CSLS Conference on the State of Living Standards and the Quality of Life in Canada

October 30 - 31, 1998 Château Laurier Hotel, Ottawa, Ontario

Centre for the
Study of Living Standards
Centre d'étude des
niveaux de vie

GDP and Its Derivatives as Welfare Measure: A Selective Look at the Literature

Abe Tarasofsky
Statistics Canada

Session 2: "Indicators of Well-being"
October 30 10:30 AM - 11:30 AM

GDP AND ITS DERIVATIVES AS WELFARE MEASURES: A SELECTIVE LOOK AT THE LITERATURE

Abraham Tarasofsky

Statistics Canada

Paper prepared for the Centre for the Study of Living Standards Conference on the State of Living Standards and Quality of Life, Ottawa, Ontario, October 30-31, 1998. Helpful comments were received from Hans Adler, Zvi Griliches, Kishori Lal, Hans Messinger, Robin Rowley, David Slater, Neil Swan, Bert Waslander, Stewart Wells, and Miles Wisenthal. Patricia Sherar provided indispensable editorial assistance.

TABLE OF CONTENTS

Introduction 1

Part I: GDP Defined 3

Some Investigations of the Question: Does a Given Change in GDP Represent a Change in Welfare?

Part II: Modifying GDP with a View to Converting it into a Measure of Economic Welfare 10

Part III: Further Arguments Against Converting GDP into a Welfare Measure 33

Part IV: Concluding Questions and Answers 40

Is GDP an Adequate Measure of Economic Welfare?
Can Some Other Measure, Derived from GDP, Serve as an Adequate Measure of Economic Welfare?
Can Any Single Measure of Economic Welfare be Adequate?
How Should Economic Welfare Be Measured?

Appendix I: Index Numbers and the Measurement of Changes in the Level of Economic Welfare 55

Appendix II: Summarized Scenarios Involving Intermediate Products, Final Products and Government Services 56

Appendix III: Some Further Thoughts on Defence Expenditures 59

Appendix IV: Three Indexes 63

References

INTRODUCTION

Considerable work was done at Statistics Canada in the wake of the widely-read Cobb, Halstead and Rowe (1995), who reject GDP as an adequate welfare measure and propose their so-called "genuine progress indicator" (GPI) as an alternative. I participated in this work, which largely consisted in reproducing GPI and some other well known alternatives to GDP on the basis of Canadian data and became unhappy with some aspects of their respective rationales. Remembering from my student days that a good deal of relevant discussion had taken place some decades ago, some of it involving illustrious theorists, it occurred to me that at least some of the matters that bothered me had been addressed in that literature.

It should, of course, never be suggested that the last word has been said, or ever could be said, on any subject, but so rich is this discussion and so varied the arguments advanced, that it is hard to imagine much that is both new and of fundamental analytical importance being said in the foreseeable future on the subject of GDP's or any derivative measure's adequacy as a measure of economic welfare. At the very least, it can be confidently asserted that there is a great deal to be learned from a review of the diverse methodologies and arguments set forth in this literature. Whilst this paper is by no means an exercise in the history of the relevant economic thought, it selectively draws upon that history in order to convey a sense of some of these methodologies and arguments and, more specifically, to advance the following four propositions.

1) Being first and foremost a measure of output, GDP can only be conceived of as a measure of economic welfare in a limited, and ultimately unsatisfactory, sense;

2) none of the better- known alternative measures, derived from GDP, fares better; indeed, each, in its own way, fares worse;

3) there is no such thing as a single, and satisfactory, measure of economic welfare; only a set of complementary variables, one of which could be GDP, has the potential for conveying an

adequate sense of a country's economic welfare; and

4) there is no such thing as an entirely objective set of complementary variables; some cannot be estimated in monetary terms and subjective factors cannot be eliminated.

The paper is in four parts. After defining GDP, Part I summarizes the methodologies and rather discouraging results of four well-known investigations of whether, and under what conditions, changes in GDP qualify as valid measures of changes in welfare. Part II focuses on a well-known example of a genre of analyses that explicitly maintains that GDP is not a sound measure of economic welfare and modifies it in various ways with a view to producing one. Some of the most important adjustments that collectively comprise this overall modification involve government expenditures included in GDP, so some of their rationales are discussed. Part III presents various defences of the *status quo*, and the concluding Part IV deals with the questions and answers underpinning the four above propositions.

PART I

GDP DEFINED

GDP and its forerunner, GNP (henceforth, the term "GDP" will be used exclusively, although the terms actually used in most of the following references are either "GNP" or "national income"), have been conceived along very similar lines since their inception. As set forth in its latest authoritative definition (in the 1993 United Nations System of National Accounts (1993 SNA)), GDP may alternatively be conceived of as four equivalent sums:

(a) final results at market prices of the production activity of resident producer units;

(b) gross value added namely, defined as the difference between output and intermediate consumption;

(c) final uses of goods and services (all except intermediate consumption) measured in purchasers' prices, less the value of imports of goods and services; and

(d) primary incomes distributed by resident producer units.

Foreshadowing an issue that was soon to emerge, and which has loomed large down the generations to this day, Kuznets (1933) warns against the temptation of attempting to translate GDP into a welfare measure. A welfare measure:

> ...runs in terms of subjective feelings, whose commensurability for various individuals is to be doubted and whose relation to the objectively perceptible economic goods is not, in the present state of knowledge, determined with sufficient precision to permit even purely qualitative economic analysis. Consequently, the concept of *income enjoyed* has to be abandoned....(p. 7, emphasis added)

As will be seen below, Kuznets's views on this issue evolved over the years (and they also evolved on other, related issues). Kuznets (1941), for example, attributes an objective character to the foregoing subjective evaluations.

> ...Market prices are of course a far from perfect measure of how well goods satisfy society's needs. But they are the sole practicable basis *if the estimator is to follow the consensus of social opinion*. (p. 55, emphasis added)

GDP is defined above with a view to confining it to marketed economic activities that are "productive," the latter criterion serving to exclude a variety of transfers not made as recompense for productive activity. It is also implicit in these alternative versions of GDP that the primary focus is upon the measurement of specified transactions and categorizing them within an accounting framework. Whilst this focus does not, in itself, necessarily preclude welfare inferences, neither does it necessarily foster them.

SOME INVESTIGATIONS OF THE QUESTION: DOES A GIVEN CHANGE IN GDP REPRESENT A CHANGE IN WELFARE?

THREE QUALIFIED AFFIRMATIVE ANSWERS

Hicks

Hicks (1940) is one of the earliest attempts rigorously to examine whether changes in GDP necessarily imply analogous changes in welfare. By means of both an indifference map and index numbers, two situations are compared. These situations first involve an individual choosing among many combinations of two commodities; they then involve a group of individuals choosing among combinations of many commodities. Subject to the condition that everyone's wants remain constant throughout, Hicks contends that the combination located on the higher indifference curve (Situation II) is preferable to the combination located on the lower (Situation I), provided the following additional criterion is also met:

> *...Let us say that the real income of society is higher in Situation II than in Situation I, if it is impossible to make everyone as well off as he is in Situation II by any redistribution of the actual quantities acquired in Situation I.* (p. 111, emphasis added)

Restating the argument in terms of index numbers, Hicks argues that the above welfare result unambiguously holds only if the ratio of Situation II's total expenditures to those of Situation I (designated as "E") exceeds both the corresponding Laspeyre's and Passche price indexes. Any other relationship either implies that Situation II is inferior to Situation I or that the relationship between the two situations is ambiguous. The argument is set out more fully in Appendix I. Hicks also assigns a crucial role to commodity prices as weights in welfare measurement: with one or two exceptions (e.g., supply-restricted or new commodities), they are deemed proxies for the corresponding marginal utilities.

Prompted by the "apparent inconsistency" (which he considers justified) of valuing private goods and services at market prices and public services at cost, Hicks also looks at the welfare implications, now conceived in terms of productivity, of an increase in GDP from the standpoint of factor cost. He finds that so long as constant returns (namely, straight-line substitution curves) and unchanging technology are assumed the preceding index-number result holds: otherwise, it does not. Nor does it hold under conditions of perfect competition and universal diminishing returns. Now, the substitution curves are not only concave to the axis, they also intersect, a serious problem.

> ...Thus even if E is greater than both the price-indexes, we cannot be certain that there has been a *definite* increase in productivity. (p. 121, emphasis in text)

Matters are "much worse" when imperfect competition and increasing returns are contemplated. In sum, and contrary to what is widely believed:

> The measurement of economic welfare and the measurement of productivity are in fact quite different things; their coincidence under optimum conditions is a special property of that particular state of affairs. (pp. 122-123)

Kuznets

Before summarizing the comments of Kuznets (1948) on the preceding analysis, it is worth briefly citing a important caveat in the earlier Kuznets (1946) concerning income distribution.

> *...[GDP] as a whole retains meaning only in so far as the national distribution by size of personal income shown tendencies towards stable patterns*. But from any point of view such a frequency distribution is an indispensable *complement* of national income estimates if these are to throw any light on the welfare of the nation. *Welfare is an actuality only within the experience of every individual and varies materially with the size of a person's income*. (pp. 33-34, emphasis added)

This caveat is especially important because Kuznets proceeds to report that empirical analyses of historical distributional patterns of various countries find that the required tendency toward stability does not exist. This finding strengthens his already substantial doubts as to whether intertemporal comparisons of a country's GDP are likely to provide a satisfactory basis for tracking its welfare trends. (His doubts about the soundness of international comparisons for this purpose are even stronger.) The notion of reporting additional, distributional data alongside GDP, as a complement to it, has important implications that will be discussed in due course.

Suppressing his doubts, Kuznets (1948) considers Hicks's analysis and argues that two problems render insufficient his criterion for Situation II's superiority over Situation I (that no redistribution in Situation I may result in higher total welfare than the total welfare achieved in Situation II). The first problem is that nothing normative can be said until the total-expenditure comparison has been reformulated in per capita terms; the second, and more important, problem is that the comparison must also be subjected to what Kuznets calls a "base reversal" test. This test determines whether it is impossible to make *everyone* as well off as he is in Situation I by any redistribution of the actual quantities acquired in *Situation II*. If it is impossible to accomplish this, Situation I is superior to Situation II. Since a comparison that produces this result may satisfy Hicks's criterion that renders Situation II superior to Situation I, there is a clear contradiction. To illustrate how such a contradiction could arise, Kuznets divides the population into two groups, rich and poor, and postulates that although both groups consume goods called "necessities," other goods called "luxuries" are only consumed by the rich. He then conjures up circumstances in which the satisfaction of Hicks's criterion proves to be insufficient to establish unambiguously an increase in total welfare, even though the total-expenditure ratio exceeds both the corresponding Laspeyre's and Paasche indexes. The essence of these circumstances is that the poor are worse off in Situation II than in Situation I and the rich are better off (proportionately more so). Since the poor do not consume "luxuries," redistributing these goods among the rich cannot restore the poor to their Situation I condition. The only way the poor can be so restored is by breaking down the distinction between necessities and luxuries and then making the necessary redistribution. Thus:

> ...If we are to determine unequivocally an increase in welfare in Situation II over Situation I we must assume not only a constancy of wants in the sense of a constancy of each individual's appraisal of different goods but also that either (a) all goods can be substituted for one another in the full range or (b) the structure of the goods aggregate in the two situations is such that no specific good, to the extent that it cannot be replaced by another, is reduced in output....(p. 4)

This finding has two important implications. First, drastic and regressive distributional changes during the interval between the two situations, which cause a large decrease in the production of "necessities," are ruled out. They may render inapplicable Kuznets's "base reversal test" and, thereby, obscure the fact that an unambiguous normative judgment cannot be made. Second, although government output is deemed ultimately to flow from consumers' decisions, drastic changes in the distribution of output between the private and public sectors are also precluded for much the same reason.

Little

The discussion is extended by Little (1957), who argues that Kuznets's requirement of full mutual substitutability of goods is unrealistic. Little's main focus, however, is on more fundamental issues, such as the very meaning of questions about welfare. He asserts that, strictly speaking, there is no objective measure capable of capturing changes in the level of total welfare. Although he generally prefers Kuznets's formulation to Hicks's, he emphasizes that serious problems remain, especially when the patently unrealistic implicit assumption that the community has not changed, even over a short period, is abandoned. People are constantly dying and being born. Hence:

> ...If we are considering whether we are better off than our grandfathers, it is silly to ask whether everyone could be made as well off as they were before. Logically, it is just as silly to ask the same question when we are considering whether the welfare of the community has increased over last year. *If a group is changing in composition, one cannot seriously consider the question of its welfare without making interpersonal comparisons....* (p. 17-18, emphasis added)

And, of course, interpersonal comparisons are impermissible, and the previously-noted caveats concerning demographics and substitutability of goods also apply.

