

## Problems of accountancy and interpretation in assessing long-term economic performance

Kennedy, William P.

Veröffentlichungsversion / Published Version

Sammelwerksbeitrag / collection article

### Empfohlene Zitierung / Suggested Citation:

Kennedy, W. P. (1983). Problems of accountancy and interpretation in assessing long-term economic performance. In R. Fremdling, & P. K. O'Brien (Eds.), *Productivity in the economies of Europe* (pp. 57-78). Stuttgart: Klett-Cotta. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-329339>

### Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

### Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

William P. Kennedy

## Problems of Accountancy and Interpretation in Assessing Long-Term Economic Performance

“In intensity of feeling ... and not in statistics, lies the power to move the world. But by statistics must this power be guided if it would move the world aright.”

Charles Booth, 1891.

In the popular mind, economic growth has generally been regarded as a “good thing.” This attitude no doubt springs from the widely-shared feeling that if only one’s (real) income were higher—that is, if economic growth in the past had been more rapid—one would be better off in some concrete sense. It would then be possible to afford better housing, better transport, more varied and interesting leisure, and so on. A little daydreaming will generally be sufficient to lengthen this list to almost any desired limit. Although such attitudes not only suggest a natural and feasible means of measuring economic activity but also accurately reflect important aspects of that activity, historical analysis, in common with the demands of public administration, requires a more searching evaluation of economic effort. For the historian, this requirement stems perhaps most fundamentally from the need to measure change over time combined with the realization that societies are complex entities made up of many individuals whose preferences differ and whose welfares are therefore differentially affected by any given change. Thus important conceptual problems arise both in aggregating for one individual the value of the different goods consumed (because of the necessity of assuming identical marginal utility of expenditures on all goods) and in aggregating across many individuals the value of the consumption of the same good (because personal circumstances differ). These problems are then enormously compounded by the aggregation of all goods across all consumers. The use of market prices to accomplish these essential tasks of aggregation, without which no analysis is possible, often imposes such unacceptable assumptions upon the historian as the belief that all observed prices are full-information equilibrium prices, that the distribution of wealth is “optimal”, or that all individuals’ preferences are identical.<sup>1</sup> But if observed market prices are rejected, how else is measurement to be made? Moreover, factors such as productivity, widely recognised to be

---

1. Sen, Amartya K., *The Welfare Basis of Real Income Comparisons: A Survey*, in: *Journal of Economic Literature*, 17 (1979), pp. 3-35.

of signal importance, depend for measurement critically upon the precise designation and valuation of economic input and outputs. Here again, any inadequacies of observed prices have very serious consequences for measurement and analysis as well as for resource allocation.

Because historians have always been sensitive to the intricacy of the past and because historians' purposes are those of broad assessment and evaluation, clumsy and insensitive systems of measuring economic change must invariably be disturbingly inadequate. Yet the construction of satisfactory measures has proved to be almost paralyzingly difficult. The result has been an uneasy, resented compromise between what can be done and what should be done. While this situation will not be altered easily or quickly, the first step in amelioration must be for all historians, as both consumers and producers of statistics, to be more consciously aware of both the difficulties that have been encountered in the past and the possible strategies that may be open for progress in the future. To that end this essay has three objectives: to review the most important inherent difficulties in measuring economic activity; to consider some of the most promising remedies offered for those difficulties; and, where possible, to consider for the remaining problems some strategies that may prove fruitful in the future.

## I

The problems of measuring economic activity may usefully be divided into two categories differentiated by their amenability to the use of observed price data.<sup>2</sup> The first and more tractable category contains those issues where prices that in principle can be observed can be used. The second, more intractable and perhaps—at least for historians—more important, category contains those issues that must be resolved by the use of estimated prices because observed prices are for a number of reasons defective. Such defects may arise because observed prices are in fundamental disequilibrium and therefore, given the preferences, technology, and resources of a particular economy, seriously deceptive; because observed prices change over time due to both general inflation (or deflation) and relative price movements; because, observed prices are deemed to be unacceptably determined by a particular, possibly arbitrary, distribution of wealth; or because observed prices are considered distorted by tariffs, quotas, or various other types of administrative action. The conventional National Income and Product Accounts are constructed mainly using observed prices, although imputations of unobserved but market related prices are made in certain crucial instances—most notably the value of the services accruing to the owner-occupiers of houses who pay no explicit rent for their accommodation and the value of agricultural output consumed on farms—and most National Income and Product Ac-

---

2. The stress placed in this essay on the use of prices should not be construed as limiting the analysis only to market economies. Centrally planned economies implicitly or explicitly, consciously or unconsciously, must produce prices in order to co-ordinate activity and those prices (or shadow prices) will reflect the strengths and weaknesses of the planning mechanism in the same way that market-based prices reflect the strengths and weaknesses of market mechanisms.

counts (NIPA) provide estimates of income and output over a span of years in constant prices even though the prices used could have at best been observed at only one instant.

At the present, however, even disregarding problems of the second category, the conventional NIPA are inadequate for two main reasons. First, and less importantly, the accounts, even in their own limited terms of reference, are not yet completely consistent logically. Broadly, the NIPA are designed to measure a nation's economic activity in three theoretically independent methods. One method measures the incomes of all factors of production (including those located abroad but owned or residing domestically and excluding those located domestically but owned or residing abroad) before transfers for taxes, gifts, or other unilateral payments not matched by a countervailing flow of goods and services. A second method measures net value added in economic activity by each of a nation's productive enterprises. The third method measures the flow of national expenditures on final goods and services, including exports but excluding imports. The accounts represent a series of compromises between what may be measured easily and accurately and what is logically required. Although the resolution of such problems is comparatively easy, extensive reworking of the conventional accounts is often necessary and the difficulties imposed by the inevitable sparseness of historical data can transform tedious but conceptually routine calculations into substantial and demanding tasks of indirect estimation.

An example of the sort of anomaly of this variety likely to cause historians difficulty can be drawn from the convention employed in the British NIPA for measuring the value added to national income by financial intermediaries. Financial intermediaries have two main sources of income: (1) service charges and commissions, which together typically account for only a small proportion of their operating income, and (2) the net revenue resulting from charging to final borrowers a higher rate of interest than they as a group pay to depositors. However, the British accounts, in common with many others, treat interest payments and receipts as transfers and not as payments and receipts for final services.<sup>3</sup> Thus the net earnings of financial intermediaries are generally understated, the net earnings of borrowing companies are correspondingly overstated and, to the extent that private individuals rather than companies are final borrowers, consumers' expenditure on final goods and services is understated. Consistency would require (1) that non-financial company profits be reduced by the amount of net interest paid to intermediaries and intermediaries' net income be similarly increased and (2) net interest payments by the personal sector to intermediaries be treated as payment for a final (consumption) service rather than as a transfer. In the belief that sufficient accessible evidence has not been available to permit such reallocations among companies and between companies and individuals, the Central Statistical Office has decided that indirect imputation of intermediaries' value added would be more misleading than the obvious paradox of flourishing intermediaries appearing to make steady annual losses, a paradox that has been preserved in Charles Feinstein's indispensable volume, *National Income, Expenditure, and Output of the United Kingdom, 1855-1965*, in order to maintain consistency be-

---

3. Maurice, Rita, (editor), *Central Statistical Office National Accounts Statistics: Sources and Methods*, London 1968, pp. 204-205.

tween current and historical data series.<sup>4</sup> However, it is made abundantly clear on all sides that the paradox in the British NIPA of intermediaries steadily making large losses exists only for expediency and that logically acceptable, albeit labour intensive, alternative accounting conventions exist, as may be seen in proposals made by the U. N. in *A System of National Accounts and Supporting Tables* (New York: United Nations, 1964) and by the OEEC, *A Standardized System of National Accounts* (OEEC: Paris, 1959). It is important to stress, however, that the fundamental problem is not one of non-existent data or of unavoidable logical inconsistency but rather a lack of commitment and interest by both the national income accounts authorities and historians to utilize what is available. Indeed one might argue that the quickest way to secure progress in this area would be for an historian to make a provisional allocation of interest payments as dictated by logical consistency, thereby exploiting and drawing attention to the available data and setting upper and lower bounds to the data series most affected. Although it is only to be expected that the first estimates will be thoroughly revised, a successful pioneering effort, by clearly defining the problem and identifying the necessary data will greatly aid subsequent work.

