority of price research carried out since the 1961 Stigler report, concerns durable goods. Within the limited sphere of durable goods, as measured in the PPI and in the household appliance and television components of the CPI, the results reported in this book decisively overturn the ambiguity of Triplett's appraisal. The bias in the durable components of both the CPI and the PPI is consistently in an upward direction. The annual percentage rate of "drift" of the ratio of my alternative price indexes to the official deflators is predominantly negative, with positive drift in only five of the sixty-six cells for subintervals in table 1.1 (which shows twenty-two PDE categories over the subintervals 1947–60, 1960–73, and 1973–83). This consistency in the direction of bias

7. The preceding sentence in the original is, "Notice that these studies do not point to a positive conclusion: we have not proved that price indexes are biased either upward or downward; rather, they establish only that the proposition that indexes are systematically upward biased is not conclusively confirmed by the available evidence" (Triplett 1971a, 31).

in the official PDE deflator is reassuring and shows that my results do not depend overwhelmingly on specific products or technologies. Very large rates of drift are identified for one or more of these subintervals for such diverse products as computers, telephone switching equipment, calculators, electric generating boilers and turbines, television sets, commercial aircraft, stationary air compressors, centrifugal pumps, and diesel engines.

1.2.1 Important Issues to Which Price Measurement is Related

While only a small minority of recent economic research has directly contributed to the improved measurement of prices, the conclusions of a much broader range of studies are affected by new estimates of durable goods prices.

1. Economic statisticians and historians have devoted much attention to accelerations or decelerations of output and productivity growth during various historical eras, for example, the acceleration in output growth after 1948 from the 1929–48 rate, or the productivity growth slowdown that started in the late 1960s and continues to this day. But few economists have attempted to determine the contribution of measurement error to these apparent historical events. If, for instance, the BLS made more detailed and widespread adjustments for quality change after 1947 than before (as seems to have been the case), much of the acceleration of productivity growth after World War II may have been caused by the improved statistical techniques of the BLS, rather than any event in the "real" economy outside Washington, D.C.[8] While the data in this book do not extend before 1947, and cannot answer questions about changes in growth rates that occurred on or about that date, the new measurement techniques have been applied consistently within the 1947–83 period and should thus eliminate spurious movements in durable goods output caused by intraperiod changes in the BLS deflation techniques. The only similar broad-scale effort to correct price data for a historical subperiod was Rees's (1961b) demonstration that the apparent cessation of real wage growth between 1890 and 1914 was an artifact of incorrect cost-of-living data and was eliminated by improved measurement techniques. Neither this study nor Rees's, of course, has anything to say about the true growth of prices and output before or after the historical subperiod under investigation. When we read that, in the forty years after 1865, the life of a rail increased from two years to ten, and the car weight it could bear from eight tons to seventy, we suspect that existing capital goods

8. Mills (1936) proves a perceptive early review of problems of price research that anticipates some of the themes of the Stigler report, written twenty-five years later. He cites (294–95) an engineering test showing that agricultural machinery in use in the United States in 1932 was 70 percent more efficient than that in use in 1910–14 and speculates that substantial improvements in the quality of "motor cars" had been missed in the WPI. However, Mills does not anticipate the potential contribution of the hedonic and other methods to improve the measurement of quality change and ends on a pessimistic note: "All commodities except a restricted number of staples are subject to quality changes that may not be reduced to quantitative form" (308).

deflators significantly overstate quality-corrected inflation in capital goods prices, and understate the increase in capital input, all the way back to the dawn of the industrial age.[9]

2. The growth rate of output minus the contributions of growth in labor and capital inputs is the growth rate in total factor productivity (TFP), often taken to represent the contribution of technical change or, more generally, advances in knowledge to economic growth (the latter contribution is sometimes called more neutrally the "residual" or "measure of our ignorance"). TFP growth is affected by errors in price measurement; in a steady state with constant growth rates of output and investment, the effect of price measurement errors on TFP growth is equal to the estimated bias in the deflator for investment goods times the difference between the share of capital in total income and the share of investment in total output. Assuming that the income share of capital exceeds the investment share of output, then the correction of any upward bias in price indexes for investment goods implies slower growth in TFP. However, this does not mean that advances in knowledge are less important. Slower growth of the "residual" is then more properly interpreted as the portion of labor productivity growth that is not contributed by technical change "embodied" in capital input. Solow's (1960) embodiment hypothesis held that productivity gains result, in large part, from the installation of capital goods that embody new technologies. Some studies, most recently Maddison (1987, 662–64), have interpreted this hypothesis to imply that explicit adjustments should be made to the BEA measures of the capital stock to account for embodied quality improvements.[10] However, it was shown long ago by Jorgenson (1966) and others that the Solow vintage argument could not contribute much to the explanation of economic growth, since the embodied quality improvements had the effect of raising both the growth rate of output (through faster investment growth) and the growth rate of capital input, leaving a minimum net effect on the growth rate of total factor productivity that depended on the small difference between the share of investment in output and the share of capital income in total income. The results in this book confirm Jorgenson's theoretical argument, in the sense that the very substantial revisions to the price deflators for PDE in the end "explain" only 13 percent of the 1.34 percent annual growth rate in conventionally measured total factor productivity over the postwar (1947–83) period. As in Jorgenson's model, the explanation of total factor productivity here is relatively small because the new deflators not only raise the growth rate of capital input, but also raise the growth rate of output.

3. Closely related to the previous paragraph are the behavior of the capital/output and investment/output ratios. The former exhibits puzzling

9. The source on rail quality is Rosenberg and Birdzell (1986, 247).
10. Some of this literature has been reviewed recently by Oliner (1988).

historical behavior, declining precipitously between 1929 and 1948, declining slowly from 1948 to 1966, and then reversing course to exhibit a slow increase from 1966 to 1983. The revised durable goods price data in this book suggest that the capital/output ratio rose dramatically rather than fell after 1948. These results naturally raise questions about the effect of possible measurement errors on its behavior in prior historical epochs.

4. A stimulus to several early hedonic regression studies of price and quality change was the need for accurate measures of relative prices for studies of the demand for durable goods. Indeed, the first such study of refrigerator prices (Burstein 1960) appeared in a book entitled *The Demand for Durable Goods*. The same need formed part of the motivation for Griliches's early work on automobile prices, including that in the Stigler report and that included in his well-known (1964) survey of measurement problems. Probably the best-known study in which the aim of the hedonic price research is the provision of a relative price measure for the purpose of demand analysis is Chow's (1967) paper on computer prices. Chow's work influenced me to urge (1971c) that the NIPA incorporate a hedonic price index for computer equipment, advice that was rejected at the time but then later accepted with a fifteen-year lag.[11]

5. Broader issues in macroeconomics that hinge on changes in the relative price of investment and consumption goods have been identified by R. A. Gordon (1961). Of more immediate interest may be the implications of this research for the terms of trade of different groups of nations. Assuming relatively accurate measurement of the prices of crude commodities exported by less-developed countries, the results reported here imply a substantial upward bias in the secular growth rate of the terms of trade of the developed world relative to the less-developed world, which imports much of its capital equipment.

6. Related to the previous topic is the dependence of measures of the user cost of capital on accurate indexes for the price of capital goods relative to output. The demand for a capital good depends on its user cost, which in the usual Hall-Jorgenson (1967) formula includes a term for the rate of change of the relative price of capital goods. The results in this study imply a substantial decline in the relative price of capital equipment, in contrast to the roughly constant relative price indicated in the NIPA over the period 1948–82. No new price indexes for structures are developed, and so the new results apply only to the equipment component of fixed investment (the nominal share of PDE in fixed nonresidential investment was 60 percent in 1948 and 70 percent in 1987).

11. One unfortunate side-effect of the fifteen-year lag is that the BEA would have discovered much earlier the pitfalls of its fixed-base-year Paasche index number formula. If a computer price index had been incorporated into the NIPA in the mid-1970s, the implicit GNP and PDE deflators would have been biased toward an increasingly severe understatement of inflation as a result of the fixed 1972 base year, used in the NIPA until the 1986 benchmark revision.