TWO FIRMLY NEGATIVE ANSWERS

Samuelson

Samuelson (1950) presents an ostensibly more comprehensive model than the Hicks-Kuznets models that rely upon index numbers. Its conceptual heart consists in two aggregate functions applicable to all members of society: a so-called "utility-possibility function" and a so-called "utility-feasibility function," the latter function lying within the former in the region of the two situations being compared. Samuelson's findings, which explicitly recognize, *inter alia,* that income-distribution policies necessarily impose costs, a fact generally ignored by other analysts, are particularly discouraging.

> ...[W]e come to the paradoxical conclusion that a policy which seems to make possible greater production of all goods and a uniformly better utility-possibility function for society may result in so great (and ethically undesirable) a change in the imputation of different individuals' incomes, that we may have to judge such a policy 'bad'...A policy that shifts society's utility-possibility function uniformly outward *may not at the same time shift the utility-feasibility function uniformly outward, instead causing it to twist inward in some places*. (pp. 20-21, emphasis added)

<u>Chipman and Moore</u>

Also in the context of a purportedly comprehensive model, and building upon preceding analyses, Chipman and Moore (1971) present their own discouraging conclusion: the statement that *potential* total welfare (the only welfare that matters) has increased because real GDP is greater in Situation II than in Situation I depends crucially upon all individuals having *identical* tastes that are also homothetic. This is an even more stringent condition than the condition of constant tastes. Consequently:

> ...[T]he sort of identification of increases in potential welfare with increases in real [GDP] which has been traditional in the profession is on very shaky ground...[A]n increase in [GDP] is a *necessary* condition for an improvement in potential welfare, even though it is not, in general, a sufficient condition for inferring that such an improvement has taken place. (p. 415, emphasis in text)

Chipman and Moore further put matters in perspective by grouping into two categories the difficulties that arise from the attempt to infer changes in welfare from changes in GDP.

> That [GDP] is a highly imperfect measure [of welfare] has never been denied...How best to evaluate nonmarket goods such as housewives' services, imputed rent, leisure, or externalities such as smoke, noise and other kinds of pollution, has never been satisfactorily resolved. Another big problem is posed by technological change, change in product quality, and changes in tastes. More difficult still is the valuation of public goods such as parks, police services, and national defense. However...it has generally been believed that while

these are difficult conceptual and practical problems, nevertheless they are capable of solution, and...then we would have an index which would provide a correct indication, if not of actual welfare, at least of potential welfare.

It is our contention that even if all the above problems...could be solved, there remain further difficulties of a fundamental nature in the use of [GDP] as a measure of potential welfare, let alone actual welfare...*In short, if an index of welfare is what we want, we cannot rely on [GDP] alone, and sometimes we cannot rely on [GDP] at all*. (*P. 393,* emphasis added.)

PART II

MODIFYING GDP WITH A VIEW TO CONVERTING IT INTO A MEASURE OF ECONOMIC WELFARE

Much attention is given in somewhat more recent literature to various problems included in Chipman and Moore’s first category, in the belief that solving them will produce a satisfactory welfare measure. With few exceptions, the analysts in question tend, in effect, to proceed as if Chipman and Moore’s “fundamental,” second-category problems do not exist: they are not discussed, much less resolved. In spite of this regrettable situation it is highly instructive to suspend disbelief and consider their work on its own merits, largely, though not entirely, in isolation of the earlier work.

Unlike of the analysts discussed above, these analysts typically do not compare the GDPs of two successive periods and consider whether the change implies an analogous ordinal change in welfare. Instead, they typically begin with that major component of a given year's GDP namely, consumption, subtract some of GDP’s components, wholly or partly, add or subtract other variables and, then, arrive at their respective welfare measure for that year. In doing so, these analysts in effect reject the proposition that individual and collective market choices are the *sole* determinants of welfare, and supplement these choices with various choices of their own, which vary greatly among analysts. In addition, they in effect deem each year's measure a cardinal measure and, therefore, deem the year-to-year changes in it as being cardinal as well. It is impossible here to consider all the additions and subtractions that can be found in this genre of welfare measures (three examples of such measures are presented in Appendix IV) but some of the most important and far-reaching ones can be discussed collectively under the general rubric of government services. Since, as a rule, their analyses do not approach the preceding ones in rigour, it is instructive briefly to consider some closely-reasoned, conceptual justifications for using a derivative of GDP as a welfare measure and, in particular, the assumptions upon which they rest.

Usher

Usher (1980), takes passing cognizance of the analyses summarized above and declares them peripheral to his own. His main focus is upon the measurement of economic growth. He holds, however, that a given increase in real income, deemed synonymous with GDP in an initial context of an artificial world in which consumers spend their entire income, reflects a corresponding *cardinal* increase in welfare. He specifies four alternative definitions of real income and assesses the realism of the assumptions upon which, under each definition, intertemporal and international welfare comparisons rest, both with respect to his so-called "representative consumer" and to the aggregation of all consumers that is the community. Under the first definition, the support he gives to two key assumptions, that the consumption of his representative consumer and that of the average consumer in the community are the same, and that the former's rates of substitution in use among commodities are reflected in observed relative prices, seems rather lukewarm. Under the second definition, he acknowledges that the informational requirements of aggregating from individual consumers to all consumers are enormous, but contends that this difficulty can be surmounted by resorting to a Lespeyres index for the community. He also acknowledges that the coherence of this measure of real income is sensitive to changes in income distribution, but advises that the actual degree of sensitivity not be exaggerated.

Usher's first two definitions of income differ only in choice of base year, but his third definition, cast in terms of certainty-equivalent income, is not only substantively different but considerably more esoteric, although he maintains, and supports with a hypothetical illustration, that changes in it can, in principle, measure changes in real income (and, hence, in welfare). Here, too, he acknowledges that the measure is sensitive to changes in income distribution and also that his crucial assumption namely, that the elasticity of utility with respect to income is the same for all values of income, is not based upon actual evidence and that calculating this elasticity in the real world would be an onerous task. In another relatively esoteric vein, Usher's fourth definition conceives of real income in terms of how much it would cost, per person, in the base year to attain the value of the social welfare function attained in the current year. In order to "link" together individuals present in different years, he assumes that the welfare function is "anonymous" in the sense that its value is unchanged if incomes are interchanged between individuals, and if demographic changes occur over the interval, so long as average income is invariant. And, as always, the specific assumption with respect to income distribution, now that it not only remains unchanged but that there exists complete equality of income within each decile class of earners, plays an important role in his analysis.

Usher reiterates his underlying premises.

> ...We...[maintain]...the hard and fast distinction between the size of income and the distribution of income. We shall not...impose a social welfare function onto the spectrum of incomes or...take account...of differences in taste among consumers or of special characteristics of the rich or the poor....On an empirical plane, it has to be recognized that adequate time series of the distribution of money income are not available over long periods of time, that...we cannot in practice account for differences in utility functions among consumers, *and that allowances for changes in the distribution of income would complicate matters greatly when we come to consider real saving, imputations, quality change and other modifications to the circumstances in which real income is defined*....(p. 50, emphasis added)

The contrast between the immediately preceding words and the caveats specified by Hicks and Kuznets, and essentially endorsed by Little (to say nothing of the far more prohibitive ones set forth by Samuelson and Chipman and Moore), is striking. Whereas their caveats are intended to caution against facile intertemporal GDP comparisons, whether for a given country or a group of countries, they ultimately validate the comparative exercise, properly done. Their attitude may be summed-up thus: do it if you must, but be very careful and don't claim too much, for our caveats are restrictive, and don't forget to complement GDP with income distribution and other data. Also, keep the comparisons ordinal. Usher's message is both more positive and more negative with respect to GDP's suitability as a welfare measure. He carefully notes the analytical importance of trends of taste, income-distribution and demography, but what are restrictive caveats for Hicks, Kuznets and Little are for him adequately-realistic assumptions, whose limitations must be acknowledged but not permitted to preclude cardinal welfare comparisons on the basis of GDP. This is Usher's positive message: his negative message is also important, for it casts a long shadow. He not only recognizes the analytical complexity of allowing consumers to save, invest and do other things that people do in the real world, he explicitly recognizes the need to incorporate into welfare comparisons the imputed values of quite a few variables, ranging from home grown food, through depreciation (this appears to render the comparison in terms of NDP, not GDP) and environmental changes, all the way to travel to and from work and changes in mortality rates. This part of Usher's overall message points us in a very different direction from the one in which we are pointed by Hicks, Kuznets and Little namely, in the general direction chosen, *mutatis mutandis,* by the above-mentioned analysts that make cardinal welfare comparisons based upon derivatives of GDP. His imputations include an estimate of the value of the services provided by publicly-owned capital goods, and he explicitly refuses to do the same for those government services devoted to "defence and public administration." Thus, we now return to the subject of government services and, in due course, we will give special attention to these excluded services.

THE CATEGORIZATION OF GOVERNMENT SERVICES

The problem of how best to categorize government expenditures in measuring GDP has preoccupied students of national accounting from the very outset. Its complexity may be inferred from the fact that some leading analysts in the field have taken different positions at various times, in response to criticisms of their previous positions. Studenski (1958), who studies at length the debate surrounding the problem, goes so far as to declare it one of the most difficult in the entire field of national accounting.

THE OFFICIAL VIEW

GDP, viewed in terms of expenditures, includes final government consumption expenditures at purchasers' prices; and it is worth considering what the 1993 SNA says on the subject.

> The collective services provided by government consist mostly of the provision of security and defence, the maintenance of law and order, legislation and regulation, the maintenance of public health, the protection of the environment, research and development, etc. All members of the community benefit from such services. As the individual usage of collective services cannot be recorded, the individuals cannot be charged according to their usage or

the benefits they derive.... (p. 214)

This reasoning is extended to certain non-collective (i.e., quasi public) goods, such as health, education and services "concerned with the formulation and administration of government policy, the setting and enforcement of public standards, the regulation, licensing or supervision of producers, etc." The value of specific government services is defined to be equal to the amount of the corresponding expenditures. Also:

> Many government expenditures benefit enterprises as much as households; expenditures on the cleaning, maintenance and repair of public roads, bridges, tunnels, etc....These are individual services whose consumption can be monitored....*However, it would be difficult to separate the services provided free to households from those provided free to enterprises and, by convention, all these expenditures are treated as collective final expenditures*. (*ibid*., emphasis added)

There are only three possible positions on the appropriate categorization of government services in GDP: all are intermediate products; some are intermediate and some are final products; and all are final products. In spite of a certain amount of overlap, both of rationale and critique, it is useful to consider each position in turn.

ALL GOVERNMENT SERVICES ARE INTERMEDIATE PRODUCTS

Studenski

To illustrate this view (but without endorsing it), Studenski cites the rationale offered by economists that estimated Hungary's national income, during the late thirties, for deeming all government services, with very few exceptions, as being intermediate products. This rationale is repeated, *mutatis mutandis*, by later analysts that share their outlook in varying degrees.

> *...[T]he machinery required for keeping order and securing safety...do not produce values in addition to the flow of consumers' goods, but ensure only the maintenance of the present economic and social order and the maintenance of the present level of production*. (emphasis in text)....If the cost of public administration increases, the national income does not therefore become greater, just as it would not become smaller if the cost of public administration could be reduced, provided the prices of consumers' goods remain unchanged....We do not dispute the usefulness of the public service but....the results of its usefulness appear in the value of the goods and services produced, *and an inclusion of the*

> *cost of public services as such would mean double counting*.... (p. 196, emphasis added)

The underlying concept is unequivocal: no activity undertaken during a given fiscal period contributes to social welfare unless its fruits are directly and immediately manifested in additional consumption or investment goods. Man lives by and for bread alone.

The basic argument presented by Studenski, in rejecting this reasoning, is couched in terms that are likewise echoed by other analysts sharing his perspective in varying degrees.

> ...It [the concept] cannot reasonably be said to apply to modern government, which in a thousand ways seeks to provide necessities and conveniences to its citizens *complementary* to those that business enterprise provides or that they provide for themselves. *Can it be said that public health, hospital, and clinical services that combat disease, promote hygiene, and prolong human life "do not produce values in addition to the flow of consumers' goods," and that they do not contribute to a "general rise in the standard of living"? Or is it true of the services of public museums , parks, highways, street lighting, street cleaning? Is not the feeling of security in the enjoyment of one's life, provided by the various protective services of government, a good that is additional to that provided by the flow of consumers' goods?* (*ibid*., emphasis added)

Studenski also points out a serious inconsistency in the methodology advocated by the Hungarian economists. Goods and services purchased by the government from business are treated as final products and, therefore, are included in GDP, a treatment that is hardly appropriate if the resulting government services are themselves deemed to be intermediate products. As will be seen in due course, this is not the only such inconsistency in the argument that all government services are intermediate products.