Another anomaly of a related type likely to create difficulties for historians, simply because it creates grave problems for conventional accountants, arises in the treatment of profits, which conceptually are calculated net of depreciation to distinguish them clearly from cash flow but are often reported gross of depreciation because depreciation is so difficult to estimate, particularly when the quality and the relative price of capital goods changes as a result of technological progress. At the very least, the timing of profit peaks and troughs will be affected by the procedures for measuring depreciation, but it can easily be seen that other important issues related, for example, to the distribution of income and the size, composition, and productivity of the capital stock are also involved. The treatment of depreciation also affects the measured levels of consumption, most notably the services derived from consumer durables such as automobiles and household appliances. Logically, these items should be treated in the same way as houses, being noted as additions to the capital stock—that is, as investment—upon completion and thereafter yielding a flow of consumption services gradually diminishing as depreciation occurs. Consumer durables other than houses, however, are not treated as additions to the capital stock that subsequently yield flows of services but as consumption items that are counted as if they were consumed immediately upon acquisition.<sup>5</sup> Such treatment, if accepted without reflection, leads to nonsensical results for it implies that for identically priced goods the rate of depreciation does not matter whereas in reality it matters a great deal. Cars that last for four years without major repairs yield a greater flow of services than do cars that last only for two, although this fact is ignored in the conventional accounts except to the extent that the price of the more durable car is greater than the price of the less durable one. In this case, as in the case of the value added by financial intermediaries, the remedy—allocating the value of the durable good over the time period in which it depreciates while yielding service—is straight-

---

4. Feinstein, Charles H., *National Income, Expenditure and Output of the United Kingdom, 1855-1965*, Cambridge 1972, pp. 141-43.

5. Maurice, *Sources and Methods*, p. 365.

forward, albeit tedious in terms of calculation and demanding in terms of data requirements.<sup>6</sup>

The more important inadequacy of the conventional accounts, however, arises not from logical inconsistency but from an inappropriate but understandable choice of objective. The conventional accounts are designed to measure marketed production whereas historians, in common with most other users of the accounts, are ultimately interested in sustainable consumption, the obvious objective of all economic activity.<sup>7</sup> To be sure, marketed production is an important component, perhaps even the most important single component, of consumption and in addition substantial benefits are derived from a choice of objective which lends itself—as marketed production does—to relatively straightforward extrinsic measurement. Nevertheless, marketed production is not in itself an objective of economic effort and is therefore inherently a poor indicator of it. Furthermore, in precisely those periods when economic activity is undergoing important structural changes—periods such as the classical Industrial Revolution or the emergence of post-industrial society—the relationship between marketed production and non-marketed economic activity is most likely to be changing as well. At such times, the conventional accounts will not be merely an indirect and imprecise means of monitoring economic activity but will also be systematically misleading. As long as the relationship between measured and unmeasured activity is constant, the conventional accounts will at least reflect reasonably faithfully changes in overall economic activity, but if the relationship itself is changing, it is no longer possible, without additional information, to infer the characteristics of aggregate economic change.

The nature of the problems that arise from the consideration of marketed production rather than sustainable consumption might best be conveyed by illustration. It has been stressed, for example, that virtually all societies at all times have provided themselves by one means or another with textiles, tools and other simple manufactures. For much of human history these simple goods have been produced directly by those, or the near kin of those, who were ultimately to consume them and this production was often totally removed from any sort of market transaction.<sup>8</sup> Gradually,

---

6. It perhaps should be noted that William D. Nordhaus and James Tobin, in a preliminary calculation, show that the depreciation of consumer durables other than houses is more significant for its logical implications than for its practical consequences. See Nordhaus, William D., and Tobin, James, *Is Growth Obsolete?* in: National Bureau of Economic Research, *Economic Research: Retrospect and Prospect—Economic Growth, Fiftieth Anniversary Colloquium V*, New York 1972. There is another complication, however, which Tobin and Nordhaus do not consider. This concerns the incidence of capital gains and losses that occurs because of unanticipated changes in the relative prices of capital goods. Such relative price changes cause the anticipated time profile of depreciation to differ from the actual pattern. If relative capital goods prices rise, the firm simultaneously realizes a capital gain (because it bought the equipment relatively cheaply) but also must adjust upwards its depreciation allowances (because the equipment is more expensive to replace). The reverse occurs when relative capital goods prices fall: the firm suffers a capital loss and must adjust downwards the appropriate depreciation allowance.

7. It is necessary to stress “sustainable” consumption in order to exclude from consideration consumption that is made possible by a temporary deterioration in the capital stock.

8. Hymer, Stephen, and Resnick, Stephen, *A Model of an Agrarian Economy with Non-Agricultural Activities*, in: *American Economic Review*, 59 (1969), pp. 493–506.

however, this domestic production became more specialized in certain regions and surplus producers began to sell an increasing fraction of their output in formal markets, often organized by a merchant entrepreneur who would “put out” raw materials to cottage workers and market the finished goods produced by those workers. Eventually, when the technology of manufacture was sufficiently sophisticated, specialization would advance further and workers would no longer toil in their own cottages but were forced or lured into factories.<sup>9</sup> Thus over time an increasing proportion of economic activity came to be mediated by markets. If this process were to be measured by the conventional NIPA, the rate of increase of output, which was often quite rapid in any case due to undeniable technical progress, would be greatly overstated, for self-sufficient production would not be counted. Thus the changes recorded in the NIPA would include the effects *both* of more productive techniques *and* of changes in the proportion of total output marketed. The timing and intensity of an “industrial revolution” can, by the mechanical application of conventional methods, be more apparent than real, with output levels much higher at an earlier date than indicated by the conventional accounts and growing much more slowly.