7. Studies of the payoff to research and development (R&D) are hamstrung by the misallocation of the fruits of research implied by the official price deflators for durable goods. Airlines do no research but "receive credit" for the R&D accomplished by the airframe and aircraft engine manufacturers. Electric utilities receive credit for the R&D performed by the makers of turbo-generator systems. The telephone communications industry receives credit for the achievements of the companies that manufacture telephone transmission and switching equipment. The approach here to the measurement of quality change creates a greater upward revision of the growth rate of output in industries where technical change has been rapid than in industries where it has been slow. This approach stresses the symmetric treatment of research on performance and operating attributes of capital goods. If R&D investment is shifted from making aircraft go faster to making them more fuel efficient, that R&D is still productive if it yields improvements in the ability of the new-model aircraft to earn profits for airlines. While the new theory calls for universal adjustments for improved energy efficiency that is not accompanied by proportional increases in the cost of durable goods, such adjustments are feasible only for a subset of products, and the improvements for other products that are not treated explicitly stand as an example of unmeasured quality change in this study.[12]

8. Accurate measures of the relation of transaction to list prices, that is, of discounts and premiums, are necessary to evaluate Stigler's contention that "it is not possible to make a direct test for price rigidity, in part, because the prices at which the products of oligopolists sell are not generally known. . . . Nominal price quotations may be stable although the prices at which sales are taking place fluctuate often and widely" (Stigler 1947). If discounting and premiums are important, vary procyclically, and are not adequately measured in present official statistics, existing national accounts may understate cyclical fluctuations in prices and overstate fluctuations in real output. To test this hypothesis, all price data collected for this book are either true transaction prices or close proxies, for example, prices of late-model used cars and tractors. This issue resurfaces in chapters 10 and 11.

1.3 The Scope of This Study

The original intention of the research project was simply to collect previous research and summarize its implications for the two major issues of "quality-adjustment" and "transaction-list" bias in the official producers' durable deflators. Primary attention to producers' durables was motivated by my longstanding interest in the measurement of capital input and in the

12. Additional research issues in the cross-industry allocation of the returns to R&D are considered by Griliches (1979, esp. 104–5).

sources of economic growth. After the initial results of the collection and summarization exercises were published (see Gordon 1971b), it became apparent that a new, larger-scale study was needed. Previous regression studies of the quality-adjustment problem were available for only short sample periods, in some cases were defective in methodology, and in almost all cases relied on list prices rather than transaction prices. Previous evidence on changes in unit values was unsuitable for application to the transaction-list problem, because inadequate attention had been paid to shifting product mix across size classes as an alternative source of fluctuations in unit values. The only solution was an effort to collect a large amount of new data for a long period of years and to perform an analysis that attempted to correct for quality change as thoroughly as possible, and to correct unit value data for shifts in product mix.

After the first draft of this project was completed in 1974, the scope of the study widened further. The stimulus for an explicit treatment of changes in energy efficiency came from the 1974 and 1979 oil shocks. Detailed case studies on aircraft and electric utility generating equipment (the latter treated only sketchily in the 1974 version) were added toward this end, and new evidence on changes in automobile fuel consumption was incorporated into that chapter. The criticism that the first-draft results were too dependent on mail-order catalogs led to the development of parallel evidence from *CR* for numerous appliances and other products, including typewriters, outboard motors, and home power tools, which results in a significant improvement in the robustness and credibility of my conclusions. Further, the scope of the study was broadened to include more "large" equipment that is sold directly to firms rather than to households. The improved evidence on electric utility generating equipment and the new evidence for the aircraft, computer, telephone, and railroad categories helps address this earlier criticism.

The following basic ground rules were set down to delimit the scope of this study.

1. The basic aim was to create a new index suitable for replacing the existing deflator for PDE in the NIPAs. Thus, the particular commodities eligible for data collection were those that are a component of the NIPA deflator for PDE.

2. Nineteen forty-seven was selected as the initial year of the sample period, since almost no capital goods were included before then in the PPI (on which the NIPA deflator is based). While an extension of the study back to 1929 or earlier would have been interesting and potentially important, the pre-1947 years carry the disadvantage for the present purpose that no estimates of official deflator bias are possible for individual commodities for those years. Any future study of the pre-1947 period will have to follow Rees (1961b) by developing price indexes from scratch, using available information on weights from the Census of Manufactures and other sources, since there is no product-by-product set of NIPA weights that can be used as

in this study. The terminal year of 1983 was chosen arbitrarily to bring the project to a halt, and work continued throughout the 1984–88 period to update to 1983 all the many data sources listed above.

3. Data were eligible only if they were obtained from sources independent of the PPI or CPI. The only exception is the inclusion of the CPI for new autos as part of the final auto index developed in chapter 8.

4. Data were eligible only if they could reasonably be interpreted either as transaction prices or as close proxies for transaction prices. This includes catalog prices, unit values, *CR* data, which after 1959 are based on market surveys of actual prices paid, buyers' prices for aircraft, electric generating equipment, telephone equipment, railroad equipment, and late-model used prices for automobiles and tractors.

Once the data were collected, they were adjusted for changes in quality by two basic methods. In the case of commodities where numerous observations were available in each year, hedonic regression studies were performed. Products for which hedonic regression estimates were obtained include electric generating equipment, computer processors, refrigerators, room air conditioners, washing machines, new and used automobiles, new and used wheel and crawler tractors, hot water heaters, and outboard motors. Given the fact that almost all the hedonic regression studies cover the full 1947–83 period and are thus based on a much longer span of time than the first generation of hedonic studies performed in the 1960s, it is likely that there are more "product years" covered in the hedonic regression studies of this book than in all previous hedonic research on durable goods combined.

The second basic method of quality adjustment was the conventional specification method, mainly used for the mail-order catalog items, where (with the exceptions of refrigerators, air conditioners, washing machines, hot water heaters, and outboard motors) only a few observations were available in each year for those products included in the study. Catalog specifications and illustrations were examined carefully to limit price comparisons to the subset of models that were absolutely identical in any given pair of years. An innovation in the discussion of both the hedonic regression results and the catalog indexes is the cross-checking of the hedonic price indexes by comparisons of "closely similar" models over long spans of time. This is done for automobiles with the aid of data on the value of specific options supplied by General Motors; in chapter 10 on the catalog data for almost thirty separate products; on diesel engines through correspondence with manufacturers; and on home appliances and several other products by developing parallel price indexes from *CR*.

The *CR* indexes are a relatively recent addition to the project and have proved to be so useful that they are the basis for all conclusions regarding home appliances. One of the most innovative and interesting aspects of the book is the set of explicit adjustments for energy efficiency and repair frequency of home appliances based on the *CR* data. Explicit energy

adjustments are also carried out for individual pairs of aircraft models, electric generating equipment, automobiles, and railroad locomotives. Examples of products where improvements in fuel efficiency have occurred, but where no explicit adjustment is made, are tractors and diesel engines.

After all the individual product price indexes were developed, they were weighted together into deflators for the twenty-two categories of PDE and then for PDE as a whole. Weights within the twenty-two categories are identical to those used in the NIPA deflator for the same products, but across the twenty-two were calculated with the Törnqvist index number methodology discussed above rather than with the flawed implicit deflator method currently used in the NIPA. The only exception to this statement is that within the office machinery category computer and noncomputer equipment is aggregated with the Törnqvist formula. Chapters 6 and 12 show the substantial differences that emerge when the Törnqvist method is used in place of the implicit deflator method of aggregation.

A portion of the weight in the NIPA deflator for PDE is allocated to official price indexes for goods purchased not only by businessmen, but also by consumers. Chief among these are automobiles, large and small household appliances, radios, and television sets. The set of alternative price indexes developed for these products in this book can be used not just as a component in the alternative PDE deflator, but also as the basis for computation of an alternative deflator for consumer expenditures on durable goods. From the total of 105 unduplicated product indexes developed in this book for use in the alternative PDE deflator, a subset of eleven product indexes is used in the alternative deflator for consumer durables. Because nominal sales data can be easily obtained for these products, aggregation for consumer durables uses the Törnqvist formula throughout, in contrast to PDE, where NIPA weights are used within group categories and the Törnqvist formula is the basis for aggregating across those categories.

1.4 Summary of the Results

1.4.1 Methodological Innovations and Conclusions

In the deflation of durable goods, the proper criterion for quality adjustment is to consider as identical two different items if their ability to produce services for consumers, or their ability to generate net revenue for producers at a fixed set of output and variable input prices, is identical. Leaving aside changes in energy efficiency and repair frequency, there is no difference in principle between deflation based on equating goods in terms of their ability to generate net revenue and deflation based on equating goods in terms of their cost, as long as both concepts are measured for the same units, for example, the ability of a computer to perform calculations. Because firms may discover technological improvements that improve the net-revenue

generating ability of a capital good by an amount greater than the effect of the improvement on the selling price of that capital good, the theory of quality adjustment must incorporate such changes, called "nonproportional" quality improvements in chapter 2.