SOME GOVERNMENT SERVICES ARE INTERMEDIATE AND SOME ARE FINAL PRODUCTS

<u>Mainly Kuznets</u>

Hicks (J.R. and U.K.) (1939) argue in favour of including all government services in GDP, treating them in effect as final products, even though they consider many of them intermediate products. Their grounds are the purely pragmatic ones that the measurement problems that attend any attempt to separate the one type of product from the other, with respect to each of the many services in question, are insurmountable. However, Hicks (1948) reconsiders the question, having in the interval further consulted his economist wife (U.K.) and become convinced that the measurement problems are not in fact insurmountable: so, intermediate government products should be identified, measured and excluded from GDP. Hicks is not the only analyst to reconsider his position. Colm (1954) records his own concurrent journey in the opposite direction, wherein he abandons his original position (namely, Hicks's revised one) and adopts Hicks's original (and now abandoned) position, on the very same grounds of intractability of measurement problems. It appears that tractability and intractability of measurement problems in this area are very much in the eye of the analyst (and his advisor, if any).

Unlike the shifting pragmatic grounds of these two writers, Kuznets's own shifting grounds are essentially conceptual. He frequently couches the issue in terms of intermediate and final products, but not always and not entirely. Consider Kuznets (1941):

> It could hardly be denied that the services of the post office, judiciary, etc. represent productive activities and contribute to the satisfaction of the needs of society at large. Doubts, however, have often been expressed concerning the validity of including the services or armed forces in national income. One can easily see the reason for such treatment if an investigator adopts criteria of productivity in the light of which he can modify judgments expressed by an overt act of the body social. However, only the acceptance of criteria of productivity different from those applied to society at large would justify the treatment. Since the estimators or analysts who advocate it usually profess to accept the dictates of the market place, the exclusion of services of government agencies such as police or armed forces...seems inconsistent....
>
> ...[I]f the activities of private police used by many large corporations are productive, why not those of municipal police? And if of domestic police, why not of international police, i.e., the armed forces of the nation? *If capital invested in private plants is productive, why not capital sunk in the preservation of the country's economic system or in securing to it economic privileges that affect the welfare of all enterprises or inhabitants? Indeed, there is considerable parallelism between governments and private corporations in their expenditures on policing, economic warfare, their financial structure*.... (p. 31n, emphasis added)

For Kuznets, on this occasion, both private and public expenditures reflect conscious and, above all, *complementary* social choices, the former collective, the latter individual, which are made in pursuit of utility in a manifestly imperfect world. Both kinds of outlay therefore warrant inclusion in GDP.

However, a decade later, in Kuznets (1951), his reasoning is quite different. GDP is now described as "an approximation to social welfare," and it is explained that some social goals are both society-specific and time-specific. If, therefore, comparisons are to be made over time and across societies, they must be made in terms of goals that "transcend differences in economic and social organization across time and space." Kuznets now explicitly echoes Pigou's view that only with the "measuring-rod of money" (albeit "stretched to its utmost") can the transcendental goal of economic activity namely, the satisfaction of the wants of the individual consumers that make up the society, be made operational. Hence, with respect to government services, only those services directly provided to consumers, plus those that increase society's capital stock, should enter into GDP. Kuznets is now thoroughly at odds, not only with his earlier outlook, but also with the official view of GDP, which maintains that unlike purchases of inputs by firms that will later be resold at a price, most government purchases of inputs are made on behalf of the population as a whole, not for resale but to provide it with free services. Hence, they should be included in GDP. For Kuznets, now, the not-for-resale criterion is adequate to identify final expenditures but inadequate to identify final *products*. Some consumers' purchases are made because they directly produce individuals' happiness, and therefore are final products. Others, however, are made for resale, in that they enable individuals to sell their labour and, therefore, are intermediate products.

> ...[M]ost government activities are designed to preserve and maintain the basic social framework and are thus a species of repair and maintenance *which cannot in and of itself produce net economic returns*. (p. 184, emphasis added)

But, and this decisively sets him apart from other analysts making analogous assertions, Kuznets immediately proceeds greatly to restrict the generality of this position by specifying a major government activity to which it does not apply (and he later specifies others). That activity is the waging of "a crucial war," a war that is a "life and death struggle" in which individuals' current welfare is less important than the "survival of the social framework." For the duration of this "temporary" situation all defence expenditures are final products.

More generally, Kuznets sets forth three criteria for distinguishing intermediate from final government products. The first criterion is that the individual recipient of the service pays no price for it, or only a token price (the service is financed by taxes or borrowing). Thus, not only do free education, health care, legal aid, etc. qualify as final products, but so do most other free government services. The second criterion is intended greatly to narrow the field: the service only becomes available to individuals after they have taken an overt initiative to obtain it. The three services just mentioned again qualify as final products, but the services of the legislature and the military (to mention only these) do not.

> ...The activities by government designed to preserve or expand the framework involve economic costs to society at large; but any net returns cannot be associated directly with any changes in the framework, certainly not in terms of services to individuals. (p. 193)

Kuznets, believes, however, that a criterion which excludes from the final-product category all government services not supplied at the overt behest of individual beneficiaries is insufficient to exclude all governmental intermediate products. He gives the example of the work of an appellate court, which although set in motion by overt individual initiative, belongs to the same category as other services designed to preserve the social order. Hence, a third criterion is needed, which requires, in addition to the other two criteria, that the particular service-to-individuals have an analogue in private markets.

> ...This permits the inclusion of such services by government as education, which obviously finds its analogue in purchases of private education; medical services, with similar analogues in private medical service; parks, theatres, public tourist centers, amusements, etc.... (p. 195)

Kuznets admits, however, that, if stretched too far, this criterion breaks down. For example, the fact that individuals hire bodyguards must not be allowed to justify declaring police services as final products. The analogue in private markets must be "widespread," although Kuznets undercuts this position by admitting that the presence of a widespread analogue may imply that the social framework "does not recognize an overriding need for internal peace." In such circumstances, police service to individuals qualify as final products. He also acknowledges another problem with his private-market-analogue criterion: government services that obviously are services to individuals may be so well delivered that pre-existing private suppliers vacate the field. Examples include free medical care, free right to participate in elections, and free police protection to individuals.

It is important to note that Kuznets also believes that all government spending on capital goods, like all private investment spending, should be included in GDP; and he includes changes in the "net balance" of claims on foreign countries under that heading, although he acknowledges a difficulty with claims resulting from war. Even when the services provided by governmental capital goods consist in intermediate goods investment in them should be included in GDP.

> ...There is no inconsistency in including in the final product of government changes in the stock of armament, and yet excluding from final product such government activities as are carried on by the country's armed services; as there is no inconsistency in including additions to the stock of blast furnaces in net output, and yet excluding pig iron from the flow of finished goods to the country's ultimate consumers. (pp. 197-198)

In sum, expenditures to feed, pay and otherwise maintain the police and [peacetime] armed forces as well as the corresponding bureaucracies, as well as expenditures to hire and manage road maintenance crews, etc., are expenditures for intermediate products, but expenditures to equip them with the tools of their respective trades, in order to enable them to do their job are expenditures for final products.

Abandoning previous positions, Kuznets now implicitly agrees (subject, again, to the crucial reservation that his agreement only applies to peacetime conditions) that individuals live by and for bread alone. Being consumers and nothing more, they view all other material and institutional aspects of their existence as but means to that end, not as ends in themselves. If, under these conditions, they want liberty and protection of life and property at home, protection against foreign enemies, reasonably honest and efficient judicial institutions, and the host of institutions, rights and regulations that preserve individual security and liberty, as they obviously do, it is not because these things are valuable in themselves but merely because they serve present and future consumption appetites.

Like most other participants in the debate, Kuznets provides no factual evidence to substantiate his new (or his old, for that matter) outlook on human priorities, whether individual or collective. Nor does he provide practical criteria to facilitate making his crucial distinctions operational. Although he is at pains to exempt expenditures to wage a "major war" or maintain a "chronic state of crucial struggles" from his restrictions on the final utility of national defence, he provides no criteria for distinguishing major from minor wars, chronic from nonchronic states of struggle, or crucial from noncrucial struggles. If specified, such criteria would surely qualify as final products the military expenditures made during (and at least shortly before) the two world wars and the Cold War of this century. Whether they would similarly qualify the expenditures made during the probably-less crucial Korean War and Vietnam wars is less certain. On balance, it is fair to say that Kuznets's "peacetime" allocation of defence expenditures to the category of intermediate products must await the advent of a future epoch, one that is very different from the contemporary one.

It must also be noted that Kuznets's list of government services that qualify for inclusion in GDP is significantly incomplete. It does not include services like street lighting and garbage removal or the services of roads, bridges, canals, etc., all of them services that meet his three criteria. Individuals must take overt action to use them, little or no price is charged, and the private sector that once provided them much more extensively than it does currently has withdrawn as a result of the successful entry of government into the field.

Nor, in an important omission, does Kuznets consider the possibility that there exists a private sector analogue to the security services that he designates as intermediate products on the grounds that do not directly produce consumer goods: they only create the environment in which consumer goods can be produced and consumed. The overhead and other fixed costs that firms necessarily incur in their everyday operations are incurred for reasons that are entirely analogous to the government's reasons for incurring security costs: to preserve the firms as institutional environments in which production can take place. Indispensable though these "environmental" costs are, they do not in themselves directly result in output. They do, however, enter into the market prices of the products that constitute that output and, thereby, into GDP, and, quite properly, no one finds this practice objectionable. This is an anomaly that will be returned to below. All things considered, Kuznets is a lone voice in this debate, a voice that belongs neither in the analytical camp of the above-mentioned Hungarian analysts nor in that of the analysts discussed next.

Nordhaus and Tobin

As, *mutatis mutandis*, do other derivatives of GDP (including the above-mentioned GPI), the well-known "measure of economic welfare" (MEW) developed by Nordhaus and Tobin (1972, henceforth N-T), excludes a variety of private and government expenditures (of which the latter are much more quantitatively important) on the grounds that they are "instrumental," or "regrettable" expenditures. The excluded private instrumental expenditures are the sum of the items in one consumption-expenditures category of the US national accounts, personal business, and one-fifth of the items in another, personal transportation expenses. The first excluded amount refers to a variety of personal services ranging from brokerage charges and investment counselling, through bank service charges, other services provided by financial intermediaries, life insurance handling and legal services, all the way to funeral expenses. The second is an estimate of the cost of commuting to work.

The quantitatively more important government instrumental expenditures not considered contributions to MEW include police services, sanitation services, road maintenance and national defence: N-T describe them as "among the necessary overhead costs of a complex industrial nation-state...[for]...activities that are not directly sources of utility themselves but are a regrettably necessity to activities that may yield utility."

N-T's discussion of instrumental expenditures is largely couched in terms of defence expenditures because, in their view, they most fully illuminate the issues raised by the genre. Their case for excluding them begins with two specific reasons.

> ...First, we see no direct effect of defense expenditures on household economic welfare. No reasonable country (or household) buys "national defense" for its own sake. *If there were no war or risk of war, there would be no need for defense expenditures and no one would be the worse without them.* Conceptually, then, defense expenditures are gross but not net output. (p.7, emphasis added)

Their second reason is that defense expenditures are input rather than output data. Further, measurable output is especially elusive in the case of national (or other) security.

> ...Has the value of the nation's [US] security risen from $0.5 to $50 billion over the period 1929 to 1965? Obviously not. It is patently more reasonable to assume that the rise in expenditure was due to deterioration in international relations and to changes in military

> technology. *The cost of providing a given level of security has risen enormously. If there has been no corresponding gain in security since 1929, the defense cost series is a very misleading indicator of improvement in welfare*. (p. 8, emphasis added)

Implicitly rejecting Kuznets's distinctions, N-T acknowledge that when prompted by "unfavourable circumstances," defence expenditures can make consumers better off, but they do not merit inclusion in MEW because their benefits are only indirect.

> ...*The diversion of productive capacity to [defence production] cannot be regarded simply as a shift of national preferences and the product mix*. Just as we count technological progress, managerial innovation, and environmental change when they work in our favour (consider new business machines or mineral discoveries) so we must count a deterioration in the environment when it works against us (consider bad weather and war). From the point of view of economic welfare, an arms control agreement which would free resources and raise consumption by 10 per cent would be just as significant as new industrial processes yielding the same gains. (p. 8, emphasis added)

It is worth mentioning, in passing, that the slightly earlier Lekachman (1971) reflects, in his own terms, much the same reasoning with respect to regrettable expenditures. A self-proclaimed "deteriorationist at heart," he advocates the development of an "index of amenity," whose purpose would be to deflate GDP. By his own admission, this index is subject to "severe flaws in logical consistency," but it would still be better than nothing. The main issue is the extent to which purchases made by consumers can be regarded as "personal enjoyment," rather than as "costs of personal enjoyment," a distinction he regards as crucial (he eschews the term "instrumental" expenditures).