A problem similar in nature is currently affecting contemporary national income accounting as female labour force participation rates rise.<sup>10</sup> Imagine two neighbouring households, in each of which a housewife does her own cooking and cleaning. The NIPA would record only the intermediate inputs (e.g. unprocessed foods and soap) that each bought, the value of the transformed inputs—the hot meals on the table and the clean sheets on the bed—being ignored. Now suppose that each housewife decides to specialize, one doing only cooking and the other only cleaning, and that each sells her surplus in a formal market and buys in the same market the good she no longer produces. In this case, the market value of the labour of each woman is duly recorded in the NIPA, along with, as before, any intermediate purchases. The NIPA will record a great increase in activity but clearly a large fraction of this recorded increase is nominal and not real because home labour was previously ignored by convention. To be sure, this change may also be accompanied by a legitimate output increase as a result of economies of scale and technological progress, but such an increase is likely to be small compared to the recorded change. This deliberately simple example captures the essence of important structural changes currently taking place in the labour forces of modern industrial economies. As an increased proportion of women leave their children in day-care centres, buy prepared foods in supermarkets and fast-food carry-outs, send their cleaning and mending to specialist firms, recorded national income rises, but the increase is obviously greater than the increase in real output simply because such a large proportion of home-centered production had previously been ignored. This problem is the exact reverse of the one encountered in the nineteenth century British NIPA which show the rate of growth of household production falling with the decline in the rate of increase of domestic servants.<sup>11</sup>

---

9. This process has recently been reviewed by Pollard, Sidney, *Peaceful Conquest*, London 1981, pp. 63–78.

10. The extent of this problem has been explicitly recognised by the British Central Statistical Office. See Maurice, *Sources and Methods*, p. 8.

11. Ebury, Marke, and Preston, Brian, *Domestic Service in Late Victorian and Edwardian England, 1871–1914: Reading Geographical Papers No. 42*, Whiteknights/Reading/England

It is safe to assume that consumption of household production did not fall as the recorded number of domestic servants fell but rather that it was the market mediation involved in household production that changed.

The corrections of the obvious anomalies introduced into historical analysis by changes in the extent of market mediation are not conceptually difficult to comprehend but are likely to prove difficult to implement because the amount of indirect estimation required is likely to be large and because the resulting estimates are unlikely to be highly robust with respect to various assumptions. Clearly what is needed is an estimate in the one example of the manufacturing output of self sufficient farms and in the other an estimate of the value of various household tasks. What is being sought is a measure of final output for consumption that is independent of the degree of market mediation. If such a measure can be found, only real output changes that are independent of marketing changes will be recorded. Although this requirement is a daunting one, before despairing, it should be recalled that it is with precisely such matters and details that economic and social historians have recently concerned themselves. While the necessary data may at present be highly fragmentary and incomplete, there is every prospect that it will become more complete in the future, especially as research attention is directed to issues where competing historical interpretations are particularly sensitive to the choice of analytical assumption. Furthermore, recent advances in simulation modeling offer means of utilizing fragmentary information much more efficiently than has been possible previously.<sup>13</sup> Simulation modeling involves describing a fragmentary data series by the known distribution (such as the normal, exponential, or gamma distribution) which on both empirical and theoretical grounds is most consistent with the currently available evidence and then using combinations of such distributions to yield a distribution of operating results for the process being studied. For example, Jeremy Atack used the procedure to assess the relative capital and operating costs of steam and water power in the early nineteenth century American economy but it can easily be seen that the same techniques can be used to estimate the average costs of providing, for example, various types of household consumption, using manuals of domestic management and isolated wage data for servants where Atack used contemporary engine price lists and engineering estimates.

Much more serious problems, both conceptually and quantitatively, arise in the treatment of leisure, obviously a highly esteemed element of consumption. The nature of this problem, which is clearly related to the general problem of assessing non-marketed output, is easily seen. Imagine two economies, identical in all respects ex-

---

1976, Table 5a, p. 23 and Lewis, W. Arthur, *Growth and Fluctuations, 1870-1913*, London, 1978, Table A3.

12. On pre-industrial manufacturing, see for example, Mendels, Franklin, F., *Proto-industrialization: The First Phase of the Industrialization Process*, in: *Journal of Economic History*, 32 (1972), pp. 241-261. On household labour, see Goldin, Claudia, *Household and Market Production of Families in a Late Nineteenth Century American City*, in: *Explorations in Economic History*, 16 (1979), pp. 111-131; Ebury and Preston, *Domestic Service*, pp. 85-104 and Horn, P., *The Rise and Fall of the Victorian Servant*, London, 1975.
13. See, for example, Atack, Jeremy, *Fact in Fiction? The Relative Costs of Steam and Water Power: a Simulation Approach* in: *Explorations in Economic History*, 16 (1979), pp. 409-437.



cept the way in which the benefits of technological change in a particular year are consumed. Suppose that in the first economy the hours of work and intensity of effort remain unchanged and that all of the benefits of the technological advance accrue in the form of more goods and services. By contrast, suppose that in the second economy the initial output of goods and services is maintained and that all of the benefits of the technological advance accrue in the form of fewer hours worked in order to obtain an unchanged level of material output. While by assumption, the two economies differ only in the composition of consumption, they appear markedly different in the conventional NIPA. In the first case, measured marketed income rises by the maximum amount permitted by the hypothesized technological advance. In the second case, measured marketed income does not rise at all although by assumption the increase in real productive capacity in the two economies was identical. It is obviously a critical weakness of the NIPA that the measure of economic activity should be so sensitive to its structure and composition. Moreover, in their pioneer reworking of the conventional NIPA, William Nordhaus and James Tobin found that the assumptions they employed to evaluate changes in available leisure dominated their measure of sustainable economic welfare.<sup>14</sup>

This result stemmed from the authors' inability to determine whether leisure time was itself a final consumption good or whether leisure time was only one of several inputs into a consumption process. If leisure time itself were the final consumption good, the necessary adjustments to the NIPA are straightforward. The change in the number of leisure hours, measured most plausibly as reductions in standard working hours but strictly excluding involuntary unemployment, is estimated for the economy as a whole and weighted by the average hourly earnings of those workers who obtain such reductions.<sup>15</sup> Note that this procedure assumes that workers are indifferent at the margin between earning another hour's income with which to consume more material goods or forgoing the material goods in favour of leisure. In this case, where leisure time itself is the final consumption good, comparisons across time are quite easy. An hour of leisure in 1880 is worth in ultimate consumption exactly as much as an hour in 1913 or in 1980 (assuming a constant marginal utility of leisure).

If on the other hand, however, the historian wishes to argue that leisure time is only one of a number of inputs into a consumption process, changes in the economy over time, most notably technological and demographic changes, make intertemporal comparisons for a particular economy or contemporaneous comparisons among economies with different technological capabilities much more difficult. Consider, for example, the impact of cheap rail transport in the nineteenth century on the leisure activities of the British working class.<sup>16</sup> The rapid growth of seaside resorts and other amusement centres following the advent of cheap rail travel is strong testimony of the contribution this form of technological change had on the enjoyment of leisure

---

14. Nordhaus and Tobin, *Is Growth Obsolete?*, pp. 38–48. Involuntary unemployment is strictly excluded from measures of leisure time. (pp. 44–45).

15. Standard working hours may change through variations in the working hours per day, in the working days per week, in the working weeks per year, or in the working years per lifetime.