A simple method for carrying out such an analysis is to calculate the "real" price change implied by the introduction of a new model as the percentage difference for the new and old models in the selling price and the ability to generate net revenue at a fixed set of output and variable input prices. If the introduction of a new model raises the selling price and ability to generate net revenue in exact proportion, then this "proportional" quality change implies no change in the "real" price caused by the new model introduction, and the nominal price change for this product can be measured by the behavior of nominal prices for models that remain unchanged in the two adjacent time periods affected by the new-model introduction. If, however, the introduction of a new model raises the selling price by less than the ability to generate net revenue, as occurred, for instance, with the introduction of the jet plane, a "real" price decline is recorded, and this negative change in the real price is added to the recorded nominal price change for models that remain identical in the time period affected by the introduction of the new model.

This technique amounts to "crediting" manufacturers for improvements in knowledge and techniques that achieve any such "nonproportional" improvements in quality. The approach is a straightforward extension of the treatment of computers already incorporated into the NIPA, which treats nonproportional changes in computer calculating power relative to computer prices as a price decrease, that is, which essentially measures changes in computer prices per unit of calculating power rather than per "computer box."

Problems in the implementation of the conceptual framework are the subject of chapter 3, which compares the advantages and disadvantages of the conventional "specification" technique used in the official deflators with those of the newer hedonic regression technique. The general conclusions of the chapter run counter to the view held in the 1960s that regarded the hedonic technique as a panacea that would make the conventional "specification approach" obsolete. While both methods are identical in principle, in practice their advantages are complementary. The conventional method is inferior to the hedonic approach in its inability to provide values or "shadow prices" for changes in overall dimensions or other physical characteristics, for example, length or horsepower. But the hedonic technique falls short when there is a change in the relation between quality characteristics that are, respectively, included in and excluded from the regression equation. Inclusion of all relevant variables in the regression is prevented in practice by multicollinearity, and by quality changes that occur

on all models simultaneously and hence cannot be identified in a cross-sectional regression equation.

The emphasis in this book on improvements in energy efficiency leads to a further downgrading of the hedonic technique, since improvements in energy efficiency tend to be made on all models simultaneously (except in the interesting case of room air conditioners, where a variety of high-efficiency high-priced and low-efficiency low- priced units is for sale simultaneously). The hedonic technique is not well suited to measuring price for complex machines like tractors and automobiles, where handling, comfort, and other intangible characteristics are important attributes that are "missed" by the hedonic approach.[13] Instead, the hedonic approach is much more useful for simple products producing a homogeneous, easily measured output, represented in this book by the examples of electric utility generating equipment, hot water heaters, and outboard motors. Electronic computers fall somewhere in between the extremes of complex and simple, having two major quality attributes that can be measured relatively well, speed and memory, but also possessing other quality attributes that may cause shifts in the ratio of excluded to included variables in hedonic regressions, for example, multitasking ability and virtual memory.

Overall, a mix of measurement techniques is likely to be superior to the use of a single technique for all products. For products having a small number of characteristics that are valued by users and relatively easy to measure, and for which a substantial number of different models is available in each cross section for regression analysis, the hedonic technique may work well. In extreme cases where price variation is almost entirely explained by a single size or performance characteristic, for example, horsepower for diesel engines or number of circuits for telephone switching equipment, the price per unit of that quality characteristic, which amounts to a single-dimension hedonic function forced through the origin, may serve as an adequate price index.[14] For other products, and indeed for most of the Sears catalog data analyzed in this book, the standard BLS specification method may be superior to the hedonic method, since for many products the number of relevant quality characteristics seems to be too large relative to the number of available model observations in a single cross section to allow a regression to be run. The experience of developing Sears catalog indexes by the specification method for more than seventy products over thirty-seven

13. Triplett (1988, 57–58) also argues that "the automobile may be too complicated for hedonic studies." He concludes that "changes in the omitted characteristics can, without providing a clear signal to the investigator, swamp the effect of the included ones."

14. Mills (1936) speculates that a price index for automobiles based on price per pound or per horsepower might be a more adequate approximation to the unattainable true price index than the WPI methods used at the time he wrote. Triplett (1989) cites approvingly and uses as part of his final "best practice price index" an index of price per unit of computer calculating power developed by Flamm (1987).

years suggests that it is almost always possible to match models exactly along all listed quality characteristics without the need for linking or explicit cost adjustments. New models are usually introduced on a random basis in a particular year, replacing some models for a given product but leaving other models unchanged and available for price comparisons. This tacit endorsement of the specification method for some products is subject to the qualification that, even for low-tech products, there may be a tendency for quality improvements on new models to be "missed" in the linking procedure, just as matched-model (i.e., specification) indexes for computers show markedly slower rates of price decline than hedonic indexes.

Because the new price indexes often exhibit markedly different rates of growth than the PPI for the same product, I always ask a simple question of the indexes, one that seems rarely to have been asked in previous research. Does a comparison of the raw nominal price change at the beginning and end of the sample period for models of roughly similar quality "make sense"? Of course, quality does not remain unchanged over a span of two or three decades for most products. Nevertheless, for a surprising number of products in the Sears catalog study (chap. 10), it is possible to make such comparisons of "closely similar models" over periods of two decades or more. On average, the price change measured in such comparisons is smaller than that in the equivalent Sears catalog price index pieced together by cumulating year-to-year price changes for identical models. This implies that prices are more often reduced than increased relative to the quality change implicit in the new model introductions that are omitted from my matched-model Sears catalog indexes.

Whether the hedonic or the specification technique is used, however, it is often impossible to avoid a shift in the ratio of excluded to included quality attributes. This problem is inherent to the hedonic technique, as discussed above, but can also plague the conventional specification technique. The most important omitted quality changes in this study are improvements in fuel efficiency, repair frequency, and durability. With a full set of data from equipment users on the operating characteristics of new and old models, as in the studies of commercial aircraft and electric utility generating equipment, the value of such improvements can be calculated directly, although changes in durability may not be revealed until decades after the introduction of a new model. Analogous direct calculations can also be made for a few products, especially household appliances, where laboratory test data are available on energy efficiency, and where user questionnaires measure repair incidence. For other products, however, such supplementary data are not available, and the presumption must be that the new price indexes are on average biased upward by calculating fuel efficiency and repair adjustments for only a subset of the covered products.

Table 1.1 **Drift of the Ratio of Törnqvist Indexes, This Study and Corresponding PPIs, Over Selected Intervals**

NIPA Categories	Annual Growth Rates			
	Full Period of Data (1)	1947–60 (2)	1960–73 (3)	1973–83 (4)
Office, computing, and accounting machinery	−9.32	−3.94	−16.61	−6.83
Communication equipment	−5.84	−8.44	−2.89	−6.28
Instruments, photocopy and related equipment	−3.49	−3.18	−4.21	−2.97
Fabricated metal products	−1.80	−4.08	−1.28	0.49
Engines and turbines	−3.53	−7.16	−2.27	−0.44
Metalworking machinery	−1.15	−3.01	0.58	−0.96
Special industry machinery, n.e.c.	−2.48	−3.70	−1.01	−2.79
General industrial, including materials handling, equipment	−1.81	−2.87	−1.15	−1.29
Electrical transmission, distribution, and industrial apparatus	−2.11	−3.62	−1.89	−0.43
Trucks, buses, and truck trailers	−2.97	−5.74	−2.04	−0.59
Autos	−1.33	−5.02	−0.27	2.09
Aircraft	−8.29	−12.69	−7.48	−3.63
Ships and boats	−1.93	−3.17	−1.11	−1.39
Railroad equipment	−1.45	−1.24	−1.43	−1.76
Furniture and fixtures	−1.41	−2.72	−0.84	−0.46
Tractors	−1.35	−0.05	−1.28	−3.14
Agricultural machinery, except tractors	−0.70	−2.80	0.69	0.21
Construction machinery, except tractors	−1.63	−2.35	−1.63	−0.68
Mining and oilfield machinery	−1.63	−2.35	−1.63	−0.68
Service industry machinery	−3.15	−4.06	−1.91	−3.59
Electrical equipment, n.e.c.	−1.01	−2.56	−0.09	−0.20
Other	−1.99	−3.90	−0.30	−1.69
Total:				
Törnqvist	−2.96	−4.13	−2.44	−2.07
Implicit deflator, 1972 base	−2.90	−3.17	−1.88	−3.87
Implicit deflator, 1982 base	−1.97	−3.12	−1.20	−1.48

Sources: See the notes to table 12.2.