> ...[Much of what silent majoritarians purchase is perceived *by them* (not just by...snobbish elites) as a grudging response to personal danger (as in the acquisition of fierce dogs, burglar alarms, locks, and firearms), environmental hazard (as in the purchase of air conditioning and home insulation against intolerable noise), or a direct defense (sic) of one's personal health (emphasis in text)... *Such expenditures are quite likely to strike the men and women who make them as coerced. They are among the costs of personal enjoyment rather than the enjoyment itself.*
>
> It is a deficiency of current national income calculation that the general rubric of market activity is made to cover both the pluses of personal gratification and its costs....
> (p. 123, emphasis added)

Returning to N-T, although they largely disregard them, they paradoxically register some important caveats to their own approach.

> We must admit...that the line between final and instrumental outlays is very hard to draw. For example, the philosophical problems raised by the malleability of consumer wants are too deep to be resolved by economic accounting. Consumers are susceptible to influence by the examples and tastes of other consumers and by the sales efforts of producers. *Maybe all our wants are just regrettable necessities, maybe productive activity does no better than to satisfy the wants which it generates; maybe our net welfare product is tautologically zero. More seriously, we cannot measure welfare exclusively by the quantitative flows of goods and services. These, too, will be relative to the value systems which determine whether symptoms indicate health or disease*. (pp. 8-9, emphasis added)

There are two problems with N-T's first reason for excluding defence expenditures from MEW, the first of which can readily be seen by considering national defence from a historical perspective. It is the merest of truisms to assert, as they do, that without war and risk of war there is no rational justification for defence expenditures when, ever since the dawn of recorded history, the world has never been free of either. This sad but incontrovertible fact is all the more apparent if risk of war is defined as risk of armed conflict with another country that, if not imminent, is significant over, say, the medium term. In a word, N-T's first reason for excluding defence-related expenditures from their welfare measure depends upon a state of human affairs that has yet to exist. The second problem with this reason is as or more serious. National security not only provides the welfare inherent in the protection of the economic system, it also simultaneously provides the welfare inherent in additional social desiderata, some perhaps even more cherished than a preserved economy, such as peace and liberty. Wedded to undefined "economic welfare," N-T's do not reckon with this consideration. Since their approach is shared by other analysts, this very important issue will further be taken up towards the end of the paper.

N-T's second reason for excluding defence expenditures from MEW is also unpersuasive, as can be seen by generalizing it to other expenditure areas , as does Jaszi (1958) while also offering what is at once an example of sound national accounting and anomalous welfare measurement.

> ... [S]uppose defense production increases but the power of the potential enemy rises *pari passu*. Other things being equal, the population is no better off than before, but national product measured inclusive of defense expenditures would be larger. *This is the sort of ground on which it is argued that the inclusion of defense expenditures in national output constitutes duplication*. By analogy, in the private sphere there may occur a rise in medical

> expenditures which merely offsets an increased danger of infectious disease. Other things remaining the same, the population is no better off than before, but again the national product as currently defined would rise. (p. 71, emphasis added)

Jaszi here touches on a matter the importance of which extends far beyond defence and health expenditures. Expenditures captured by national accounting only reflect the fact that they were considered necessary by the individuals that made them, a fact that tells us nothing of their circumstances or consequences. Since these can and do vary widely, without knowledge of them there is no way of distinguishing expenditures that enhance welfare from those that merely restore it by remedying a previous loss or, indeed, from those that reduce it because they are perverse.

Although they are by no means the only arguments that can be brought to bear, Kuznets's above-mentioned arguments suffice to show that N-T's general approach to national defence is unsatisfactory. Their discussion is also sometimes mistaken in purely factual terms, as for example when they in effect deny that American defence expenditures had a major impact on the welfare of the United States (to speak only of that country) during this century. To give but two of several possible examples to the contrary, the historical evidence is overwhelming that the insignificant American defence expenditures of the early-thirties strongly encouraged the aggressive designs of the governments of Germany, Japan and Italy that made World War II inevitable. Further, the year 1965 was on the morrow of the Cuban Missile Crisis, an occasion when overtly manifested American military strength prevented a very dangerous shift in the international balance of power that would greatly have increased the risk of all-out war between the Western democracies and the Soviet Bloc.

Finally, N-T behave inconsistently in failing to remove from MEW a variety of private-sector analogues of governmental defence expenditures, including expenditures by households and businesses for security, whether in the tangible form of locks and alarms systems, in the human form of security guards, and the like, or in the intangible form of insurance coverage against theft and vandalism. As does Kuznets, they also fail to consider whether there is an analogy between the public costs of protecting and preserving the environment in which production takes place and the private fixed costs routinely, and inevitably, incurred by firms: although they likewise do not directly result in output, they accomplish the same end of making output possible. Granted, the private costs are (and will always be) technically obligatory and the public ones are not; but, the world being what it is, they are what may be termed “socially obligatory.” Until that devoutly-to-be-hoped-for dawn of a better world, the word “obligatory” will remain operative. The fact that the private costs enter into the prices of consumer and investment goods bought and sold in private markets, and that the public costs may or may not (depending upon how they are financed by the government), should not be the determining factor with respect to inclusion in a welfare measure.

The full relationship between defence expenditures and national security is exceedingly complex and far beyond our scope. The rules of thumb sketched by Kuznets certainly have their uses but the most important consideration was just implied: the welfare implications of a given level of defence expenditures depend crucially upon the attitudes and responses of other countries, especially of actual or potential enemies. Without information on these attitudes and responses, the implications are indeterminate: they can be positive, negative or nonexistent (the last is unlikely). The only thing that can be asserted with confidence is that it is inappropriate to exclude defence expenditures *a priori* from a welfare measure, whether welfare is defined broadly or narrowly. This, of course, also implies that the expenditure data must be accompanied by additional information, and that in turn raises additional issues. Several of these will be taken up in various places below, and particularly in Appendix III.

ALL GOVERNMENT EXPENDITURES ARE FINAL PRODUCTS

Jaszi

Proceeding from a very different perspective from that of analysts he describes as "welfare-oriented," Jaszi (1946) contends that semantic confusion is an important contributing factor to the controversy surrounding the proper place of government expenditures in GDP, a controversy that was already heated in his day. Citing the unenviable experience of various countries, he denies that a useful distinction can be drawn between intermediate and final products in the case of government services. He insists that even in the private sector, distinguishing intermediate from final products is far from being a straightforward exercise. Offering an example involving the production of men's suits, he shows that the composition of national product is affected (rightly so under established national accounting conventions) by whether suits are made from cloth customers have bought and supplied to their tailors or whether the tailors buy their own cloth. Consequently:

> ...[I]t is not possible from the physical characteristics of yarn to determine whether it will be enumerated as a final product or not. *The criterion which is used to distinguish between final output and intermediate output is whether the item in question is bought by a consumer or bought by a business. A final product is output bought that is bought by consumers. An intermediate product is output that is bought by a business [and either charged to current cost or to fixed assets]. A definition of consumer goods, which runs in terms of physical characteristics that yield satisfaction to consumers, is not operative.* (p. 171, emphasis added)

However,

> ...government services are neither bought by consumers nor by business and hence the definitions of consumer good and intermediate product used in the measurement of national product in the private sphere are not directly applicable. (p 172)

Jaszi points out that economic theory postulates that consumers, at any given time, are faced with a specific variety of goods and their respective prices, from among which they make choices (subject to their budget constraints) to maximize their total utility; and he insists that the operative term here is "specific variety," a term that refers to a particular institutional setting. Staying with the example of cloth and mens' suits, there invariably exist several different, and alternative, institutional settings in which cloth could be bought.

> ...For instance, if the customary procedure is to buy uncut cloth and to have it tailored to measure, uncut cloth is a consumer good, *because the customer would take it into account in the course of his exchange transactions*. If, on the other hand, it is the practice to wear ready-made suits, uncut cloth will disappear from the list of consumer goods, *because it is no longer an item in the list of exchange transactions by which consumers maximize their transactions*. (p. 174, emphasis added)

Economic theory does not, and need not, differentiate between institutional settings in order to establish that consumers always seek to maximize utility. Consequently, "[economic] theory does not provide tools which would allow national product measurement to transcend the market place."

All this is argued in support of Jaszi's main thesis that, more than anything else, GDP is an accounting concept far more suitable for measuring output than welfare. (It will be seen in due course, that this view is overdrawn.) Pointing out, in passing, that GDP ignores the contributions of family services, he offers the example of shelter, heat and light provided to workers in factories, and suggests that these commodities enhance workers' welfare at work in precisely the same way as they enhance it at home. "Yet the latter are, and the former are not, part of [GDP]." He also provides additional examples. Recognizing that inputs used by business enter into the prices of products, he insists that:

> ...It will not do to say that the heat and light consumed in places of work is [sic] indirectly included in [GDP] by being an element of the price of the products sold by the factory. This can be made clear as follows. If a simple shift should occur between heat and light consumption places of work and heat and light consumption in homes, *deflated* [GDP] would show a change (an increase if the shift were towards the home, and a decrease if the shift were away from it) *in spite of the fact that no change had occurred in the total physical consumption of heat, light, and factory products*. (pp. 176-177, emphasis added)

After reviewing some conundrums presented by certain imputations (ranging from food and board of farmworkers and members of the military, through owner-occupied premises, to self-administered shaves and haircuts) Jaszi insists that as long as analysts are under the "hypnosis" of a welfare-based definition of GDP, the problems of imputed income and occupational expense will be a constant source of irritation. Glaring omissions that cannot possibly be reconciled with a welfare definition will always be encountered.

> ...To give only one striking example, the services rendered to consumers by the resources engaged in the...broadcasting industry do not appear in any of the existing measures of [GDP] even though that industry is just as important in providing recreational enjoyment as are motion picture theatres, etc. (pp. 103-104)

Jaszi goes on to review some attempts to establish a satisfactory criterion for distinguishing intermediate from final products, and he finds them all wanting.

> ...The reason for this impasse is that the proposed criterion involves the asking of a basic philosophical question which never is raised in economic theory, viz. what things are wanted "for their own sake", "are valuable in themselves", and what things are desired merely

> because they are instrumental in getting other things. *Economic theory merely assumes that individuals will buy the collection of goods which maximizes their satisfaction. But it does not analyze the internal production process, as it were, by which consumers maximize their satisfaction; it does not attempt to distinguish* between...raw materials in that process and goods that are "valuable in themselves." Introspection will show that this question often cannot be answered.... (pp. 187-188, emphasis added)

The insoluble difficulties in the private sphere also exist in the public sphere, Jaszi maintains, offering the illustration of public education. Some people think it valuable for its own sake; others think it valuable for increasing earning power; and still others think it valuable for both these reasons. So, in the event that the government ceases to provide free education, and assuming that the consumption of marketed consumer goods is initially unaffected, it may be inferred that some, but only some, proportion of total public education is an intermediate good at that point in time. This raises a serious problem because:

> ...different individuals will feel differently about this matter, that the same individuals will feel differently about it at different times, and that most individuals...will not have a clearly defined attitude at all. *No objective measure can be constructed from this evidence.* Other examples of a similar nature would arise in cases when government provided free transportation to individuals. (p. 191, emphasis added)

Jaszi agrees that, as a practical matter and unlike firms, the government often cannot distinguish individual users from business users of a given service. (Nor should it do so, in his view, even if it could.) He sees no analogy between the supplying of government services and the allocation of certain costs, such as electricity, in a firm with two or more production divisions, each division producing a different and unrelated product from specific inputs that are also independent of the inputs of the other products. For the firm with this kind of production to allocate its electricity costs among its products is rational, but it would not be rational, in his view, if, instead of being different and unrelated, the firm's products consisted of joint products or a primary product and various byproducts. Such products are a set of two or more related products that emerge from the same inputs, and the output of the entire set increases when any of these inputs is increased. There is no rational reason to allocate the cost of any single input among such products, nor should overhead costs be allocated among them. Whilst some government services can validly be likened to products to which input costs can be specifically allocated, most government services are analogous to joint products or primary and byproducts.

Like other analysts, Jaszi offers the example of national defence but, unlike many of them, as an input that serves more than one important end. Two of these are the preservation of liberty and the provision of the peaceful conditions essential to a stable and efficient economy. How then are the costs of national defence to be allocated between individuals and business? To ask this question, Jaszi argues, is to demonstrate its vacuity.

> ...[It does] not seem more reasonable than the question whether sunshine is desired for its own sake or because it facilitates production; and, if it is desired for both purposes, what is the proper allocation between business and consumers? Yet, in many studies, assignments of basic government functions either to business or to consumers, or allocations between the two, have been advocated or actually made. (p. 194)

Jaszi's broaching of the subject of cost-accounting for joint products and primary and by products is interesting and will be considered more fully in the paper's concluding section.