16. Hawke, G. R., *Railways and Economic Growth in England and Wales, 1840–1870*, Oxford 1970, pp. 37–40, 52–54.

time. Similar arguments can be made for the impact of cheap books and magazines, bicycles, automobiles, cinemas, television, sports facilities and equipment and so forth. To the extent that technological progress enlarges and enriches the consumption of leisure time, a comprehensive set of NIPA must value leisure more highly over time as technological progress occurs. Nordhaus and Tobin propose doing this by deflating the nominal value of leisure hours over time (and by implication, between countries) by the price index of consumer goods, an index which over long periods of time has risen less rapidly than an index of nominal wages or earnings. The conceptual problems inherent in choice of index cannot readily be evaded because the decision taken makes a crucial difference in outcome. Nordhaus and Tobin's estimate of the per capita increase in measured economic welfare in the U.S. between 1929 and 1962 is 18.6% if leisure itself is a final good but 126.4% if leisure is considered a process fully participating in the benefits of technological change.<sup>17</sup>

Nordhaus and Tobin do not venture a resolution of the uncertainty created by the need to devise an appropriate measure over time of the value of leisure. Their purpose rather was to illustrate a means by which a complex, vital problem could be approached, in the belief that sustained investigation would ultimately yield greater understanding. In pursuing this problem further, historians may very reasonably employ a much more detailed index of leisure activities than the illustrative one used by Nordhaus and Tobin. Each component of a more detailed index would have its own separate price deflator and the weights attached to each component would be chosen to reflect the relative significance, as contemporaries are believed to have seen it, of each component. Here again is an opportunity to use systematically and quantitatively the results of recent research in social history. Such work has added greatly to the knowledge of how the past was actually lived by most people and the revision of the conventional NIPA offers an opportunity to use this new knowledge extensively.

Unfortunately, beneficial technological change is not the only influence on leisure enjoyment which must be assessed. Non-market costs, particularly those associated with congestion and overcrowding as more people tried to take advantage of new leisure facilities, must also be considered. There are a number of ways this task could be approached. For example, cross-section studies could be used which would relate, say, rent charges in various resort areas to the density of vacationers. The steepness of the slope of such a relationship would permit appropriate adjustments for the effect of congestion. The purpose of such adjustments would be more in the nature of ascertaining the relative magnitudes of the considerations involved than of generating precise estimates of what must ultimately be arbitrary magnitudes.

It is of great importance that historians do not expect resolution of the problems raised by the consumption of goods and services whose value cannot be directly calculated—goods such as proto-industrial manufactures and ill-defined but highly desirable services such as leisure—to be achieved quickly. Rather resolution will occur through the slow, controversy-prone process of creating a consensus among historians regarding the significance of the various assumptions made to produce quantitative estimates. Progress will not occur because a correct answer can be found—for there is unlikely to be a unique correct answer—but because the *process* of investiga-

---

17. Nordhaus and Tobin, *Is Growth Obsolete?*, Table A16, line 16, pp. 52–53.

tion will methodically expose important issues and problems and allow the quantitative significance of different assumptions to be carefully recorded.

## II

Concentration on marketed production rather than sustainable consumption encourages the blurring of the critical distinction between intermediate and final goods. It is well-known that failure to preserve this distinction results in serious double counting. The most obvious example of an intermediate service routinely recorded in the British NIPA as a final output is the expense incurred by workers commuting to work.<sup>18</sup> As in many other similar situations, this convention has been adopted due to a desire to obtain a precise measurement rather than engage in what amounts to speculation, even if this requires an inappropriate definition of what is to be measured. The problem is that it is difficult to distinguish travel for pleasure, indisputably part of final consumption, from travel required for work. Thus expenditures on final goods and services are overstated by the amount spent on commuting, which is not a consumption item but an intermediate input; factor incomes, which should be calculated net of commuting expenses, are overstated to the extent such expenses are not deducted; and value added by productive enterprises is overstated because the payments for the intermediate inputs of commuters' travel have not been properly deducted. Double counting has taken place in this case just as if both the cost of bread and the cost of the flour that went into the bread were added together to estimate total expenditure. Commuters' expenditures, however, are only the most obvious form of double counting. There are quantitatively much more important sources of this error, the most important of which are connected with government expenditure. Such government services as national defense, police, and public health and sanitation are not reasonably enjoyed for their own sake but because their provision makes possible genuine consumption. It is of critical importance to realize that the decision to disallow such government expenditure as final consumption in no way implies that such expenditures are not useful or important. Indeed, an inadequate provision of them will invariably result in a sharp reduction in properly measured final consumption. At the same time, the consistent classification of such expenditures as intermediate prevents any deterioration of a society's position which necessitates higher levels of national defense and police spending from being recorded as a condition that improves social welfare. Accordingly, expenditures on intermediate government services should be excluded from national income, factor incomes should be calculated net of the taxes and other payments necessary to finance these services, and calculations of value added must be made net of the value of such services which are properly considered intermediate inputs. It remains an interesting exercise to recalculate conventional historical NIPA for the major European countries excluding such instrumental and regrettably necessary intermediate expenditure as defense and police services. For example, the levels of regrettably necessary defense expenditures required by the diplomatic and military positions of France and Britain respectively in the years between 1871 and 1914 were sufficiently different to suggest that the differ-

18. Maurice, *Sources and Methods*, pp. 173-76.

ence between the two countries in the level of net output, properly defined, and the level of conventional output, counting defense services in national output as a final good, was sufficiently great that choice of accounting procedure, if applied uniformly in the two countries and making proper allowance for the value of conscripted soldiers' services, would substantially alter perceptions of economic performance.

The existence of important intermediate goods such as national defense and police protection present no difficult conceptual problems. The interpretive decisions are made in determining what is to be designated a final good or service and what intermediate. Very little need be done to the present procedures for collecting statistics for the NIPA. The conventional NIPA can be easily altered to yield estimates of final output, expenditure, and related net factor incomes by re-allocation. However, concern with intermediate goods and services that are properly considered regrettable necessities and instrumental expenditures is closely related to negative externalities and disamenities, the significance of which are much harder to measure, Nordhaus and Tobin, who otherwise are able generally to suggest attractive, operational approaches to national income accounting problems, have no systematic solution to offer for the most fundamental problem that historians must consider in this regard: whether population growth should be seen ultimately as the source of most negative externalities and disamenities or whether population growth genuinely reflects both a society's conscious desire and its physical ability to support more people. Because population-related issues appear so frequently in the conceptually difficult area of negative externalities and disamenities, consideration of this question is fundamental. Wherever decreasing returns to scale exist, or where important resources such as fuels, minerals, and arable land are in inelastic supply, a society's ability simultaneously to support more people *and* to raise real living standards is an impressive economic achievement. Clearly, it costs societies a great deal in terms of forgone consumption to nurture and equip a growing labour force. Using a constant returns to scale Cobb-Douglas production function, which assumes an economy-wide ease of factor substitution rarely encountered in specific industries, and making a variety of assumptions about desired wealth-income ratios for given steady-state rates of population increase, Nordhaus and Tobin estimated that for the relatively moderately growing U.S. population of 1960, a move from an equilibrium population growth rate of 2.14% per year to zero population growth would have raised *per capita* consumption levels by the order of 10%.<sup>19</sup> The assessment of this figure obviously depends upon how off-spring are regarded. The material value of off-spring can be estimated by a variety of means. One of the most promising, for example, would use a sample of linked Census and tax records to calculate age-specific fertility rates across income groups.<sup>20</sup> A finding that wealthier families had on average fewer children and that the fertility patterns of wealthier families were imitated in the rest of society with a lag would lend support to the argument that in such circumstances the observed material cost of child-rearing was greater than the anticipated gain. If the patterns were reversed, with wealthier families for a long period of time having more children

19. Nordhaus and Tobin, *Is Growth Obsolete?*, pp. 18-24.

20. For the possibilities of record linkage in general, see McCloskey, Donald N., *Does the Past Have Useful Economics?* in: *Journal of Economic Literature*, 14 (1976), pp. 441-448.