1.4.2 The Magnitude of Quality Bias in the Official Indexes

The overall results for PDE are summarized in table 1.1, which presents the difference in the annual rates of change of the new and official indexes for the twenty-two categories of PDE and for three different aggregations over the twenty-two categories. The overall "drift" in the ratio of the alternative to the official deflators, aggregated by the Törnqvist technique, occurs at the rate of 2.9 percent per year, which cumulates exponentially over thirty-six years to a ratio of the alternative to official price index in 1947 of 286 on a base of 1983 = 100. This Törnqvist-weighted average index for all PDE is compared to the official PDE deflator in figure 1.1. As shown at the bottom of table 1.1, an alternative PDE deflator aggregated over the twenty-two categories with the NIPA method and a fixed 1972 base

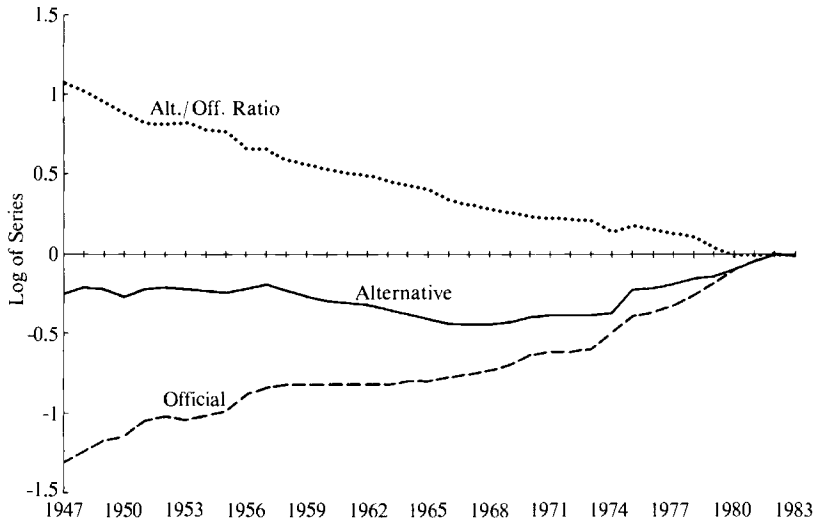

Fig. 1.1 Comparison of Törnqvist price indexes, this study and PPI, 22 categories, 1982 = 1.0

year yields a rate of drift for the full 1947–83 period that is almost exactly the same as the preferred Törnqvist index. But the flaws inherent in using a fixed base year lead the fixed 1972-base index to record too small a rate of drift before 1972 and too large a rate after 1972. In contrast, the fixed 1982-base index records too small a rate of drift throughout the 1947–83 interval.

The results displayed in table 1.1 indicate that the drift occurs at widely differing rates both across the twenty-two categories and across the three subintervals shown. Using either a fixed 1972 or fixed 1982 base year, the overall average rate of drift is smaller in 1960–73 then before or after. However, using the preferred Törnqvist method, the rate of drift is smaller in 1973–83 than in 1960–73. Several products stand out that have double-digit rates of drift for particular subintervals, for example, computers in 1960–73 and aircraft in 1947–60. Chapter 12 shows that there is a significant correlation between the growth rate of real investment already present in the NIPA data and the magnitude of the drift, lending credibility to the hypothesis that the measured drift is not a "fluke" but rather represents a systematic failure of the PPI to capture quality change adequately for those types of durable equipment having the most rapid rates of growth of nominal demand. This hypothesis is further supported by a regression analysis in chapter 10, which for the subset of indexes based on catalog data finds that the drift across sixty-eight products is significantly related to the technological complexity of the product and to the presence of electronic components,

Table 1.2 **Drift of the Ratio of Törnqvist Indexes, This Study and Corresponding NIPA Implicit Deflators for Selected Consumer Durables, 1982 Base—Over Selected Intervals**

NIPA Categories	Annual Growth Rates: Full Period of Data (1)	1947–60 (2)	1960–73 (3)	1973–83 (4)
1. Motor vehicles and parts	−1.71	−2.39	−1.69	−0.85
2. Furniture and household equipment:	−1.79	−2.52	−1.26	−1.55
2.1 Kitchen and other household appliances	−3.22	−4.39	−2.37	−2.83
2.2 Radios and TVs	−5.94	−9.07	−3.77	−4.69
3. Total consumer durables	−1.54	−2.21	−1.24	−1.05

Sources: See the notes to table 12.8.

but is not related at all to the closeness with which the individual Sears products are matched to individual components of the PPI.

Table 1.2 displays the estimated rates of drift for new consumer durable goods deflator relative to the official deflator. Since the new consumer durables index represents an alternative aggregation of a small subset of the individual price indexes originally collected for the purpose of deflating PDE, its coverage is substantially less complete. Fully half of consumer durable expenditures remain uncovered by the new indexes, and, at least for now, the existing NIPA deflators for these uncovered components are accepted as accurate. This limited coverage explains why the rate of drift for all consumer durables in row 3 of table 1.2 is smaller than the average of the rates of drift shown on rows 1 and 2, and also why the rate of drift on row 2 is less than the average of rows 2.1 and 2.2. As indicated above, the negative rate of drift for the consumer durable auto category is larger than for PDE, because the upward-biased official price index for used cars enters with a negative weight in the official PDE index but with a positive weight in the official consumer durable deflator. The rates of drift for the appliance and radio-television category may appear to be surprisingly high but are based on one of the most comprehensive and accurate parts of this study, the indexes developed from *CR*, which incorporate explicit energy efficiency adjustments for all major appliances, both energy efficiency and frequency of repair adjustments for television sets, and which since 1959 are based on market surveys of transaction prices.

1.4.3 Summary of Implications

This section briefly summarizes the main implications of the new PDE and consumer durable deflators developed here. Each of the implications is documented and discussed in more detail in chapter 12, which begins by treating issues involved in weighting together the new price indexes and in

converting them into new measures of real investment, the capital stock, and total factor productivity.

One of the longstanding stylized facts about the U.S. economy is the stability of the investment/GNP ratio. The ratio of structures and equipment investment to real GNP in the current official NIPA was almost identical in 1947 and 1983, with ratios of 11.4 and 11.1 percent, respectively. One of the most important implications of this study is that this stylized fact has been far from the truth for the entire postwar era. In contrast to the NIPA ratio, which remains constant, the new ratio of real equipment investment to real GNP *triples*. Even though this study does not contain any new results on structures prices or real investment in structures, the ratio of real structures plus PDE investment to real GNP doubles when the new data on real PDE investment are substituted for the official data.

Because the new data on the investment/GNP ratios overturn standard impressions of the process of economic growth, they raise new issues that will need to be discussed. First among these is the most obvious from a longer historical perspective. If the investment/GNP ratio doubled between 1947–83, what happened in the 100 years prior to that? Is the large quality bias in official price indexes for capital goods a new problem or an age-old problem? More important for the near future, how long will investment continue to grow faster than GNP? Will the rapid growth rate of real equipment investment eventually lift the U.S. economy out of the productivity doldrums? Answers to these questions will require substantial further research, using the results of this book as a point of departure.

The alternative and official real investment series are then converted into capital stocks, using a simplified version of the perpetual inventory method used by the BEA. The equipment capital stock is calculated separately for the twenty-two categories of PDE, using an algorithm that extracts from the present official data the implicit service life for each category. The alternative and official NIPA capital stock growth rates for equipment, that is, PDE, differ by about 3 percent per annum, roughly the same amount as the difference between the annual growth rates of the alternative and official equipment investment series. For the subintervals, the differences in the growth rates of the equipment capital stock are, respectively, 3.06, 3.25, and 2.63 percent per annum. The growth rate of the equipment capital stock differs most, then, in the middle interval (1960–73), in contrast to the growth rate of equipment investment, which differs most in the first period (1947–60). This difference in timing reflects the fact that this study starts only in 1947 and has little effect on the growth of the capital stock in the first few years after 1947, when the majority of equipment capital consisted of items purchased before 1947.

Effects of the new price data on the growth of the capital stock of equipment and structures together are much smaller than for equipment alone, as would be expected in view of the facts that there are no new

deflators for structures in this study, and that the stock of structures represents well over half the total capital stock. Despite the absence of new evidence on structures, however, this study has major implications for the behavior of the aggregate capital-output ratio. In the official data, the capital-output ratio is almost the same in 1947 and 1983, after displaying a substantial decline during 1947–67 and an increase during 1967–83. The capital-output ratio implied by this study is quite different. The increase in the capital-output ratio for equipment between 1947 and 1983 is 406 percent, compared to a much smaller 75 percent in the official data. For the aggregate of structures and equipment, the 1947–83 increase in the capital-output ratio is 44 percent, as contrasted to only 5 percent in the official data.