Jaszi also tackles the claim that including all government services in GDP involves double counting (since many of these services are allegedly intermediate goods) first by pointing out that national accounting with respect to the private sphere is itself hardly devoid of double counting.

His most basic rebuttal of the double-counting charge, however, runs along different lines. Analysts that decry government services as being intermediate products tend to conceive of a welfare measure as an stand-alone aggregate, calculated at a certain date, say a fiscal year end. This, in his view, is inappropriate, for the very good reason, allegedly recognized in economic theory, that an *absolute* welfare measure is conceptually impossible: only a *relative* measure is possible. We cannot speak, in other words, of this year's absolute level of welfare; we can only compare this year's welfare measure, say GDP, with last year's, and then say whether total welfare has risen or fallen during the year (presumably ordinally). From this perspective, the answer to the question of whether the inclusion in GDP of both cloth and the suits made from it distorts welfare measurement depends upon whether a GDP that only includes suits is the best basis for measuring relative welfare. Put another way, a GDP that includes both cloth and suits is certainly a GDP that includes an intermediate product, but this does not in itself suffice, according to Jaszi, to render it a flawed basis for measuring *relative* welfare, the only welfare that can validly be measured.

To further illustrate the complexity of the problem of satisfactorily delineating government services that are intermediate products, the inclusion of which in GDP distorts it as a welfare measure, Jaszi conjures up eleven different production scenarios involving cloth and suits, with various forms of government participation financed by various tax regimes. These are summarized in Appendix I.

Gilbert, Jaszi, Denison and Schwartz

Some of Jaszi's arguments are further developed by Gilbert *et al* (1948), which is specifically intended to respond to Kuznets (1948). Denying that a precise line between intermediate and final goods can ever be drawn, these American experts in national accounting suggest that most consumers' expenditures are not for "Sunday clothes and holiday dinners," which are obvious examples of "the good life" and, therefore, of final products. These other expenditures could as legitimately be viewed as having been made to maintain consumers' earning power as having been made for final consumption. Since much the same is true, in their view, of many business expenditures, conventions are needed in order to get on with calculating GDP, any version of which:

> ...must overlook the fact that the expenditures of individuals in their business capacity are influenced by their standards as consumers, and that the expenditures of consumers are influenced by their activities as producers. It must overlook also the fact that the conditions under which work is performed have an important bearing on the welfare on individuals. *These conditions are affected by business expenditures on goods and services that are classified as intermediate just because there is no satisfactory way to take account of their benefits in a quantitative measure of final output*. (p. 184, emphasis added)

As does Jaszi, Gilbert *et al.* dispute Kuznets's (later abandoned) claim that the classification of intermediate and final products can readily be extended to the sphere of government services. They deny, for example, that his "all types of economic legislation, administration, and adjudication" really are services to business. They regard them as being, instead, services to individuals.

> ...Activities such as regulation of security dealings, administration of anti-trust laws, control of child labor, provision of minimum wages, and regulation of public utilities are certainly services to the individuals that make up the nation. *Economic legislation generally provides rules for the operation of the private enterprise system in order that it may function with greater protection to the individual against fraud and the hazards of industrial life*.... (p. 185, emphasis added)

Gilbert *et al*. also conjure up hypothetical examples to make their point. In one, the government supplies free flour to bakers, and they ask whether it would be double counting to include the government's purchases of flour in GDP. Since the flour is not a cost element to bakers it does not enter into the market values of their final products: hence, the question must be answered negatively. Nor would there be double counting if, instead, the flour were distributed free to consumers and the government's flour purchases were included in GDP. They point out that if these flour purchases are not included in GDP, as some analysts propose, comparing two consecutive years, in one of which the flour is purchased by bakers and in the other it is obtained free from the government, would result in distortion.

They also reject Kuznets's (later modified) argument that many government services (e.g., to provide internal and external security) are not direct contributions to welfare because their purpose is to preserve the state of society in the context of the existing civilization. They point out that the costs of these myriad services are highly variable, because the circumstances that render them essential in the eyes of the government acting on behalf of society, are themselves highly variable. Succinctly put, different social needs demand different policy responses.

> ...For example, to decide whether armaments should be included in a measure of economic welfare it is not enough to consider the comparison of a peacetime situation with a wartime situation. One must also consider two situations in which there is a comparable need for armaments but a different quantity of armament production. *Under this circumstance, there is no question but that a larger war output means a larger measure of economic welfare. Omitting armaments from [GDP] would yield entirely false results for this type of comparison.* (p. 189, emphasis added)

The reality that social needs tend to change over time is of great analytical significance, in their view (as well as in the respective views of Hicks, Kuznets and others, as noted above): among other consequences, it precludes the making of intertermporal welfare comparisons. Consequently, although the costs of government services belong in GDP, any notion of using the temporal trend of GDP to measure the corresponding trend of society's level of welfare must also be abandoned. To express this strong conclusion in the conceptual terms of Hicks, Kuznets and the others, the analytical requirement that tastes and wants remain constant over the interval (one of their *sines qua non* of inferring changes in welfare from changes in GDP) is effectively incapable of being satisfied over an interval of any length.

Granted, Jaszi and Gilbert *et al.* have a profound knowledge of the practical exigencies and tradeoffs that routinely confront practitioners of the art of national accounting in situations in which no fully satisfactory alternative exists. Nonetheless, some questions arise when considering their resolute defence of the *status quo.* One of these concerns the *net* magnitude of the accumulated distortions, in any given fiscal year, that result from the reporting inconsistencies they attribute to varying institutional arrangements in the production and delivery of consumer goods. Like the analysts they criticize, neither Jaszi or Gilbert *et. al.* provide any empirical evidence to support their position. If the accumulated distortions prove significant, it could be argued that the solution lies in correcting the accounts, not in perpetuating inherently dubious practices with respect to government services. On balance, however, their arguments add verisimilitude and weight to the case that man does not live by and for bread alone. Hence, government services that to a substantial degree serve purposes that have their own utility, a utility separate from and independent of that of consumer and investment goods, warrant their own place in a welfare measure, whether it be a measure of economic or of general and comprehensive welfare. Their arguments also serve collectively to give pause to analysts that are too zealous in pursuing the otherwise laudable objective of eliminating double counting of intermediate and final products from their respective GDP-derived welfare measures.

PART III

FURTHER ARGUMENTS AGAINST CONVERTING GDP INTO A WELFARE MEASURE

Okun

The pragmatic arguments just considered avowedly do not conceive of GDP as a welfare measure. Instead, they conceive it as an aggregate economic statistic generated by a national accounting system that records multifarious private economic activities, the most important feature of which is that they occur within the institutional structure of markets, along with most government expenditures on goods and services. Somewhat paradoxically, the enormous reliance upon GDP by a wide variety of analysts in business, government and academia has not resulted in correspondingly wide acceptance of this conception. Instead, according to Okun (1971), it often attracts a "fundamental criticism."

> ...Put simply...the criticism is that, even after correction for price and population change, [GDP] does not yield an unambiguous measure of national welfare; a rise in real [GDP] per capita does not necessarily mean that the Nation has become better off, nor does a decline imply that it has become worse off. (p. 129)

This criticism strikes Okun as incomprehensible.

> ...Obviously, any number of things would make the nation better off without raising its real [GDP] as measured today: we might start the list with peace, equality of opportunity, the elimination of injustice and violence, greater brotherhood among Americans of different racial and ethnic backgrounds, better understanding between parents and children and between husbands and wives, and we could go on endlessly. *To suggest that [GDP] could become the indicator of social welfare is to imply that an appropriate price tag could be put on changes in all these social factors from one year to the next*.... It is asking...the national income statistician to play the role of philosopher-king...and it is absurd to suggest that, if the national income statistician can't do that job, the figure he writes for [GDP] is not interesting. (pp. 129-130, emphasis added)

Okun tackles head on the issue of "regrettable necessities," noting, as was discussed above in some detail, that it has been argued that these are not really final products and should, therefore, be excluded from GDP.

> ...Don't start down that path. If you should do so, regrettable and unnecessary as it would be, you would find it winds along forever. Physicians' services and all other medical care costs are obvious regrettable necessities. So are the services of lawyers, policemen, firemen, sanitation workers, and economists (including national income statisticians). So are heating and air-conditioning outlays. Except for the few people who live to eat, rather than eat to live food is a regrettable necessity. *Indeed, it is hard to imagine any output which clearly serves the purpose of pure, unmitigated enjoyment*.... (p. 132, emphasis added)

Echoing Denison (1971), Okun now advances a rather profound consideration, which was alluded to earlier in the context of national defence and war risks.

> ...But even if you could invent some arbitrary definition that kept final-product consumption from falling to zero, the exclusion of regrettables would make no sense. *It would deny the distinction between meeting one's necessities or failing to meet them. (ibid.,* emphasis added)

Using the example of air conditioning, he argues that if it is a regrettable necessity to those who have spent money to obtain it, it is also a regrettable necessity to those who do not have it. Furthermore:

> ...it must have been a regrettable necessity fifty years ago when nobody had it. Excluding the services of such items as might be deemed regrettable necessities is palpably unsatisfactory. (p. 132)

In other words, to exclude regrettable expenditures from a welfare measure is to erase the difference between meeting social needs and not meeting them.

On the subject of externalities (e.g, environmental degradation), Okun notes that GDP has been criticized for its failure to capture them when they occur. He wonders, however, how this could be done "if the market and the democratic process don't generate the necessary price tags."

> ...[GDP's present rules] reflect the costs and benefits that society recognizes and responds to. If a ban is placed on activity that is inherently dangerous, or fees and taxes are imposed, [national accountants] will follow the signals and properly reflect them in [their] valuation of output. If society changes its mind. [they] will make...changes in [their] definition and coverage of outputs. *But any puzzles that arise concern the volatility of the Nation's collective judgment, not of [their] practices*.... (p. 132-133, emphasis added)

He substantiates this point by referring to the national-accounting practice of excluding the output of illegal activity from GDP.

> ...[This practice] abides by the social judgment that some activities have such important negative externalities that they subtract from society's output even though somebody is willing to pay for them as an ultimate consumer. *However sensible or foolish it was for the Nation to decide that the sale of alcoholic beverages was illegal and then it was legal again, it was completely sensible for the national income accountant to follow these verdicts....* (p. 133, emphasis added)

These words are pertinent, but they are not the whole story. The wisdom or lack thereof of declaring certain expenditures legal or illegal has a bearing on the level of social welfare, as does the wisdom or lack thereof of the expenditures themselves. These matters will further be considered below.

It is precisely because Okun is so firmly convinced that GDP, however modified, is inherently incapable of serving as a satisfactory welfare measure, rather than because he fears that a modified GDP will deprive analysts of very useful information (which would still be available), that he so strongly opposes the very attempt to modify it.

> ...[A]s I see it, the big danger is that, by taking a few steps in the direction of an allegedly more comprehensive measure of welfare, a reformulation of the accounts might mislead the Nation into supposing that [GDP] was at last measuring social welfare. *And that would impede the progress which we so urgently need toward better measurement and evaluation of various changes in our social and physical environment, our health, and the diffusion of well-being across the country....(ibid.,* emphasis added)

Far from being something to deplore, GDP's unsuitability as a welfare measure is actually a matter of gratification to Okun.

> In short, [GDP] is not the whole story of our society or even of our economy, *and no conceivable redefinition can turn it into the whole story.* National accountants can help put together some of the other pieces required...But [they] would not assist by compromising on the proposition that [GDP] is *not* a measure of total social welfare. (Emphasis in text.) *The beauty of [the present] practice is that no sensible person could seriously mistake [GDP] for such a measure.* (*ibid.,* emphasis added)

<u>Denison</u>

Denison (1971) shares Okun's opposition to the undeniably attractive and convenient concept of a single, generally accepted welfare index. If such an index could be constructed:

> ...A glance at it would tell us how much better or worse off we had become each year, and

> each decade. We could judge the desirability of any proposed action by asking whether it would raise or lower this index. (p. 13)

To demonstrate the impossibility of constructing such an index Denison lists some of the social changes it would have to capture and the problems its compilers would face. He agrees that relative prices are useful gauges of consumers' relative preferences, and that they provide a sound basis for aggregating innumerable and disparate goods and services. He also endorses the value of measuring GDP at constant prices in order to shed light on changes in productivity. He cautions, however, against overburdening these positive features of GDP in vain attempts at welfare measurement, and in so doing propounds a very important theme that will be taken up below.