than poorer families, the conclusion that children were viewed as income elastic consumption goods would be appropriate. In any case, a careful study of the relationship between income levels and population growth for different countries at different periods would offer a much richer data base from which to assess, as Nordhaus and Tobin have done for the U.S., the material cost of child-rearing. These estimates in turn would allow a calculation of the material gains (or losses) that should be contrasted with the negative externalities and disamenities of population growth.<sup>21</sup>

Although at present there are only the crude estimates of consumption forgone in favour of child-rearing derived from simple growth models, Nordhaus and Tobin do offer an illustration inviting imitation of how the complex question of negative externalities might be handled.<sup>22</sup> They note that there appears to be a systematic variation in earnings across U.S. cities of different sizes, with earnings highest in the largest, most densely populated cities most fully exposed to the costs and disadvantages of congestion, pollution, and other negative externalities and disamenities. They then suggest that two sets of factors might explain this pattern. In the first set are those factors unrelated to negative externalities and disamenities: these are taken to be (1) median years of schooling achieved by the labour force; (2) proportion of the population over 65 and presumed to be retired with low earned incomes; (3) proportion of the population Negro and presumed to suffer from non-environmentally related discrimination; (4) the migration rate where a net inflow is presumed to reflect anticipation of high and rising real incomes and a net outflow the reverse; (5) property taxes *per capita*, included to capture the impact of physical capital which can be expected to cause patterns of observed earned income, especially those of self-employed shopkeepers, merchants, and various other types of local businessmen, to vary; (6) local government expenditures *per capita*, included to capture the benefits of public services. The remaining factors are those related to environmental costs for which an earnings premium would be necessary, other things being equal, in order to induce people to work in a less pleasant environment. This second set of factors includes: (1) population size; (2) population density; and (3) proportion of the population in a metropolitan county living within urban boundaries, a variable dictated by the manner in which the data were available. The logarithm of median family income in a sample of metropolitan counties was regressed against the nine independent variables listed above. The estimated coefficients on the three variables in the second set, presumed to reflect negative externalities and disamenities, were then used to calculate the implicit premiums necessary in order to compensate people for living in a more crowded, dangerous, noisy, dirty urban environment. Nordhaus and Tobin estimated that for the U.S. in 1965 the premium was equal to 8% of average U.S. disposable family income and that this figure would have risen to about 30% had the entire U.S. population been concentrated in the most densely populated cities.

The great value of Nordhaus and Tobin's pioneering work is not so much the specific quantitative estimates obtained, although those estimates are of great interest because they represent the most informed evaluation currently available of the im-

---

21. Interestingly, Nordhaus and Tobin suggest that population growth in the recent U.S. past has been as rapid as it has been because the social costs of children have not been borne by parents but by society at large. See Nordhaus and Tobin, *Is Growth Obsolete?*, pp. 18-24.

22. Nordhaus and Tobin, *Is Growth Obsolete?*, pp. 48-54.

pect of many important but hard to measure factors, but rather the opportunity it offers to enrich and extend the already elaborate historical collections of national income statistics.

It is very reasonable to expect, on the basis of what has been done so far, that the process of enrichment and extension will give new meaning and significance to historical data that has not been heavily drawn upon for lack of a systematic means of assimilation and assessment.<sup>23</sup> The NIPA, modified to ensure logical consistency and refocused to measure consumption rather than marketed production, provide a framework capable of processing, categorizing and evaluating data on a much greater scale than has been attempted so far. Furthermore, the national income accounts, in both modified and unmodified form, can be combined with what Mervyn King has described as "social indicators" to assess more broadly and more searchingly trends in welfare.<sup>24</sup> King compiled an index of 17 social indicators scaled such that high values reflected improvement and low values deterioration of welfare. The indicators, chosen for wide coverage across countries rather than for intrinsic importance, included: public expenditure on education as proportion of GNP, students per 100,000 of the population, proportion of total students who were female, doctors per 10,000 persons, infant mortality rate, suicide rate, stomach ulcer death rate, and telephones per 100 persons. King's procedures could be easily extended to cover such indicators as: male and female life expectancy, average length of work week and work

---

23. A variant of Nordhaus and Tobin's procedure has already been applied to nineteenth century Britain by Jeffrey G. Williamson, *Urban Disamenities, Dark Satanic Mills, and the British Standard of Living Debate*, in: *Journal of Economic History*, 41 (1981), pp. 75-83. His conclusions are not dissimilar to Nordhaus and Tobin's. Williamson found that the disamenity premium required to induce workers to endure harsh urban environments was no more than 8% of observed urban wage rates. However, in two important aspects of his study, Williamson appears to have biased downwards his estimate of the disamenity premium. First, he included a cost of living index as an independent variable to capture wage rate variations *not* related to environmental disamenities. But, since the object of estimating the disamenity premium was to determine what proportion of higher nominal wage rates went to compensate for the disadvantage of urban life, which would include high site rents as one aspect of congestion, the highly significant coefficient on the cost of living index should have been used in rather than excluded from the estimation of the disamenity premium, particularly since the cost of living variable is picking up influences that would otherwise be captured by the population density and population size variables. Secondly, he assumed that the variable infant mortality would pick up the main impact of urban disamenities. But many other factors, especially overall fertility levels, which in the short term are at best only remotely related to either wage rates or urban disamenities will also effect infant mortality and this remote relationship will be reflected in a small, relatively insignificant coefficient. On the other hand, Williamson did not control, as Nordhaus and Tobin did, for other factors, notably returns to education and skills, differential labour force participation rates, and migration rates, that would cause wage rates to vary, thus making his results incompatible with Nordhaus and Tobin's and hard to interpret.

Williamson's effort does hold out the promise that further, systematic exploitation of historical data will provide more illumination on this issue. To the extent that Williamson's objective was to provoke further research, his paper is certain to be a success.

24. King, Mervyn A., *Economic Growth and Social Development: A Statistical Investigation*. in: *The Review of Income and Wealth*, 20 (1974), pp. 251-272.

year, occupational accidents, unemployment, divorce rates, strike records, public expenditure on leisure and the arts, and net migration. Such greater coverage would make interpretation easier. For the 17 indicators that he did choose, King found that the movements of his unweighted composite index tended to be positively correlated with conventionally measured economic growth, although the strength of this relationship showed some signs of diminution over time. Such findings, if confirmed more widely in historical studies for a broader range of social indicators would allow researchers to use the conventional NIPA with increased confidence.

Such analysis also suggests a natural and practical way to modify the interpretation of the conventional or modified NIPA when a substantial and persistent divergence in measures does occur. It would thus be possible to isolate those social indicators whose behaviour was particularly badly reflected by the income measures and to identify the causes and to estimate the quantitative consequence of that behaviour.