Just as the postwar price changes in producer and consumer durable goods have been much closer to each other with the new deflators than with the official deflators, the same conclusion applies to the shares of durable spending in GNP. A familiar fact in the existing NIPA is that the share of real consumer durable expenditures has increased substantially over the postwar period, by about two-thirds. Yet there has been virtually no increase in the NIPA share of real PDE in GNP. In contrast, the alternative deflators imply much greater 1947–83 increases in both investment shares, 178 percent for producer durables and 164 percent for consumer durables. As is true for the capital-output ratios discussed above, the denominator of these new investment/output ratios is not NIPA real GNP, but rather GNP recalculated with the alternative real PDE and consumer durable expenditure series in place of the official series for those two components.

The corollary of the rapid increase in the share of durable goods spending in the new data is, of course, a decline in the share of spending on categories of GNP other than durable goods. In contrast to the NIPA, in which the share of real nondurable consumption falls only by 0.7 percent from 1947 to 1983, in the new data that decline becomes 7.4 percent. The NIPAs already register a substantial 10.3 percent decline in the share of real GNP other than PDE and personal consumption expenditures, and this decline becomes 16.4 percent in the new data. In short, durable goods have been the most dynamic component of spending on GNP, and use of the new data to compute new industry productivity measures would indicate that the growth rate of productivity in durable manufacturing has been greatly understated, although by more before 1973 than since 1973.

An important question to ask of the new data is whether they contribute any explanation of the post-1973 productivity growth slowdown in the United States. It seems evident already from the results presented earlier in this chapter that there cannot be a major contribution, simply because the drift of the alternative relative to the official deflator extends over the whole postwar period. The 1947–73 annual rate of drift of the alternative relative to the official deflator (from table 1.1) is 3.3 percent, and a lower 2.1 percent for 1973–83. One would think that a finding of more rapid growth in

Table 1.3 **Effect of Alternative Durable Goods Deflators in Sources of Growth Calculation (annual growth rate over interval)**

	1947–60 (1)	1960–73 (2)	1947–73 (3)	1973–83 (4)	1947–83 (5)
Private GNP:					
Alternative	3.68	4.14	3.91	2.08	3.40
NIPA	3.35	4.02	3.68	1.82	3.17
Alternative—NIPA	0.33	0.12	0.23	0.26	0.23
Capital input:					
Alternative	4.60	5.73	5.17	4.97	5.11
NIPA	3.10	3.87	3.49	3.56	3.51
Alternative—NIPA	1.50	1.86	1.68	1.41	1.60
Capital contribution:[a]					
Alternative	1.15	1.43	1.29	1.24	1.28
NIPA	0.78	0.97	0.87	0.89	0.88
Alternative—NIPA	0.37	0.46	0.42	0.35	0.40
Private business labor hours	0.79	1.93	1.36	1.00	1.26
Labor Contribution:[b]	0.59	1.45	1.02	0.75	0.95
Multifactor productivity:[c]					
Alternative	1.94	1.26	1.60	0.09	1.17
NIPA	1.98	1.60	1.79	0.18	1.34
Alternative—NIPA	−0.04	−0.34	−0.19	−0.09	−0.17

Sources: See table 12.11.

[a]Equals capital input times 0.25.

[b]Equals labor input times 0.75.

[c]Equals growth in output minus capital contribution minus labor contribution.

capital input before 1973 would contribute to the productivity puzzle, since new data showing a greater slowdown in the growth rate of capital input would leave a smaller slowdown in total factor productivity ("Denison's residual").

This presumption, however, ignores the effect of the new deflators for producer and consumer durables on the growth rate of output; since the growth rate of output is increased more prior to 1973 than after, this implication of the new data deepens the productivity puzzle. The balance between the two effects, a greater speedup of the growth of both output and capital input before 1973 than after 1973, depends on the weight assigned to capital input as a source of growth. Unfortunately, existing estimates of capital's share differ widely. One can find weights for fixed capital input ranging from as low as 20 percent (Denison 1985, table G-2 for 1967) to 41 percent (Jorgenson and Griliches 1972, table 20 for 1962). Here, a compromise position sets the capital share at an arbitrary 25 percent. The results displayed in table 1.3, are set out in a format that allows them to be recalculated easily for any other assumed income share of capital.

Table 1.3 is arranged into five columns, corresponding to the three standard subintervals, the full 1947–83 period, and the extra 1947–73 subinterval that is of interest in discussions of the post-1973 productivity

growth slowdown. Comparing columns 3 and 4, we can assess the effect of the new price indexes on the growth of GNP, capital input, the contribution of capital to output growth (using the arbitrary 0.25 weight), and multifactor productivity (MFP), that is, the growth rate of output minus the contributions of capital and labor input. The row labeled "Alternative—NIPA" in the top section indicates that the effect of the new deflators on the growth rate of output is 0.23 percentage points in 1947–73 and 0.26 points in 1973–83, that is, almost exactly the same. The respective figures for capital's contribution to growth are 0.42 and 0.35 points. Using the same series on labor input to compute both alternative and NIPA versions of MFP growth, the new price deflators reduce the growth rate of MFP by 0.19 points before 1973 and 0.09 points after 1973. The MFP growth slowdown after 1973 is 1.61 points in the official data and 1.51 points in the new data. Thus, the new data contribute less than one-sixteenth (0.10/1.61) of the needed explanation of the post-1973 productivity growth slowdown. This conclusion is quite robust to variations in the arbitrary 0.25 capital share used in table 1.3; the explanation of the productivity slowdown is 0.08 points with a share of 0.15, 0.10 with a share of 0.25, and 0.13 with a share of 0.35.

1.5 What Is Wrong with the PPI?

The enormous difference between the alternative price indexes and the PPIs for the same products raises difficult questions for specialists at the BLS. Here, I can only speculate, for a fundamental limitation on anyone's ability to resolve this quandary is the confidentiality rule that prevents outsiders from inspecting individual price quotations. I corresponded in the late 1970s with BLS officials about the specific case of diesel engines, where the PPI shows prices more than doubling during 1947–72, while officials contacted at the leading firm producing the engines, which supplies data to the BLS, provided price schedules showing no change in price. Even in this specific case, BLS officials were unwilling or unable to provide a single example of a price change for a particular model number over that period and in their response offered only generalities about how in principle unit value data (the basis for my diesel engine index) could be flawed, without noting that my inquiry had been based not on unit value evidence but rather on explicit communications with one of their own price reporters.

As a second example, the BLS published an index for nonprinting electronic calculators over the period 1969–75 that fell only by about one-half, while Sears, *CR*, and everyone's personal recollections register a price decline by a factor of at least *ten* while at the same time a marked improvement in quality occurred over several dimensions, particularly the switch from LED to LCD displays, which eliminated the previous need for frequent battery recharging (the LED display soon became obsolete). What

was the BLS pricing? How could it reach a conclusion so far from facts that are familiar to everyone, including BLS officials?

These examples suggest puzzles that this study identifies rather than resolves. However, even without any specific time series of BLS price quotations from individual reporters, several specific statements about the sources of bias in the PPI can be made.

1. The most obvious flaw in the PPI is that many price changes are compared without required quality adjustments being made at all. Mills (1936) was concerned about this long before the Stigler committee report, and most descriptions of BLS methods indicate that quality adjustments for products other than automobiles were spotty or nonexistent before 1960. For instance, an official with a diesel engine manufacturer (quoted in chap. 9 below) states explicitly that the BLS failed to make adjustments for the fact that his firm's engine was steadily upgraded in horsepower and fuel efficiency in the 1950s and 1960s even though it carried the same model number. Accordingly, it is not surprising that the negative drift of my alternative price indexes relative to PPIs for the same products is most rapid before 1960.