> ...[H]ouseholds, governments, and nonprofit organizations are regarded as the final users of the economy's output, and [GDP] measures the goods and services they buy. *How effectively they use their purchases is outside the purview of [GDP].* Soap, vacuum cleaners, washing machines, and the time of domestic servants bought by the housewife are measured, not how clean her house and linen may be. Similarly, the teachers' services, books, school buildings, etc., purchased by school systems are measured, as are the planes ammunition, and soldiers' services bought by the Department of Defense; *[GDP] does not tell us how much education and national security are obtained per dollar...of expenditures for such items.* (pp. 13-14, emphasis added)

Denison identifies some of the intractable problems that inevitably arise in any attempt to convert GDP into a welfare measure. One is the futility of any attempt to value government services by regarding governments as being, in effect, as he puts it, in the business of "selling" their services to the public. There is no acceptable way, in his view, of measuring "prices" of services such as education and defence, to speak only of these services. He is equally pessimistic about the feasibility of soundly, and generally-acceptably, measuring unpaid household services. Accounting for quality change is a third problem area , largely, though not entirely, due to the absence of an acceptable way of ascribing value to the ever-increasing range of new products that the economy generates: at present, no value is ascribed to this important and never-ending phenomenon.

> Such [problems], which in my view are not remediable, limit the accuracy of real product as a measure of changes over time in the ability of output to satisfy our wants. Nevertheless,

> real product is a very useful measure. *But to evaluate welfare we would need additional measures which would be far more difficult to construct.* (p. 14, emphasis added)

Denison identifies some of the many such measures, beginning with an index of real costs of production, which is necessary "because we are better off if we get the same output at less cost." Constructing it for, say, labour would require the "unreasonable" assumption that all changes, say increases, in required working hours per week cause equal burdens on all workers, irrespective of the numbers of hours they are already working. "A usable measure of the real costs of working would consider that the welfare benefits from working fewer hours decline as hours are shortened and may even disappear." Another dimension that would need to be captured is what is generally described as "working conditions," (the importance of which was recognized in a different context by Little, as noted above). Denison not only believes that the necessary data are lacking, but also that such highly complex variables as "the disutility of involuntary service in the armed forces" would need to be counted, along with the deaths and injuries suffered in the armed forces during both war and peace. Data on saving exist, but there is no measure of the real costs of what was once called "abstinence." Nor, in his view, does an acceptable way exist to combine the real costs of labour with those of abstinence.

To meet Denison's version of the familiar criterion of Hicks, Kuznets and the others, welfare measurement requires, in effect, measuring changes in the wants that goods and services must satisfy. This, he makes clear, is another very tall order. Consider the impact of population change, "now handled crudely" in per capita terms.

> ...[The assumption is] that, other things being equal, twice as many people need twice as many goods and services to be equally well off. Beyond this, an index of needs would account for differences in the requirements for living as the population becomes more urbanized or suburbanized; for the effect of weather changes on requirements for heat, air conditioning, and clothing; for medical requirements occasioned by epidemics or new diseases; and, most of all, for changes in national defense requirements. *Such an index would have to tell us the difference between the cost of meeting our needs to the extent that we do, in a base year, and the cost of meeting them equally well under the circumstances prevailing in every other year.* (*ibid*., emphasis added)

Repeating one of the most serious of the reservations summarized earlier, Denison, like Okun, gives short shrift to arguments in favour of excluding instrumental or regrettable variables from GDP.

> ...This procedure fails utterly. It yields the false result that we are equally well off whether, in the same circumstances, we ride or walk to work, freeze or are comfortable, do or do not obtain medical care when we are sick, or provide or do not provide for national security. *Needs and provision to meet them must be separately evaluated.* (p. 15, emphasis added.)

Put another way, the question is: if the expenditures made to satisfy certain wants is deemed regrettable, how much more regrettable should be the refusal to make them, and thus leave the wants unsatisfied, be deemed? As implied above, this question goes to the heart of the question of the feasibility of *any single* welfare measure.

Denison recognizes that measures of needs necessarily:

> ...shade into measures of the human and physical environment in which we live; perhaps it is here that the concept of economic welfare broadens to encompass "social welfare." (*ibid.*)

However, he strongly discourages optimism about the feasibility of incorporating environmental data and data on social pathologies into single a welfare measure. The data problems are enormously difficult, but they are less serious than two other problems, the first of which is the fact that technical relations between environmental conditions and welfare are usually nonlinear. This makes the relations very hard to analyze and assess in welfare terms.

> ...A little air pollution is harmless, more an annoyance, a great deal lethal. Discrimination against Jews by a random 10 percent of employers, landlords and operators of public places might be merely an annoyance to those affected; by 40 percent, a real hardship; by 90 percent, an economic and social catastrophe. (*ibid.*)

The second intractable problem arises from the fact that weighting is usually required:

> ...To combine robberies and murders in a crime index; to combine pollution of the Potomac and pollution of Lake Erie in a water pollution index [etc.]. An expert in the field may be able to provide judgments...that would permit an interesting index to be calculated. *However, the necessity for numerous individual judgments that are difficult to assess or even to describe must impair general acceptability of measures based on them*. (*ibid.,* emphasis added)

Already very complicated, matters become still more complicated when the construction of an index of indexes is contemplated. Denison flatly rules out the possibly of accomplishing such a construction a generally-acceptable way. This prompts him to make another statement of special importance:

> ...I can imagine only letting each individual compute his own index with his own personal weights, and then averaging them. But even this procedure is almost sure to be biased because we are all concerned with the aspects of the environment that *currently* are problems. Who would now think to consider the dangers of attack by hostile Indians?...Even the very recent elimination of refrigerator doors that cannot be opened from within that cost the lives of so many children is almost forgotten....(*ibid.*, emphasis added)

As has been seen, the relevance of income distribution has been recognized by analysts since national accounting's earliest days. Denison endorses the centrality of some notion of the "goodness" of income distribution to any welfare measure but, unlike other analysts, he only does so to cite the lack of agreement as to the "ideal distribution from which departures could be measured" as further evidence of the inherent futility of attempts to construct such a measure. To round out his list of obstacles, Denison repeats the familiar observation that "tastes differ among individuals and change over time" and, then, touches on another consideration:

> ...I have not yet recalled that welfare is affected by people's perception of reality as well as the objective facts; one's fear of crime on the streets need not be closely related to actual risks. The authors of "Towards a Social Report" [U.S. Dept. of Health Education, and Welfare, 1969] stressed the need for attitudinal data to develop welfare measures. (*ibid.*)

Denison reiterates that even if the foregoing indexes (on real costs, needs, etc.) could be constructed, the absence of an acceptable weighting system rules out a composite index. "Certainly statisticians and social scientists are in no position to assign weights." He then extends this dilemma to the classic dilemma faced by all decisionmakers.

> ...Most decisions that might be made have favorable and unfavorable effects on various aspects of life. Decisionmakers must try to determine the favorable and unfavorable effects of alternatives and then decide their course of action. Economists [etc.] can help determine what the effects are likely to be, but the responsible decisionmaker must decide how [these diverse effects] balance out *and different persons will decide differently*. This is another way of saying that a generally acceptable weighting system does not exist. (*ibid.,* emphasis added)

There is some overlap among the views reported in this section but they are in many respects compelling, and they will be drawn upon in formulating and answering the questions addressed in the concluding Part.

PART IV

CONCLUDING QUESTIONS AND ANSWERS

IS GDP AN ADEQUATE MEASURE OF ECONOMIC WELFARE?

If the term "adequate measure" is meant to refer to a measure that is suitable both for intertemporal comparisons, with respect to a given country, and international comparisons involving two or more countries, the answer is clearly no, on both counts. Taking the second comparison first, it is evident that even between societies that have a great deal in common, such as the Canadian and American, there will always be enough differences in tastes, wants and product mix to render at least somewhat dubious rigorous comparisons for even a single year based on their respective GDPs, let alone comparisons over longer intervals. As to intertemporal comparisons with respect to a given country, so onerous are the conditions that must be met before it can confidently be inferred that a given change in GDP *unambiguously* implies an analogous change (in ordinal terms) in economic (and therefore in general) welfare, that, Usher's more optimistic view notwithstanding, prudence dictates a negative answer. The same prudence dictates a negative answer ever more strongly the longer the interval of comparison. Whether a given requirement (say, constant tastes) is more or less onerous than another (say, full substitutability of goods) is immaterial. All the requirements must be met if comparative GDP is to provide a rigorous basis of comparative welfare, and the probability of that occurring, even over a fairly short interval, seems low.

It is important to recognize, however, that serious though these caveats are, they do not entirely deprive annual changes in GDP of normative significance. As Little and Kuznets both point out, each in his own way, a *substantial* annual increase in GDP would probably prompt most people to agree that most, though by no means all, members of society are, in varying degree, better off as a result. The normative significance of this agreement, though limited, is real.

Of greater, though still limited, normative significance is the fact that a given year's GDP is produced by a national-accounting system designed to report the interaction of the main components of a market economy. Thus, GDP is the sum of the monetary values of innumerable, freely-made choices of individuals pursuing their own self-interests as consumers and/or entrepreneurs. In a different sense and to a lesser degree, the same may be said of the government expenditures included in GDP. Without endorsing the almost frivolous suggestion of Hicks (1940) that the expenditures of even the least representative of governments can ultimately be attributed to the choices of individuals, it is reasonable (with due regard for Arrow's impossibility theorem) to suggest that in North American and Western European democracies, such attribution is defensible. And without denying that the connections are sometimes tenuous, it is also a reasonable assumption that these market economies are sufficiently competitive to ensure that prices bear enough of a relationship to the corresponding marginal utilities to justify the preceding majoritarian welfare comparison.

Although these considerations increase the normative significance of substantial changes in GDP, that aggregate remains an inadequate welfare measure. One reason is the existence of significant market imperfections, be they restrictions on competition in specific markets or market failures of one kind or another. Another, and more important, reason is the fact that individual decisions, freely-made in markets for private goods though they may be, can be wise or unwise, far-sighted or short-sighted, and responsible or irresponsible with respect to the best interests of both the individuals themselves and the society to which they belong. Likewise, the volume of expenditures for government services, ultimately endorsed though it may be by collective decisions, provides no information as to the efficacity of the services. More generally, it is in the nature of both private-market transactions and the provision of government services that although prompted by individual and collective wants, they are both inherently incapable of illuminating either the circumstances that induced them or their consequences. These dimensions can only be indicated, and then only incompletely, as will be seen, by additional information, which for the most part is not generated by the national accounts.

CAN SOME OTHER MEASURE, DERIVED FROM GDP, SERVE AS AN ADEQUATE MEASURE OF ECONOMIC WELFARE?

In addressing this question, a great deal depends upon how the term "welfare" is understood. If it is understood as a generalized level of satisfaction of human wants, the caveats concerning changes in GDP that collectively disqualify them as a measure of a country's comparative-welfare trend, over more than a short period, apply *a fortiori* to measures derived from GDP. This remains true even if a distinction is made between "welfare" and one of its components "economic welfare," but the question takes on a different character. This difference was alluded to when N-T's MEW was discussed and it was noted that they do not delineate economic welfare from generic welfare. It was also noted that analysts belonging to their school implicitly couch the issue, not in the ordinal terms that characterize analyses of the Hicks-Kuznets type, but in cardinal terms. The caveats carefully developed by Hicks, Kuznets and the others, which collectively are so damaging to the enterprise of inferring welfare changes from changes in GDP, tend to go unrecognized (Usher is a rare exception), much less satisfied. To this worrisome situation must be added the concerns arising from various problematic components of their respective measures. Among the most problematic is the widespread exclusion of expenditures on many government services included in GDP. Whether these expenditures are described as "instrumental expenditures" or "expenditures on necessary evils," it is their involuntary character that allegedly disqualifies them. They are further disqualified by their alleged inability directly to produce utility.

Although these two disqualifying grounds tend to be combined by analysts of this school, they are in fact distinct. The second of these grounds were considered above with respect to national defence, a prime example of such a government service, and it was shown that they depend heavily upon the assumption that man lives by, and for, bread alone. Ignoring the very different treatment of business expenditures on fixed costs, which are no less indispensable to output and whose contributions to it are no less indirect, they insist that this renders national defence an intermediate rather than a final product. If, however, this unidimensional view of the human condition is rejected, the argument that, in addition to preserving productive capacities, national defence provides desiderata of great and independent utility (such as peace and liberty) comes into its own and implies that defence expenditures warrant inclusion in a welfare measure, economic or general, as fully as any others.

The first grounds were also considered above, and the key question was seen to be: which consumption expenditure is not incurred in response to a "necessary evil" of one kind or another? This is another way of asking whether there actually exist consumer goods that are not only final products which directly produce utility, but are also so luxurious and unessential that they are only purchased at the whim of consumers that can painlessly forego them. To put the question thus is to demonstrate that quite apart from the separate and rather profound concern (separately expressed by both Denison and Okun) about the evil informational consequences of obscuring the distinction between providing for necessities and failing to do so, very few, if any, commodities fall into that happy category.