In this regard, because of the quantitative importance of leisure in Nordhaus and Tobin's study, it is interesting to note the high degree of correlation in late Victorian Britain between the reduction of the length of the average full-employment workweek—a reduction which should be seen as increasing the amount of leisure time potentially available to workers—and the strength of conventionally measured economic growth. Between 1860 and 1914, the length of the average British workweek was reduced by 10%.<sup>25</sup> The reductions occurred in discontinuous bursts concentrated in periods of vigorous conventionally measured growth.<sup>26</sup> The period of greatest reduction, accounting for perhaps two-thirds of the entire reduction in normal working hours achieved between 1860 and 1914 occurred during the intense boom years of 1872–74. On the other hand, virtually no reduction at all occurred when the rate of conventional growth perceptibly declined between the Boer War and the First World War.<sup>27</sup> The most plausible explanation of this pattern is that a reduction in the standard workweek was viewed by employers as a major and essentially irreversible concession to the labour force and hence granted only in exceptional circumstances. In the stable, competitive, environment of the pre-1914 British economy, only during those infrequent times when both workers and employers could anticipate sufficient technological advance to make the bargain feasible were normal hours reduced. Without such expectations employers would, during periods of normal cyclical expansion, “buy off” demand for reduced hours by a combination of higher wages for the normal workweek and an insistence that any reduction in hours be accompanied by a proportional reduction in pay. During periods of cyclical contraction the workers' position was too delicate to withstand the effort to win a major concession. As long as such factors are generally involved in the process of reducing normal working hours, it is reasonable to anticipate that movements in conventionally measured economic growth will generally be found, as they were in Victorian Britain, to be correlated with increases in leisure.

---

25. Bienefeld, Manfred Alfons, *Working Hours in British Industry: An Economic History*, London 1972, pp. 98, 111, 121, 146.

26. Bienefeld, *Working Hours*, pp. 193–94, 197, 201,

27. Bienefeld, *Working Hours*, pp. 146–48.

### III

The problems considered so far have all relied for their resolution, at least implicitly, either upon observed prices directly or upon imputations based on observed prices. The value of leisure, for example, was determined with reference to average income payments for additional work. Where adjustment for enhanced leisure benefit due to advances in transport or consumer durables was necessary, the degree of adjustment was calculated by deflating a suitable consumer price index. Allowances for urban disamenities were evaluated by the observed earnings premium which had to be paid to induce workers voluntarily to endure harsh urban conditions. The use of social indicators was linked to the degree of correlation between measured income and an index of non-market indicators. Thus one way or another, directly or indirectly, observed market prices have represented an indispensable component of the assessment and evaluation of economic activity.

But there are important circumstances where the historian may feel that observed prices are fundamentally distorted and that any economic measure ultimately based on such prices must be misleading. Fred Hirsch has recently presented a detailed critique of economic growth based on a belief in the fundamental fallibility of observed market prices.<sup>28</sup> Hirsch argues that much of the consumption desired in modern economies is centred on "positional" goods and services whose value is determined by the satisfaction they can provide through relative position alone, from the quality of being in front or from the fact of others being behind.<sup>29</sup> In this view the real value of, for example, a desirable home site or the rewards of a responsible and fulfilling job cannot greatly change. The value of purely positional goods is taken to be almost completely independent of technological change. If consumers can use only so many television sets or so many cars, after which the value of another such good becomes virtually zero, then the limits to growth are clearly and unalterably set.

This argument is tantamount to the claim that if equilibrium prices could only be known, the high and rising relative value attached to those things whose output could not be increased would be clear, as would the low and falling relative value of those things whose output could be increased. Ultimately, in "true equilibrium" value terms, real growth would be impossible as long as substitution in consumption between positional and material (or reproducible) goods was strictly limited. In this situation, the appearance of growth can only be an illusion created by weighting the various categories of output by prices in fundamental disequilibrium. Such a procedure would give current period weights to positional and reproducible goods, ignoring the fact that as reproducible output is increased the relative value of such output would fall while that of positional goods would rise. Current period disequilibrium prices undervalue positional goods and overvalue reproducible goods thereby creating a "mirage" of economic growth, since the price weights attached to positional goods, whose output cannot increase, rise over time whereas the price weights of reproducible goods, whose output can increase, fall over time. The presumed benefits of growth thus continually recede even as strenuous efforts are made to realize them.

---

28. Hirsch, Fred, *Social Limits to Growth*, London, 1977.

29. Hirsch, *Social Limits*, p. 20.



The implications of this line of argument for the interpretation of the national income accounts is serious, for it implies that the use of the observed prices upon which the accounts are necessarily based are fundamentally incorrect and misleading. Clearly Hirsch has identified a process which accounts at least in part for the inability of even very rapid sustained economic growth to provide the full anticipated benefits. Yet the claim that positional goods are so dominant in aggregate consumption and so immune to substitution by reproducible goods as largely to remove the possibility of real economic growth is an empirical question which must be resolved not by assertion but by extensive historical and international comparisons of economic development. Such research would probe the strangely ahistorical nature of Hirsch's critique. On the one hand, the benefits of economic growth in the past are acknowledged but on the other it is felt that continued benefits of growth cannot be expected in the future. But why should such a situation occur now and not a quarter of a century earlier or later? Comparative historical research would establish the extent to which expenditure on specific positional goods actually has comprised, after allowing for the effects of population growth which amplify the inherent scarcity of positional goods, a stable or rising proportion of total real income, as Hirsch's analysis predicts should occur.

Although it is not possible to anticipate completely the results of research still to be carried out, it would appear Hirsch underestimated the ability of modern technology to create substitutes for positional goods. Since Hirsch often illustrated his argument by reference to the example of a limited number of desirable sites for houses, it is useful to recall how transport improvements, trains in the nineteenth century and cars in the twentieth, by making accessible desirable sites that were previously too remote or inaccessible for extensive use have increased the stock of choice sites, thereby undermining in this instance the very concept of positional goods. Similarly while it remains an open question whether the relative availability of rewarding and fulfilling jobs and occupations has increased as measured growth has occurred, it is clear that modern household appliances, like refrigerators, washing machines and dishwashers, whose production and sale if not use is fully recorded in the conventional NIPA, have removed a substantial amount of the drudgery that blighted and limited the lives of most people in the past. Furthermore, it would appear that the greater variety of consumption goods that has become available over time has acted to diffuse both the pleasure of possessing highly esteemed positional goods and the dissatisfaction of not doing so. After all, as R. C. O. Matthews noted in his generally appreciative review of Hirsch's book, in an economy where no growth takes place, all goods are positional and the possibilities for substitution among consumption goods is much more harshly limited than is true in an expanding economy.<sup>30</sup> Finally, it should be noted that health care and education, two services whose provision has moved in close parallel with movements in the conventional income accounts and are thus manifestly not positional goods, have come to account for a larger share of output in most countries over time.<sup>31</sup> Ultimately Hirsch's argument serves to stress the

---

30. Matthews, R. C. O., *Review of 'Social Limits to Growth'*, in: *The Economic Journal*, 87 (1977), p. 576.

31. Some care must be taken when measuring these services to avoid double counting. Thus education that is solely for occupational and professional advancement should be consid-

importance of careful construction of indices of economic activity, supported by close observation of consumption patterns over time, a task in which Simon Kuznets' pioneering work should offer a most useful base for further research.<sup>32</sup> If anything, Hirsch's reservations concerning the desirability or even the possibility of economic growth provides further rationale for a revision of the conventional NIPA along the consumption-oriented lines proposed by Nordhaus and Tobin.