2. A weakness that has been discussed traditionally is that the PPI introduces new models too late and tracks obsolete models for too long. Until 1971, the BLS priced a steam-turbine generator in a size class that had been obsolete since the 1940s. Fluorescent lights were not introduced until 1961. An official with the firm that publishes the used tractor guides claims that the BLS perennially tracks tractor models that are "dogs" having a small market share. Triplett and McDonald (1977) found that "refrigerators in the sample were not a good reflection of the range of industry sales; the sample consisted, disproportionately, of 'bottom-end' or 'builder-model' refrigerators with a minimum of special features."[15] Finally, I have a personal anecdote regarding "obsolete models" with appalling implications if the practice is widespread. In my role as treasurer of the Econometric Society, I have been informed by our former printing firm that in 1988 *Econometrica* was still carried as part of the PPI for the printing industry, even though printing of that journal was shifted from the United States to the United Kingdom in 1984. When informed of this fact in 1984, the BLS commodity specialist told the U.S. printer to submit an imaginary price representing what the printing bill "would have been." We are now negotiating to return *Econometrica* to the same printer, but at a price much lower than the imaginary price, reflecting competitive bids on both sides of the Atlantic.

3. Another traditional criticism is that weights are out of date. In 1983 (the last year of my data coverage), the PPI was based on 1972 weights. As a

15. This interpretation of Triplett and McDonald is quoted from Triplett (1988, 73).

particularly stark example, the weight given in 1983 to electronic receiving tubes was greater than to integrated circuits of all types.[16]

4. For many traditional product lines, where the pace of technical advance is slow, the PPI is entirely adequate, as is confirmed by the low rate of drift in numerous product categories in the 1970s and 1980s. But when technological change is rapid, the PPI is subject to a fatal weakness in its basic methodology of relying on reporters, that is, officials in manufacturing firms who have little concern with the purposes of their PPI price reports, to flag quality improvements and to introduce new models. The path of least resistance is to report the price for the same old model, even if its sales have almost disappeared. The PPI questionnaire, at least in its 1976 incarnation reproduced in the Ruggles report (U.S. Executive Office of the President 1977) fosters this fatal flaw by asking only for the value of quality changes on the current model, not for information on any new models that may have been introduced. An advantage of the catalog indexes is that the range of models covered for each product is *complete*, not a sample, and hence grows and shrinks automatically each year with the market, and no decision is necessary on when to change models (see the discussion of fig. 10.1 below).[17]

5. An obvious flaw in the PPI is that important products are not covered at all. There is no PPI for electronic computers, PCs, electric utility generating equipment, commercial aircraft, or telephone switching equipment, all of which are covered in this book.

6. Even if PPIs did exist for these products and others, the PPI specification method cannot cope with electronic products and those experiencing nonproportional changes in energy efficiency. In practice, if not in principle, the PPI is tied to the old-fashioned methodology that was implicitly rejected by the BEA's incorporation of a hedonic price index for computers into the NIPA. Technologically progressive products, like electronic computers and electronic telephone switching systems, are characterized by new model introductions in which performance is improved relative to price. The BLS questionnaire, at least its 1976 version, does not contain questions that would lead a manufacturer to report such changes in quality that are not accompanied by a change in cost. This problem is aggravated in the PPI by the linking in of new products like electronic calculators without any allowance for the price reduction implicit in the shift of users from the old model (in this example, rotary electric calculators).

How do these problems with the PPI help explain shifts in the alternative/official rates of drift over the three postwar time intervals? No

16. The 1983 relative importance of index 117801, receiving-type electron tubes, was 0.062 percent, and that of all types of integrated circuits was 0.051 percent (this is the sum of indexes 117841, 117842, 117845, and 117846).

17. Auto batteries represent the only product for which the Sears index is based on a sample of available models.

precise connection can be drawn, because some of the shifts over these intervals reflect the changing weight and pace of innovation in product categories for which there is no corresponding PPI, for example, computers and aircraft in the middle (1960–73) interval. Although shifts in the rate of drift reflect many factors, not just flaws in the PPI, I conjecture that the problems before 1960 and after 1973 were of a different nature. Prior to 1960, we know from many descriptions, which are referenced in chapter 3, that there was little systematic attempt to control for quality change in the PPI. The fact that many of the Sears catalog indexes for relatively simple products show significant drift through the early 1960s, but then little thereafter, supports this view. Another problem for the PPI is that it did not adequately capture the shift from premiums to discounts in the early postwar years, as is revealed by the behavior of the hedonic indexes for autos and tractors based on list prices of new models relative to similar indexes calculated from used-market prices of late models.

After 1973, the PPI has been undermined by electronics. Part of this comes out in the fact that a rapidly growing nominal expenditure weight in the Törnqvist indexes for aggregate PDE is allocated to the office machinery and communications categories, where electronic innovations are important yet where PPIs do not exist. Yet even in those few electronic products where there are explicit PPIs, for example, nonprinting calculators, ten-key adding machines, and electronic typewriters, the price increases registered by the PPI conflict starkly with everyday experience. The same phenomenon probably occurs in some other components of the PPI that are not covered in this study, for example, numerically controlled machine tools and electronic typesetting machines used by the printing industry.

Most of the electronic products priced by the PPI are not final goods but rather intermediate components used in the manufacture of other products, for example, electron tubes, capacitors, resistors, connectors, integrated circuits, and diodes. The ability of firms using electronic components to substitute transistors for tubes, and then integrated circuits for transistors, and then successively more powerful integrated circuits for less powerful circuits, is not taken into account at all. PPIs for integrated circuits were simply linked into the overall PPI in December 1974 as a totally new product, without any accounting for the implicit price reduction that their invention implied, just as electronic calculators were linked in as a new product in 1969 without any allowance for the price reduction when users switched from rotary electric calculators to electronic calculators. The late introduction of integrated circuits and the use of obsolete weights helps account for the remarkable fact that the PPI group index (1178) for "electronic components and accessories" displays a 1967–83 price increase of 83.6 percent, while the price of virtually every type of final electronic product covered in this study (computers, PCs, electronic calculators, electronic typewriters, VCRs) declines at a double-digit annual rate during the same interval.

This set of defects in the PPI leads to a simple explanation of the relatively low rate of drift observed in the middle period, 1960–73, for products other than electronic computers and aircraft. In these years, the early postwar readjustment was over, discounting had become established, the PPI had begun to make more systematic attempts to adjust for quality change on conventional products, and electronic products were still relatively unimportant. This is the period during which the PPI and CPI for automobiles are probably most accurate, for instance.

This study contains no explicit recommendations for reforms in organization or procedures at the BLS or BEA. As presently organized, the two agencies almost inevitably demote historical price index research to a low level of priority. The BLS has primary responsibility for price measurement but cannot revise its basic indexes, the CPI and PPI, because they are used to escalate legal wage and price contracts. Thus, the BLS is future oriented, interested in better procedures that may improve price indexes in the 1990s, but not in fixing up errors made in past time intervals.

The BEA is not bound by any limitation on historical revisions and indeed introduces major benchmark revisions every five years. But the BEA has many responsibilities, among which historical price research is low on the list. The development of a computer price index by the BEA and its introduction into the NIPA represents less a precursor of a major future commitment to historical price research than a recognition that the previous PDE deflator was seriously deficient in its assumption that computer prices were fixed. In this special case, the actual research on computer prices was carried out largely at IBM rather than at BEA. These limitations on historical price research at the BLS and BEA make it likely that primary responsibility for historical research on the accuracy of price indexes will continue to rest with the academic sector.

Nevertheless, there are several steps that the BLS and BEA might take in collaboration to use the results in the current book to improve the official indexes. Conversion of the entire set of NIPA deflators to the Törnqvist weighting methodology should occur formally at the next benchmark revision, and in the meantime the BEA should publish an article with alternative indexes of real GNP and its components based on this superior weighting method. At the level of individual prices, in the 1970s and 1980s the PPI is most prone to an upward bias in measuring the prices of electronic goods. This book shows that it is possible to create price indexes for electronic goods from catalog and *CR* data; the BLS could develop experimental indexes from those sources and others, for example, mail-order advertisements for personal computer processors and peripherals, and could obtain required specification data directly from manufacturers. The present procedure of collecting only price data from respondents, and leaving it to them to make corrections for quality change, is so clearly unsatisfactory for electronic goods that it will have to be abandoned. This book identifies a

number of products where the PPI appears to have a consistent upward bias, indicating the need for the study of experimental indexes based on nontraditional sources in order to determine the source of the bias, while for many other products the evidence developed here suggests little or no difference between the alternative indexes and the PPI, particularly in the post-1973 period.

1.6 Qualifications to the Findings

1.6.1 Actual and Potential Criticisms

Several possible weaknesses should be considered in evaluating these results. First, no data have been collected for products representing a substantial fraction of investment in producers' durable equipment; roughly 77 percent of PDE by 1967 value is covered, and 23 percent remains uncovered. The treatment of the uncovered products involves the implicit assumption that, within any one of the twenty-two PDE categories, the observed drift for the covered products can be applied to uncovered products. Here, the crucial word is *within,* for the drift for the covered products is imputed to the uncovered products only inside each of the twenty-two PDE categories, not across them. Thus, the drift for covered metalworking machinery is imputed to uncovered types of metalworking machinery; the drift for office machinery or telephone equipment is not imputed to the uncovered types of metalworking machinery.