In sum, together with the onerous caveats previously identified, these difficulties fatally undermine the grounds upon which rest GDP-derived economic-welfare measures that exclude instrumental expenditures, and the like. This leaves GDP standing supreme in the field of single measures, warts and all.

CAN ANY SINGLE MEASURE OF ECONOMIC WELFARE BE ADEQUATE?

This question may be reformulated as follows: can all the necessary components of a single measure of economic welfare be specified and quantified in monetary terms and, if so, can it be done *objectively*? The answer is again negative, for several reasons, some of which can (yet again) be exemplified by defence expenditures. To begin with the first subquestion, it was just noted that expenditures by individuals, whether acting as consumers or entrepreneurs, may be wise or unwise, responsible or irresponsible, and so on. So long as they are incurred for a "productive" reason, the expenditures qualify for inclusion in GDP (properly so, given GDP's officially avowed purpose of measuring output), but whether they increase or decrease society's economic or general welfare is another question altogether. It is not only true, during any given era, that one man's meat can be another man's poison, and it can also at the same time be the *same man's* poison, it is equally true that one era's meat can be another era's poison. Alcohol consumption and gambling are two obvious examples of expenditures that currently enter into GDP and which have both current and subsequent deleterious social consequences, some of which also enter into GDP, sometimes (interestingly) positively, and sometimes negatively. Many analogous examples could easily be provided. Presenting a more complete welfare picture than emerges, even under the best of circumstances, from GDP alone requires that additional variables reflecting associated physical- and mental-health effects, and the like, be reported alongside it. Some of these effects can be quantified monetarily; others cannot.

To illustrate the latter, recall that both Hicks and Kuznets emphasize the importance of data on changes in income distribution to any assessment of the welfare implications of changes in GDP, and both analysts (one implicitly, the other explicitly) call for the reporting of these data as complements to GDP. Recall also that Little makes an analogous point pertaining to data on changes in working conditions that accompany changes in GDP. It is a common feature of both types of data that they cannot adequately be rendered in monetary terms alone: this is especially true of the latter type. Finally, it is enough to refer, without recapitulation, to Denison's grounds for concluding that a single welfare measure which transcends the subjective values and biases of the analyst that constructs it is inherently unattainable. The fact that his grounds (or anyone else's, for that matter) are not, and cannot be, exhaustive makes the case is all the stronger.

HOW SHOULD ECONOMIC WELFARE BE MEASURED?

Parts of the answer to this final question can be found in the answers to the preceding ones. They establish that only a set of complementary variables can qualify as an adequate measure of economic welfare, but it can only qualify as a contender, one of several, perhaps many, for this is a contest in which there can be no undisputed champion. As will be argued forthwith, any proffered set of variables is bound to be, at least in part, both problematic and subjective, not only with respect to the identification and specification of its members, but also with respect to their valuation in monetary terms.

Little's view, mentioned earlier, with respect to economic welfare now requires closer scrutiny. He says:

> ...[I am] interested in *economic* welfare. There is no part of well-being called 'economic well-being'. If I am interested only in someone's economic welfare, then I interest myself

only in the economic things that may affect his well-being. (p. 6, emphasis in text)

"Economic things that may affect well-being" are better described as contributions to economic welfare.

The distinction between economic welfare and welfare *tout court* is not a factor in the analyses of Hicks, Kuznets and those who share their approach. These analyses involve a comparison between two levels of total private and public expenditures for alternative (or successive) combinations of marketable commodities and government services, because their normative relationship is expressed ordinally, there is no need to consider, let alone measure, the absolute levels of welfare associated with the combinations of commodities and government services. Matters are very different indeed in the case of the genre of measures exemplified by NT's MEW.

If we accept Little's view that economic welfare is meaningfully conceived only in terms of economic contributions to welfare, we must reckon with the fact that, as Jaszi points out, government services often produces two or more products jointly, whose contributions to welfare often fall into more than one category. (National defence, which produces peace and liberty, while preserving the economic structure, is an obvious example.) Going beyond Jaszi's discussion, it should be recognized than when contemplating an input that produces joint products or byproducts (and there are many such) from a reporting standpoint, it is necessary to know in advance which category of contributions to welfare is the primary one. In the case of joint products, sound accounting practice requires that the costs of the inputs be prorated among the products. In the byproduct case, the entire cost of the input must be attributed to its primary product, whose total cost is then reduced by the combined market value of the other products. The question thus arises as to the criteria by which one of the perhaps several categories of contributions to welfare jointly served by a given government service should be given primacy.

It is very doubtful, however, that Little's view is operationally sound. If economic contributions to welfare exist as a distinct category, so do cultural contributions, physical and health contributions, and many other contributions exist. Is it meaningful to differentiate, for purposes of welfare measurement, between categories of contributions to welfare in the expectation that all human wants are such that their satisfactions are scientifically measurable?

Put another way, is it possible to develop not only a generally-acceptable taxonomy of categories of contributions to welfare but, also, a generally-acceptable list of their respective components? Little may usefully draw a conceptual distinction between categories of contributions to welfare and welfare itself but, because he adheres to the subjective-valuation-of-combinations-of-commodities school of thought, he has no need to establish that his distinction is operationally sound. Analysts not so happily situated must face the fact that it is highly unlikely that the development of a generally-acceptable taxonomy of categories of contributions to welfare is a realistic project.

Fortunately, there is no need to undertake it. The variety of rubrics under which contributions to welfare can be categorized is substantial, but most contributions have in common the fact that they can only materialize, in economies such as ours, through economic transactions that have been consummated in markets. When we purchase theatre or concert tickets we make cultural contributions to welfare, to utilize Little's distinction. When we purchase memberships in health clubs we make physical or health contributions to welfare, and so on. It may well be that as a result of these and other contributions to welfare, we become happier and healthier people, able to perform our jobs more efficiently than before, in spite of the fact that when we decided to make them we did not do so for that reason. If the fact that these contributions to welfare belong to different categories is subordinated to the fact that they are all manifested by economic exchanges, and (Little to the contrary notwithstanding) they are therefore classified as economic contributions to welfare, or simply, in N-T’s terms, as economic welfare, the number of purely noneconomic contributions to welfare shrinks into relative insignificance. On the other hand, not only has the number of human wants whose degree of satisfaction warrants recognition in a measure of economic welfare greatly increased, so has the complexity and subjectivity of the measure. If, irrespective of its immediate purpose, every transaction is deemed to be a contribution to economic welfare simply because it is a transaction, then the sky is virtually the limit.

Laidler (1985) invokes Adam Smith's concept of *homo oeconimicus*:

> ...[He is] a simple creature. The self-interested acts of producing, trading, and consuming scarce goods are the sum total, not just of his economic life, but of his social life as well.

> *His general welfare is identical with his economic welfare....* (p. 7, emphasis added)

By contrast, *homo oeconomicus's* more realistic, and far more complex, counterpart, *homo sapiens*, must be contemplated in more comprehensive terms.

> ...In particular, his economic welfare, conceived of as the satisfaction he obtains from the consumption of goods and services (net of the effort put into obtaining them) is only one aspect of the welfare that should concern us when we discuss the design of social, political, and economic institutions and policies. Freedom has many dimensions that are important for individuals apart from the freedom to exchange goods and services in the market place; equality has many dimensions apart from the equality of income and wealth. (*ibid.*)

But, Laidler insists, the economic and noneconomic spheres are not mutually exclusive and independent of each other.

> ...[A]ny conclusions we may reach about...enhancing economic well-being by taking or not taking particular actions, must be subjected to questions about their continuing validity when the pursuit of non-economic ends is taken into account, before they are accepted as providing a firm basis for policy actions.... (*ibid*.)

By referring only to freedom and equality as ends that fall under the noneconomic rubric Laidler implies that the economic rubric subsumes most, if not all, other ends pursued by *homo sapiens*. And by insisting that the criteria for assessing the validity of economic behaviour must include these noneconomic criteria, he also legitimates "economic-welfare" measures (in this context, the terms "welfare" and "well-being" are synonymous), like that of Osberg (1985), about which more shortly. This is tantamount to endorsing the just-summarized notion that all human needs which can only be satisfied, in a market economy, by households engaging in consumption spending are, therefore, economic needs, and their satisfaction increases economic welfare. Such expenditures obviously go far beyond those made to obtain the wherewithal to earn a living; they include all goods and services purchased in the market place, be they cultural, recreational or anything else. However, whilst this approach endorses N-T's description of their welfare measure as a "measure of economic welfare," and the other analysts' analogous descriptions of their respective measures (although giving no support to the notion of instrumental or regrettable expenditures), it subjects all such measures to additional constraints, as will be seen forthwith.

It also exposes a serious limitation of this genre of welfare measures. The only market of which they take serious cognizance is the market in which households spend money directly, to satisfy various perceived needs by means of *private* goods and services. But another market, which is just as real, tends to be ignored by the measures. This is the market in which the collectivity of households that comprise the nation simultaneously spends money indirectly, via the government, to satisfy other perceived needs by means of *public* goods and services. If spending is to be the crucial criterion, the spending that takes place in the latter market is as relevant to the measurement of economic welfare as the spending that occurs in the former market, although the commodities naturally differ. (It is in the nature of private goods that every unit consumed by a given consumer reduces the supply available to all other consumers; by contrast, it is in the nature of public goods that they are consumed simultaneously by every consumer, and the units consumed by any one of them leave undiminished the supply available to all the others.) The important consideration for our purposes is the fact that expenditures in both markets represent more-or-less deliberate choices by households, whether wearing their consumer hats or their citizen hats. As Kuznets (1933) argues, collectively-consumed commodity are as deserving of inclusion in an economic-welfare measure as any other commodity, whether privately or publicly produced. Interestingly, in spite of its importance to their approach, the expenditure criterion is not always adhered to by analysts of this school. As will shortly be illustrated with specific examples, alternative measures of economic welfare commonly include differing (and often negative) amounts not reflected in market transactions but nonetheless deemed pertinent.

The challenge of dealing with distributional equity is among the most important of the constraints upon the measures under consideration. Laidler links the development of economics as a discipline with utilitarian philosophy, a philosophy that:

> ...requires us, as a matter of moral principle, to regard all individuals as having equal capacity for enjoyment and suffering, and to regard social welfare as being synonymous with the sum of individual welfare...*At the very least, we are compelled to regard those changes which enhance the economic welfare of the poor at the expense of the rich as being superior to those which have the opposite effect, always provided that there is nothing to choose between them on other grounds*. (p. 9, emphasis added)

These words underscore, albeit for a different reason, the Hicks-Kuznets view that changes in income distribution must be reported as complements to changes in GDP, so as to permit an assessment of the welfare implications of the latter changes. They also underscore Little's more general view that the very concept of welfare is at bottom an ethical one. As will be seen, however, these views do not constitute the last word on the subject.

Before considering Osberg's measure, another of Laidler's observations is worth noting.

> The economic well-being of any society is multifaceted, and it would be surprising indeed if there were any simple formula for the organization of economic life that would ensure that it is always maximized. *Indeed, given the complex nature of economic well-being, it is a mistake to think in terms of having some easily measured achievable maximum value about which everyone can agree.*
>
> *Nevertheless, the fact that an unambiguous and universally acceptable standard of assessment is not always and everywhere available is no excuse for taking the nihilistic approach to the problem of judging the consequences of varying changes for economic well-being.* (p. 11, emphasis added)

In quoting Adler (1982) as follows, Osberg explicitly recognizes the relevance of the first of these caveats.

> ...There is an almost unanimous agreement among social indicator workers that it is neither practical nor theoretically desirable nor analytically sensible to have one measure of economic welfare. (Osberg, pp. 72-73)

Osberg also draws attention to some of the pitfalls of expressing in average terms important components of economic welfare, like GDP, and suggests that public discourse on economic issues will benefit if people are enabled to look beyond the averages. Public discourse will also benefit if:

> ...we could consider separately each aspect of well-being *and assign the weight each of us considers appropriate...*

However, Osberg proceeds to discount this point by completing the preceding sentence with the following words:

> *...but public debate will not be assisted by an incomprehensible deluge of esoteric statistics.* (p. 73, emphasis added)

Such is the choice, according to Osberg, that confronts the analyst seeking to convey the level of economic welfare at any given time, as well as changes therein over time: either present a single measure reflecting both the analyst's personal conception of each aspect of economic welfare and the weight he assigns to it, or present a flood of statistics no nonspecialist can hope to comprehend. Osberg does not explain why he sees only this one stark choice, a dichotomy not necessarily seen by the large "nearly unanimous" group of analysts he mentions. More particularly, he does not explain why each analyst should not present his personal set (which need not be a flood, by any means) of unweighted statistics, such that each member of the set is fully comprehensible to the celebrated intelligent layman. This approach would encourage the latter to decide for himself whether the set and its members are satisfactorily specified and measured, and, no less importantly, also to assign his own weights to each member.