There is, however, a dynamic variant of Hirsch's argument which increases even further the importance of comparative, historical research. The variant was first given explicit expression by E. J. Mishan and subsequently was formalized by Stephen Glaister.<sup>33</sup> The phenomenon that Mishan and Glaister were concerned with was the potentially unstable nature of desirable economic equilibria and the subsequent likelihood that normal competitive behaviour would lead to the abandonment of desirable equilibria in favour of substantially less desirable equilibria which would then be very difficult to change. Indeed, in the final, undesirable, equilibrium position, the structure of prices would create a very strong disincentive for any change. The argument is most easily grasped in the form of an example, but it is readily seen that generalizations can easily be made. Consider a transport system where no private vehicles exist and only public trams and buses are available. Now suppose that one person suddenly realizes that with a private car he could reduce his travel time to work by half, so long as his were the only private car on the road. If, however, many others shortly afterwards make the same discovery independently and attempt the same action, without reckoning on the congestion costs, the anticipated benefits will prove illusory for all. The congestion caused by only one car will be negligible, but the effect of many people simultaneously switching to private cars, even if each is correct in realizing that the impact of his own action alone is trivial, is not at all negligible. The trouble is clearly that travellers are making decisions on information that will begin to change and be incorrect as soon as the decisions are made, yet no individual traveller can by himself know what the final outcome, and hence what the correct information for a rational decision, will be. In the example, eventually a new equilibrium is reached where many people use private cars, many fewer than previously use public transport, and travel time for all is increased. The previous equilibrium, where no private transport existed, is actually superior to the one that eventually emerges from the introduction of private transport because of the unregulated increase in congestion. But for the same reasons which created the problem in the first place, the original equilibrium is very difficult to regain. The benefit perceived by any one individual in taking public transport rather than his own car is negligible but if many could be persuaded to take public transport, all would gain, those continuing to use their

---

earned incomes of those who received the education and should not, therefore, be counted separately. The educational expenditure that should be recorded as final output must be limited to that which enhances leisure and living in general. Similarly, expenditures on health and medicine necessitated by environmental deterioration and occupational hazards must be excluded.

32. Kuznets, Simon, *Modern Economic Growth: Rate, Structure and Spread*, New Haven/Conn. 1966, pp. 262-284.
33. Mishan, E. J., *The Costs of Economic Growth*, Harmondsworth/Middlesex/England 1969, pp. 232-240, and Glaister, Stephen, *Transport Pricing Policies and Efficient Urban Growth*, in: *Journal of Public Economics*, 5 (1976), pp. 103-117.

own cars gaining most. Hence everyone hopes that everyone else will take public transport, but no one does and the situation never improves. These obvious difficulties are greatly compounded if public transport is competitively eliminated as travellers switch to private cars. The greater is the proportion of fixed to total costs of public transport, the more dependent it is on intensive utilization and hence the more vulnerable its revenue to any decline in traffic. The elimination of public transport, by removing any choice in transport mode, clearly is the worst outcome, yet one which may be difficult to prevent without the imposition by a central authority of a set of "shadow" transport prices which take account of the costs of congestion and the importance of choice. The calculation of efficient shadow prices is not simple. The planning authority must correctly price a variety of transport systems whose operations it cannot directly observe but must deduce from knowledge of both the technical characteristics of transport equipment and the preferences of travellers and shippers.<sup>34</sup>

In the Mishan-Glaister example, only with correct equilibrium shadow prices could the benefits of current consumption be evaluated and plans for future investment rationally made. Yet the congestion externalities that caused such price calculations to be so difficult to make are likely to be a pervasive feature of modern economic life and to intensify as development proceeds, for they are, as in the Mishan-Glaister example, created as a by-product of the same process of technological change that is the source of economic growth in the first place. But if prices are systematically distorted in the way suggested, the national income accounts will not just measure the wrong thing; movements in the NIPA may even be perverse in relation to the real underlying economy. In the example above, when private transport was introduced, measured expenditure on private transport rose, measured expenditure on public transport fell, and the benefits of private vehicle ownership increased. But the increased benefits of private vehicle ownership were due to the deterioration of public transport and it can be shown that the net change in total social welfare may easily be negative even when observed total expenditure has risen, because more consumer's surplus is being extracted by the provision of transport services than was true previously. In such a situation, the same observed prices that lead to suboptimal resource allocations will yield incorrect and inappropriate national account estimates.

As the Mishan-Glaister example suggests, and Levin's study of railroad deregulation confirms, the analytical derivation of even a few of the prices needed for rational decision-making and accurate NIPA is a difficult undertaking. For the foreseeable future it will not be realistic to anticipate plausible calculations of equilibrium prices.<sup>35</sup> Thus the only reliable means of assessing the market prices the historian ob-

---

34. See Levin, Richard C., *Railroad Rates, Profitability, and Welfare Under Deregulation*, in: *Bell Journal of Economics*, 12 (1981), pp. 1-26, for a discussion of the complex considerations involved in such analysis.

35. A recent paper by James, John A., *The Welfare Effects of the Antebellum Tariff: A General Equilibrium Analysis*, in: *Exploration in Economic History*, 15 (1978), pp. 231-256, has demonstrated that equilibrium price vectors of interest to historians can in principle be calculated. It is likely that such work will become more common in the near future, offering historians a very powerful new analytical tool.

serves require extensive international, historical comparisons. While it is, of course, always possible that all countries will make the same mistakes and move towards the same inferior equilibrium, it is not likely. Thus examples of unusually successful organization or of notably sustained advance may be effectively used to investigate the nature of feasible equilibria. While this procedure may fail to capture the best possible outcome available to an economy at any point in time, it is unlikely that a large number of societies with different economic, political and technological arrangements and capabilities will all miss the best solution equally badly. The results of such comparisons may be difficult to interpret, but, since cross-country historical comparisons are as close as the economic observer can get to a controlled experiment, there is very little choice but to use them.

If the analysis of economic performance inevitably requires extensive international comparisons, more attention must be devoted to the notoriously difficult task of improving the statistical basis for such comparisons. While no comprehensive solution exists, two recent proposals promise marked improvement. These proposals will also aid in the comparison of a given economy's performance at different points in its own history.

The first proposal advocates facing directly the problem that price structures are determined jointly by the distribution of wealth and productive capacity. When comparing the efficiency of two economies, performance is to be measured by reference to how well the needs of specific groups common to both are met. Such an approach has been used by A. K. Sen to compare inter-state disparities of welfare standards in India.<sup>36</sup> Sen explicitly gave higher consumption weights to the relatively poor so that his comparison of performance was particularly sensitive to the consumption experience of his benchmark social group. Such a procedure has the great benefit of making the welfare basis of comparisons explicit. Of course the assessment of performance may vary with the benchmark consumer group chosen, but even this factor is beneficial, for historians have often realized that common developments have differential impacts and the use of several benchmark groups, to the extent that experiences are sharply different, serves to indicate quantitatively the variety of experience the historian wants to examine. Furthermore, because the approach using benchmark groups relies on knowledge of expenditure patterns, knowledge which is largely independent of the conventional accounts, it offers a valuable supplementary cross-check to those accounts. Also it can readily be combined with the second recent proposal to aid international comparisons.