To take the opposite approach and do no imputation would amount to assuming that the PPIs are perfectly accurate for the uncovered products. This would doubtless be wrong, for it would overlook the systematic sources of bias in the PPI discussed in the previous sections. We cannot rule out the possibility that the uncovered products have a smaller "true" (and unobserved) drift than the covered products within the same category. But there are plenty of examples of uncovered products that have experienced radical decreases in price, from satellite television dishes to smoke detectors, or substantial increases in quality relative to price, including copying, photographic, and sports equipment, not to mention electronic medical testing equipment from CAT scanners to ultrasound testers.

An explicit criticism directed at the first (1974) version of this research by experts at the BLS was that several of the Sears catalog results involved the comparison of small items sold by Sears for home use with PPIs that priced much larger industrial varieties of the same product; the examples involved eight of the Sears-PPI comparisons, or about 10 percent of the products covered.[18] I have responded to this criticism in two ways. First, the

18. Memo from John Early to Jack Triplett, provided to me as an attachment to a letter from Jack Triplett, 16 December 1977.

catalog-to-PPI product matches were reviewed carefully and tightened. Some of the most convincing evidence of bias in the PPI comes from products where the specification of the PPI commodity index matches precisely what is priced here, for example, home utility quarter-inch electric power drills or stationary air compressors with PPI capacity levels exactly matched by the catalog products. In a statistical test in chapter 10, there is no correlation at all between the tightness of the catalog-PPI match and the drift in the catalog/PPI price ratios for particular products. Second, as discussed in the preface and earlier in this chapter, this criticism led me in the present version to broaden the study to include more "large" items such as aircraft and railroad locomotives.

While the study is no longer vulnerable to this criticism, it seems clear in retrospect that the earlier critique based on the "small-large" distinction rested on a false premise. It may be valid to complain that the use of Sears catalog prices involves a sample selection bias; I deal with this criticism at length in chapter 10 and argue that, if anything, the Sears catalog indexes are biased upward owing to a change in Sears's pricing policy after the mid-1970s. But it is not valid to argue that the comparison of small items with larger items creates a de facto presumption of downward bias in the alternative indexes. Internal evidence denying this claim is readily available in the PPI itself, where price changes can be tracked over long periods of time for detailed commodity indexes covering several sizes of the same product. These changes do *not* reveal any tendency for small items to increase in price less rapidly than large items. For instance, the PPI for home-model electric drills increased substantially more over 1967–83 than that for industrial-model electric drills.[19] For twelve product categories where a comparison could be made between PPIs for small and large sizes of the same machinery product, for example, centrifugal pumps with capacities of 90 and 1,000 gallons per minute, the 1967–83 price increase averaged over the twelve categories was 180.6 percent for the smallest variety listed and 178.5 for the largest variety, that is, about the same. This average disguises great differences in the price increases of small and large items for a given product category, but cases of lower price increases for small items are almost exactly balanced by cases of higher price increases for small items.[20]

Another qualification is the possible inaccuracy involved in the use of mail-order catalogs and used vehicle prices as close proxies for transaction

19. The 1967–83 increase was 128.6 percent for home utility drills (index 11320222) and 76.1 percent for industrial-line drills (index 11320301).

20. The twelve products are all those in PPI groups 11 and 12 (electrical and nonelectrical machinery) where eight-digit indexes are available for at least two size classes of the same product, where the product itself was introduced before 1980, and where the different size indexes, if introduced after the 1967 base year, were introduced at the same time. Of the twelve product groups, all but two were introduced in 1967 or before, one in 1972, and one in 1976.

prices. Improved or decreased relative operating efficiency or changed pricing strategies might have affected the price differential between catalog houses and walk-in retail stores. The relation between transaction prices of new cars and the prices of late-model used cars might have been affected by changes in tastes. But in the case of automobiles the new estimates of transaction prices correspond closely to scattered pieces of independent evidence on the magnitude of price premiums and discounts, as cited below in chapter 8. All the available evidence, both on the changes in Sears's pricing policy and on changes in the average level of Sears's prices compared to competitors, shows consistently that Sears's prices shifted from the low end of the spectrum to the middle over the postwar period, indicating that the dominant bias in our catalog indexes is likely to be upward, that is, that the negative drift of the alternative price indexes based on catalog data relative to the corresponding official indexes is understated. This verdict is supported by the *CR* indexes developed for appliances, which are based on market prices after 1959, which cover all major manufacturers, and which in almost every case exhibit an even greater drift relative to the PPI than the catalog indexes. The upward bias in the Sears catalog indexes relative to the *CR* indexes stems not from different methods of quality adjustment but from more rapid increases in Sears prices than prices of other manufacturers. Chapter 7 computes the ratio of the price of Sears models to the average of all models as listed in *CR* articles on refrigerators, air conditioners, and washing machines. This ratio increases from 80 percent in the earliest *CR* articles (published in 1949–59) to 102 percent in the most recent articles (published in 1974–84), indicating a systematic upward bias in price indexes based on Sears data and suggesting that the many results in this book based on Sears data may on average understate the upward bias in the PPI.[21]

In contrast to the many cross-checks of the results provided in this book, an outsider does not know what to make of the individual PPI commodity indexes when outside evidence reveals a major conflict. As in my example of the diesel engine index cited above, it is impossible for an outsider to obtain answers to the following basic questions. Which manufacturers submitted the BLS price quotations? Which models were included in each commodity index, and how often did they change? Precisely how were quality adjustments performed in each individual year?

In contrast to the evidence developed in this book on the basis of almost 100 separate product indexes that the PPI for machinery contains a

21. A needed qualification regarding refrigerators is that the achievement of greater energy efficiency, which is taken into account in the final indexes in this book, has come at the cost of the use of chlorofluorocarbons, or CFCs, which are members of a family of chemical that have been identified as major contributors to ozone depletion. A typical refrigerator now has 2.5 pounds of CFC 11 in its insulation and about eight ounces of CFC 12 in its cooling system. Because of concern about the ozone layer, use of these materials may be banned in the 1990s, leading to less efficient refrigerators. See Holusha (1989).

substantial upward bias, Triplett's recent (1988) review of the price measurement literature suggests only a single example of a downward bias in an official price index for a durable good.[22] This refers to the treatment in the auto component of both the CPI and the PPI of pollution and safety equipment as representing increases in quality rather than price. Triplett would prefer a treatment in which these items are treated as a "transportation tax" on consumers. In his preferred treatment, no quality adjustment would be made for these items, so that their introduction would be allowed to raise prices rather than quality. My analysis of automobiles accepts the present BLS approach, which treats pollution and safety devices as increasing quality rather than price. The only other place in this book where this issue arises is in the development of a price index for electric utility generating equipment, where I suggest a rough downward adjustment to the new price index in order to make the treatment of government-mandated scrubbers and other nonproductive components of generating equipment parallel to that of automobiles in the CPI and PPI.

Triplett's suggestion of downward bias in the official indexes for this reason involves a conceptual rather than a measurement issue. As discussed in chapter 2, legislated environmental and safety devices are the major reason why output and input price indexes need to be distinguished, since such devices drive a wedge between price indexes relevant for deflating the output of the industries producing capital goods and those relevant for deflating the capital input of the industries using those goods. Since productive inputs are used by capital goods manufacturing firms in producing environmental and safety devices, a study of output and productivity behavior in the producing industries should use price indexes that treat such devices as an increase in quality, whereas a study of productivity behavior in the capital-goods-using industries should employ price indexes that treat such devices as an increase in price, that is, as nonproductive. Whether such devices should also be treated as a price increase in the CPI and in the deflation of consumer expenditure depends on whether they yield any benefits in the form of reduced pollution and increased safety, and whether the CPI accurately measures the value of such benefits.

Because I accept the current official treatment of these devices, the new PDE deflators are suitable for the deflation of the output of industries producing producers' equipment, but are downward biased for the purpose

22. While Triplett had no access to the 1988 version of this book when writing his recent review (1988), he not only had access to the 1974 draft but also contributed copious comments on that draft, which have been incorporated into this version. In view of that access, it is astonishing to read the statement in his recent review (1988, 71) that "the PPI has a very large number of components and few of them have been examined in research," with no reference at all to the 1974 manuscript, which contained detailed annual comparisons over the 1947–70 period of alternative and PPI commodity indexes for more than 100 products.

of constructing indexes of capital input in the using industries and hence result in an overstatement of the growth of the productive capital stock in those industries. Since neither the CPI nor the consumer expenditures deflator in the NIPA contain any explicit adjustment for the benefits of mandated pollution and safety devices, the common treatment in the CPI and in this book of such devices as an increase in quality rather than price represents an arbitrary and unsupported judgment that the costs of such devices to consumers are exactly balanced by their benefits. If their costs exceed their benefits, then the alternative deflator for consumer durable goods contains a downward bias to this extent, and vice versa.

An implication of this qualification is that increased expenditures on mandated pollution and safety equipment cause the "true" growth of capital input to slow down more after 1973 than the growth rate of capital input displayed in table 1.3. While I have not developed an explicit measure of this source of bias, it does work in the direction of helping to explain part of the post-1973 productivity slowdown. For instance, if the growth rate of capital input were slowed by the adjustment for mandated equipment by 0.40 points, from 3.56 to 3.16 percent per year, the growth of multifactor productivity would be raised by one-quarter of this amount, from 0.09 to 0.19 percent per year. This would reduce the slowdown in multifactor productivity growth from 1.51 to 1.41 percent per year.

1.6.2 The Use of "Cross-Checks"

While the new results may appear startling in the magnitude of the drift that they reveal between the alternative and the official price indexes for particular products, the study has emphasized the collection of numerous "cross-checks" to make the recorded price changes "come to life" rather than lying inert on a computer printout. One source of cross-checks is contained in the popular literature on the history of particular products. These provide numerous examples of high prices in earlier years that might be compared to today's typical prices to dramatize the widespread occurrence of price declines for durable goods: $475 in 1947 for a 6,000 BTU Fedders room air conditioner that used much more energy than today's less expensive models; $425 in 1946 for a nine-inch black-and-white table-model television; $5,200 in 1948 for the first Ampex audio tape recorder; $1,000 in 1954 for a nineteen-inch console color television; $75,000 in 1956 for the first video tape recorder; $209 in 1958 for the first Smith-Corona portable electric typewriter with narrow carriage and manual return; $2,295 in 1975 for the first Betamax home videotape recorder without remote control or programming capability; and $1,000 in 1983 for a compact-disk player.[23]

23. All these examples are from Consumers' Union (1986), except for the 1958 Smith-Corona typewriter, which is the price quoted in the spring 1958 Sears catalog.

At a more formal level, developed as an integral part of the study of the mail-order catalog prices are many comparisons of "closely similar" models that provide specific examples of catalog models exhibiting price changes far below those implied by the PPI over periods of two or three decades. These comparisons undermine the credibility of many of the individual PPI commodity price indexes. In a case where the PPI doubles relative to the catalog index over two or three decades, it seems hard to believe that the PPI respondents would have any customers left except for the likelihood that the PPI price quotations are not corrected for quality change with the same exactness and consistency as are the Sears catalog price quotations. In fact, the long-term comparisons of similar models generally imply smaller price increases, and a greater drift away from the PPI, than is implied by the catalog price indexes themselves. This finding supports the view, so evident for computers, that new model introductions are often accompanied by an implicit price reduction that goes unmeasured in the conventional specification method.

Other sources of cross-checks should be stressed. The study includes both catalog and *CR* indexes for fourteen products. It includes the evidence of the market for used aircraft supporting the theory and practice of the new treatment of aircraft prices. It includes the thirty-year history of specific IBM mainframe computer models that supports the pace and timing of the new mainframe computer index. It includes written communications from the leading manufacturer of medium-sized diesel engines, itself a PPI respondent, explicitly supporting the conclusions of the case study of diesel engines, which yields one of the largest rates of drift from the PPI, and denying the validity of the PPI for diesel engines. Finally, it includes the correlations between technological complexity and the alternative/official drift that provide a rationale for the findings.

1.6.3 A Last Word on Unmeasured Quality Change

While there are substantial elements of plausibility and internal consistency in the results, there are two important reasons to suggest that the final results may actually understate the drift of the alternative indexes from the official indexes. First, as shown in table 12.1 below, 38 percent of the 1967 weight in the alternative PDE deflator is based on the Sears catalog indexes for products other than appliances. The comparisons of Sears and *CR* indexes for appliances, outboard motors, typewriters, and power hand tools reveal an upward bias in the Sears index in almost every case, on the presumption that the *CR* index is more accurate because of its greater coverage across manufacturers. Reflecting this presumption, the new price deflators for PDE and consumer durables are based on *CR* rather than Sears data for appliances. But this leaves open the possibility that the Sears indexes for products other than appliances, which are the basis for much of the final PDE index, on average contain an upward bias.

Probably more important as a source of upward bias in the final price indexes is a substantial margin of unmeasured quality change for several products. There is no better way to conclude this summary chapter than to list those elements of unmeasured quality change. In this book, no account is taken of the following dimensions of quality change, all of which are documented either in subsequent chapters, in sources cited in those chapters, or in footnotes here:

- Reduced vibration and noise of jet planes (recall that flight attendants were originally nurses on board to deal with air sickness), their ability to fly above thunderstorms, and their improved safety record relative to piston planes;
- The unmeasured value of time savings to travelers created by the transition to jet planes;
- Increased daily utilization rates of jet aircraft as compared to the piston aircraft that they replaced;
- Greatly increased realized service lifetimes of jet aircraft relative to the piston aircraft that they replaced;[24]
- A tripling of service lifetimes of automotive diesel engines between 1945 and 1975;
- Technical improvements in locomotives that reduce "wheel creep" and substantially improve hauling ability relative to horsepower, and improved fuel efficiency of locomotives made possible by lighter rail cars;
- Additional features on electronic typewriters and calculators that did not exist a decade ago;
- Reduced energy, space, and air conditioning requirements of electronic computers, photocopying machines, and other electronic products;
- Improved handling and riding capabilities of automobiles;
- Improved sound and reception quality and greater reliability of radio and audio/stereo equipment;
- Effective price reductions implied by the availability of personal computers to perform tasks formerly done on mini- and mainframe computers;

24. More than 100 of the original unstretched 727 aircraft are still in daily service with U.S. domestic trunk airlines, all having an age of twenty-four to twenty-six years in 1989. Their service lifetimes are likely to be extended well past thirty years, as a substantial number have been sold by United Airlines to Federal Express, for delivery during 1988–91. The Chairman of United Airlines recently declared that his airline's 29 DC–8–71 aircraft, built in 1968–69 and reengined in 1982–83, would last "indefinitely" and were being sold only because the price offered by the purchaser was irresistible (United Airlines quarterly financial report, second quarter, 1988). In 1988–89, there were several incidents of physical deterioration of older planes, most notably the 1988 Aloha Airlines incident, when corrosion caused part of the metal skin to fall off a twenty-year-old 737, resulting in the death of one flight attendant, and the 1989 United Airlines episode, in which a cargo door opened in flight on an eighteen-year-old 747, tearing off part of the metal skin and sucking nine passengers to their deaths. As this book went to press, however, these were isolated incidents, and there was no movement by the government or the airlines to ground old jet planes. In fact, the United plane was promptly repaired and put back into service in the fall of 1989.

- Greater ability of refrigerators to maintain a fixed low freezer temperature;
- A shift from metal to plastic parts on many products, which reduces weight and increases service lifetime by reducing corrosion (e.g., by eliminating the need to discard a dishwasher when the tub corrodes);[25]
- Double insulation and other safety devices on home power tools;
- Reduced weight of home power tools;
- Improved design of power lawn mowers, which has resulted in an order-of-magnitude reduction in injuries since the mid-1970s;
- Reduction in the variance of quality differences across brands, as implied by the great reduction in the frequency of "not acceptable" ratings by *CR;*
- Reduction of the frequency of service calls for refrigerators and washing machines (explicitly measured and taken into account only in the new index for television sets);
- Reduction of noise, weight, bulk, and installation cost of room air conditioners;
- Improved cleaning ability of automatic washing machines and dishwashers;
- Reduced size and improved safety of microwave ovens;
- Improved fuel economy of tractors, outboard motors, and diesel engines, and reduced electricity comsumption of condensers and compressors;
- And, finally, immeasurably better picture quality of color television sets.

25. Chapter 10 cites recent reports in *CR* on power tools that downgrade models containing metal rather than plastic cases.