The second proposal has used detailed calculations of purchasing power parities between countries in order to obtain a more "realistic" set of exchange rates.<sup>37</sup> Such adjustment is necessary because tariffs, quotas, and other forms of administrative intervention—such as central bank operations—distort exchange rates from the levels that would rule if only pure demand and supply factors were operating. Furthermore, even if exchange rates were in fundamental real equilibrium, countries with different productive structures or consumption patterns may have sufficiently different ratios

---

36. Sen, Amartya K., *Real National Income*, in: *Review of Economic Studies*, 43 (1976), pp. 19–39.

37. David, Paul A., *Just How Misleading Are Official Exchange Rate Conversions?* in: *The Economic Journal*, 82 (1972), pp. 979–90.

of tradeable and non-tradeable production that cross-country comparisons based on exchange rates alone may be misleading or unnecessarily incomplete. The procedure proposed by Paul David to adjust market exchange rates uses the relative price weights that obtain in the benchmark country—in his case, the U.S. on the grounds that by common consent for the period he was considering (1950–1965) the U.S. was internationally the most advanced economy, with its structure increasingly approximated by the rest of the world—to value output and consumption in other countries. The levels of output in all countries are then compared using the benchmark country's prices and the relative rankings of countries obtained on this basis are compared with the relative rankings obtained by valuing national outputs at market exchange rates. Specifically, the following equation was estimated:

$$\left\{ \frac{Y_o}{Y_i} \right\} = \frac{0.671}{(0.063)} + \frac{0.408}{(0.022)} \left\{ \frac{Y_o}{\hat{Y}_i} \right\}$$

where  $Y_o$  = *per capita* output in the benchmark country valued at the benchmark country's relative prices.

$Y_i$  = *per capita* output in the *i*th country valued at the benchmark country's relative prices.

$\hat{Y}_i$  = *per capita* output in the *i*th country valued in the benchmark country's currency using the market exchange rate rather than the benchmark country's relative prices.

The numbers in brackets under the estimated coefficients are standard errors. The equation may be interpreted to show that on average the difference in *per capita* incomes in purchasing power parity terms<sup>38</sup> between a given country and the U.S. was only 40.8% of the percentage gap indicated by a straight exchange rate conversion expressing all incomes in dollars. The explanation for this sharp reduction in income differentials using purchasing power parity price weights and market exchange rates is that the higher manufacturing productivity in the U.S. also raised the opportunity cost of providing labour intensive services which accounted for a substantial proportion of total consumption not only in the U.S. but also in all the other countries in the comparison. When the U.S. price weighting, with its relatively high cost of services and low cost of manufactures was used to value output in various developed countries, those countries gained more in the revaluation of their comparatively abundant services than they lost in the revaluation of their comparatively sparse manufacturing output.

David's procedure goes an important way towards adjusting exchange rate data to permit its use in a much more meaningful way than was previously possible. The full potential in this procedure is gained when different countries are used as bench-

38. A classic index number problem prevents a true estimation of purchasing power parity. That would require knowledge of what the citizens of one country would have bought had they faced with their incomes the price structure of another country. What purchasing power parity means here is a measurement of the bundles of output in two different countries by the same price structure. It is clear that even this restricted purchasing power parity concept allows a significant improvement on official exchange rate estimates of *per capita* income differences.

marks in repeated calculations. Comparisons of the rankings obtained for a variety of benchmark countries will offer a systematic means of identifying and assessing important differences in the consumption and production structures of different countries at different points in time. Furthermore, David's procedure can readily be combined with Sen's so that the purchasing power parity calculations are based on explicit welfare orderings rather than observed price structures which impose a welfare basis founded on the existing distribution of wealth.<sup>39</sup>

As may readily be seen, these various procedures cannot be vested with an aura of infallibility and precision. In their way, they are as arbitrary as the conventional accounting procedures they are intended to supplement and replace. Their great virtue, however, is that used in combinations they each illuminate in a different way a particular aspect of a nation's (or a region's) historical experience. They allow most of the subjective and judgemental differences that divide historians to be systematically exposed and quantified. They permit sensitivity tests to determine which assumptions are crucial to conclusions and which affect conclusions very little. They use in a structured way vast amounts of historical data that simply cannot be assimilated without an explicit analytical framework. The results of these efforts may never command universal agreement, but the issues in dispute among historians will become systematically clearer because of them. And that surely can only be considered progress.

Zusammenfassung:  
Probleme der volkswirtschaftlichen Gesamtrechnung und  
ihrer Interpretation bei der Bewertung langfristiger  
wirtschaftlicher Leistungen

Hier wird davon ausgegangen, daß es bei der Messung ökonomischer Aktivitäten im wesentlichen zwei Arten von Problemen gibt. Die erste besteht aus solchen Fragen, die zumindest prinzipiell dadurch lösbar sind, daß man beobachtete oder abgeleitete Marktpreise verwendet. Der zweite Problemtyp besteht aus Fragen, die nur über die Verwendung geschätzter (oder „synthetischer“) Preise gelöst werden können, weil die beobachteten Preise aus vielerlei Gründen fehlerhaft sind. So könnte z. B. ein fundamentales Marktungleichgewicht vorliegen oder eine nichttolerierbare Abhängigkeit von einer besonderen – möglicherweise willkürlichen – Vermögensverteilung. Oder es könnten Verzerrungen durch Schwankungen der relativen Preise oder auch durch eine allgemeine Inflation oder eine Deflation hervorgerufen worden sein. Derartige Verzerrungen können auch von Zöllen, von festgesetzten Handels- oder Produktionsquoten und unterschiedlichen anderen administrativen Eingriffen in den Markt verursacht werden. Beide Problemtypen werden diskutiert und mit neueren Lösungsversuchen dargestellt. Am Schluß der Arbeit finden sich einige Vorschläge zur weiteren Forschung.

---

39. Incomes are converted into wealth equivalents by using the present discounted value of future income streams.

Fragen, die zum ersten Problemtyp zählen, sind nach zwei übergeordneten Gruppen klassifiziert. Die erste kleinere Gruppe befaßt sich mit logischen Widersprüchen in der herkömmlichen volkswirtschaftlichen Gesamtrechnung, die bisher deshalb hingenommen wurden, weil die verfügbaren Daten so leichter zu erfassen waren. Die zweite weitaus größere und wichtigere Gruppe befaßt sich mit Problemen, die aus der theoretisch unangemessenen Zielvorstellung über die zu messende Größe erwachsen. Eigentlich sollte der Endverbrauch erfaßt werden, doch bezieht sich die herkömmliche volkswirtschaftliche Gesamtrechnung im allgemeinen auf die vermarktete Produktion. Die Diskussion der veränderten Meßmethoden, die bei einer Abänderung der Zielgröße für die volkswirtschaftliche Gesamtrechnung erforderlich wären, schließt die Darstellung der Pionierarbeit von Nordhaus und Tobin ein. In diesem Abschnitt werden Punkte abgehandelt wie die Bewertung von nicht-vermarkteter Haushaltsarbeit oder von Freizeit sowie Kosten und Nutzen des Bevölkerungswachstums. Am Schluß wird die Möglichkeit erörtert, den Bereich der bisherigen volkswirtschaftlichen Gesamtrechnung durch die Einbeziehung sozialer Indikatoren zu erweitern.

Fragen, die zum zweiten umfassenden Problemtyp zählen, werden in Anlehnung an das kürzlich erschienene Werk von Hirsch, Mishan und Glaister erörtert und mit den Lösungsvorschlägen von Sen und David vorgestellt.

Nun können Historiker keine eindeutigen definitiven Antworten auf die Mehrzahl der Probleme in der volkswirtschaftlichen Gesamtrechnung, wie sie in diesem Artikel angesprochen wurden, erwarten. Die Vorschläge, die hier zu verschiedenartigen Revisionen und Modifikationen der herkömmlichen volkswirtschaftlichen Gesamtrechnung gemacht wurden, sollen vielmehr dazu dienen, daß die subjektiven Meinungsunterschiede (die Historiker voneinander trennen) in systematischer Weise offen dargelegt und quantifiziert werden. Damit ließen sich fundamentale Fragen und Probleme viel klarer umreißen, wenn nicht endgültig lösen.