Before considering the question of which approach places the more onerous burden upon the intelligent layman, return to Osberg's measure, which he labels "index of economic well-being" (IEW). Four "main dimensions" of IEW are identified: the aggregate flow of effective consumption, net additions to the stock of productive resources, poverty/inequality of lifetime income, and economic insecurity in year-to-year income flows. Briefly, the first includes the (per capita) sum of consumption of marketed goods and consumption of household production, leisure and other unmarketed goods; the second includes the sum of business and personal capital formation, net accumulation of training and R&D investment, net changes in the value of natural resource stocks, environmental costs and net foreign debt; the third includes a poverty measure consisting in income shares, the fraction of the population with an average income below the median income, along with income needed to raise all poor households to a poverty-level income; and the fourth includes the level and rate of increase of unemployment, the percentage of the labour force that can expect over-twenty-percent variations in average real income, and the change in the rate of inflation.

There is only space here for a few brief comments on IEW, of which one of its several striking features is the fact that national-defence-type expenditures are excluded but, paradoxically, not personal spending on locks, burglar alarms, private security guards, and the like. Nor are private burglary and theft insurance excluded along with the business equivalents that enter into the prices of consumer and investment goods. Defence-type expenditures are declared to be made entirely made to acquire "bads...in the hope of avoiding something worse," but it is not explained why the avoidance of a worse "bad" is not itself a "good." Osberg goes so far as to declare categorically that people will not feel better off if increased spending on police protection reduces their holiday time in the Bahamas, although it is not obvious that, given the choice, a reduction in police protection that allows them more holiday time in the Bahamas, at the expense of more muggings and burglaries at home, will be preferred by these same people. More generally, Osberg does not deal with the reality that there are unavoidable tradeoffs between the desiderata of holidays and security (to mention only these), and that different combinations of them will be evaluated differently, both at a given time and over time, not only by individuals but by communities. Granted, all this renders the task of placing a value upon any actual combination very difficult indeed, but that is all the more reason to put the problems on the table, as it were, not in order to declare the situation hopeless, but to see what can be done in spite of them.

Like other analysts of his school, Osberg takes no cognizance of the highly varied, not to say anomalous, treatment, so laboriously set forth by Jaszi and Gilbert *et al.,* which national-accounting conventions afford economically-similar transactions completed at different points in time. At the same time, Osberg implicitly deems all private expenditures on consumption and investment goods irreproachable from an economic-welfare standpoint.

Osberg justifies his inclusion of an income-inequality measure, not in the technical terms of Hicks and Kuznets, but in ethical terms. He declares himself an adherent of the theory of justice of Rawls (1971), which places a premium on alleviating the poverty of the poorest members of society. However, as Laidler makes clear, other reputable theories of distributive justice also exist (e.g., those of Nozick (1971), Dworkin (1981) and Gilder (1981)), which have different implications for the measurement of economic welfare. An indication of the range of thoughtful opinion on the subject is given by following quote from the Encyclopedia of Philosophy (1967).

> ...Questions of [distributive justice], according to Hume, Mill and others, presuppose conflicts of interest; there would be no point in talking about [it] but for the limitations of human benevolence and the competition for scarce goods. *[Distributive justice] presupposes people pressing claims and justifying them by rules or standards....*

Whereas, the Encyclopedia continues:

> It would be generally agreed that doing [distributive justice] means treating equals equally

> and unequals according to their relevant inequalities. *Disagreements arise over the criteria of relevance -- that is, over the rule to be applied.* Distribution (for example, of income, taxation, social service, benefits, rations) may be organized on any of at least three principles of justice: arithmetical equality, merit (or desert) or need.... (p. 301, emphasis added)

The important point in the present context is that oceans of thoughtful ink can be, and have been, spilt, first in defining equality and degrees of inequality, merit and need, and, then, in developing distributional-equity arguments under these rubrics. It is no reflection upon Osberg's Rawlsian approach to say, first, that it is far from having the field to itself and, second, that the intelligent layman contemplating IEW is likely to be aware of this fact.

Taken as a whole, IEW is a mixture of approaches, all of them subjective but in different ways. On the one hand, the subjective choices of individual consumers and entrepreneurs are taken at face value. On the other hand, important government services are deemed unhelpful to economic welfare, a specific theory of justice is invoked, and economic insecurity is deemed inherently deleterious (presumably, whatever its level). Each of these positions is in its own way and in varying degree far from unanimously held. Then, each of Osberg's four "aspects of economic welfare" is assigned a weight, again according to his personal ranking, and aggregated to produce IEW.

To further demonstrate that singularity and subjectivity are the order of the day when it comes to measures of economic welfare, however labelled, consider two additional recent measures. Cobb and Cobb (1994) present their "index of sustainable economic welfare (ISEW), describing it as a "Green GDP" namely, a GDP adjusted to reflect environmental and other factors ignored by GDP. Its details can be found in Appendix IV. As usual, the measure's starting point is consumption expenditures, but these are weighted by the authors' own index of distributional inequality. The weighted consumption expenditures are then increased or decreased by a series of amounts deemed to be the values of factors ranging from household services, through public and "private defensive" expenditures on health and education, expenditures on advertising and the costs of auto accidents, to various environmental costs, including loss of wetlands and farmland. Apart from the weighting for distributional inequality, all other components of ISEW are unweighted monetary values.

ISEW's slightly later "successor index" namely, the "genuine progress indicator" developed by Cobb, Halstead and Rowe (1995) omits some items included in ISEW and adds positive or negative amounts for the new value of volunteer work, as well as for the costs of crime, family breakdown, and underemployment. As before, only consumption expenditures are weighted.

It is enough to set the respective components of NT's MEW, Osberg's IEW and the Cobb's ISEW side by side to show that, specification and measurement problems apart, there can be little in the way of social phenomena that could not plausibly be deemed either a social pathology or its opposite and, therefore, suitable for inclusion in a measure of economic welfare. This problem is quite separate from the problem of the propensity of analysts to prefer radically-different ways of addressing similarly-described phenomena. To illustrate, MEW, on the one hand, and these two measures, on the other, purport to capture the dimension of "sustainability, but they go about it in ways that differ so profoundly that it is hard to imagine that their respective authors are contemplating the same thing. Analogous incompatibility and noncomparability attend the treatment of health and education expenditures. The enormous overall incompatibility and noncomparability of the measures under consideration can be seen in Appendix IV, which lists the variables that comprise them.

It cannot be too strongly emphasized that none of the foregoing caveats and criticisms is intended to imply that efforts to measure social welfare, and changes therein, are foredoomed to failure and should therefore be abandoned. Far from it. Just as people need an analogous sense with respect to the body politic, they need a sense (based upon more than the evidence of their personal experience and hearsay) as to whether, overall, things are getting better or worse in the body socioeconomic. They will satisfy that need in one way or another; and, just as the absence in the public domain of an authoritative single measure of political welfare is not, and should not be, cause for concern, neither should the prospect that an authoritative single socioeconomic measure is unattainable be alarming.

Every individual ultimately assesses the trend of the nation's political health, at a given time, by looking at those variables he or she considers both relevant and significant, weighted on the basis of his or her personal value system and, doubtless, the variations among individuals of these variables, and especially of their weights, are as great or greater than those of the socioeconomic variables and weights under consideration. The resulting assessment (presumably in ordinal terms), represents the individual's judgment as to whether, on balance, things political are "better" or "worse" than they were "last year." So, analogously, must individuals assess, for themselves, economic welfare and its trend, not on the basis of a generally-acceptable single measure (no such measure exists, or is ever likely to exist) but on the basis of a self-chosen set of variables, subjectively weighted. The members of that set may or may not all exist in the public domain, but that consideration is secondary to the fact that it will be an individually-chosen set, subjectively weighted and not easily comparable in ordinal, still less in cardinal, terms with the sets chosen by other individuals.

To demonstrate the inevitability of such an outcome, return to our intelligent layman, still trying to form an informed opinion on how things are going socioeconomically, only now he is contemplating not only Osberg's IEW, but, also, the other two measures whose components are listed in Appendix IV. Assume that these three measures are the only ones available. What is he to make of them, since it is altogether likely that if he does not deem relevant all the listed variables, he deems many of them relevant, and he also deems relevant other variables that none of them lists? Since the "bottom line" values of the measures are fundamentally noncomparable, he is likely to disregard them. What he is bound to do is to collect together those listed variables that he deems relevant and, if he agrees with them, he will collect their assigned monetary values as well; otherwise he will supply his own estimates. He will also collect those variables not listed that he deems relevant, along with their respective values that he considers plausible. Finally, he will assign to every one of these values the weight he considers appropriate, whatever the listed weight. The end result will be the personal index that forms the basis of his judgment as to the trend of the country's socioeconomic health.

If this reasoning is correct, the most sensible approach to welfare measurement, whether welfare is defined broadly or narrowly, must be this: to a large extent, the assessment of welfare, like the assessment of beauty, is for the eye of the beholder. Therefore, let a hundred measures bloom, but let none claim that it, alone, is the truth, the whole truth and nothing but the truth.

APPENDIX I

INDEX NUMBERS AND THE MEASUREMENT OF CHANGES IN THE LEVEL OF ECONOMIC WELFARE

Hicks

Using Hicks's notation: let E = ($\sum p_2q_2/\sum p_1q_1$) represent the index number of money expenditures at the second date, using the expenditures of the first date as the base; let L = ($\sum p_2q_1/\sum p_1q_1$) represent the (Laspeyre's) index weighted by the quantities of the first date; and let P = ($\sum p_2q_2/\sum p_1q_2$) represent the (Paasche) index weighted by the quantities of the second date. The criterion for a higher welfare level at the second date thus is $\sum p_2q_2 > \sum p_2q_1$, which implies that, after dividing by $\sum p_1q_{1,}$ E > L. The criterion for a higher welfare level at the first date is $\sum p_1q_1 > \sum p_1q_2$ or $1/\sum p_1q_1 < 1/p_1q_2$, which implies, after multiplying by $\sum p_2q_2$, that E < P.

Hicks identifies four alternative situations. One, E may be greater than either L or P, which means that welfare has increased. Two, E may be less than either L or P, which means that welfare has decreased. Three, E may be less than L and greater than P, which means we cannot categorically say whether welfare has increased or decreased. Four, E may be greater than L and less than P, which means, absurdly, that welfare has both increased and decreased (Hicks suggests that the underlying problem here is that wants have changed).

Expanding upon the third situation, Hicks recognizes that although sufficient information is not available to enable us to identify them, there usually will exist some values of E < L that indicate an increase in welfare and some values of E > P that indicate a decrease. Thus, it is justifiable to suggest that welfare has probably increased when E is only slightly less than L, and that welfare has *probably* decreased when E is only slightly greater than P. "This is as far as we ought to go."

APPENDIX II

SUMMARIZED SCENARIOS INVOLVING INTERMEDIATE PRODUCTS, FINAL PRODUCTS AND GOVERNMENT SERVICES

The following alternative situations are conjured up in Jaszi (1946) to convey an important message. The problem of differentiating intermediate from final products when calculating GDP for purposes of welfare measurement, especially in an economy in which production processes change over time and in which the government participates by supplying (or not supplying) certain services, is considerably more complicated than is generally recognized.

I - There are two firms, one producing cloth, the other producing suits made from the cloth. Labour is the only other input and government is not a participant. Everything is straightforward and there is no double counting in GDP.

II - Hitherto unemployed workers are now employed to double the work force, and GDP therefore doubles. Government is still absent from the picture. Everything is still straightforward and there is no double counting.

III and IV - The government takes over the production of cloth and provides it free to the firm producing suits, which reduces its prices accordingly. If Scenario I and Scenario II are both replayed under this arrangement as Scenario III and Scenario IV respectively, Scenario IV's output in terms of the output of private business is double that of Scenario III. If an index of output is constructed for these scenarios, and the intermediate product provided by the government is included, it too shows an increase of 100 per cent, whether or not the government's production of cloth is included in GDP. (These results are independent of the particular form of taxation the government adopts to finance its cloth production.) Says Jaszi:

> ...The purpose of this example is to show that the counting of government intermediate output *per se* does not result in duplication, although this belief is...fostered by many current statements...." (p. 205)

But that is not the end of the story, because national accounting that included intermediate product in GDP would produce a distorted picture of relative welfare in these two scenarios .

Remember that suit prices fell in response to the provision of free cloth (although consumers' welfare in term of suits was unaffected thereby), and if the value of those suits in <u>constant</u> dollars were now augmented by the value of the governmental intermediate product, the sum would show an increase in consumers' welfare that is really spurious. This result puts in question the preceding statement that the inclusion of the intermediate product does not distort welfare measurement.

V - Jaszi now hypothesizes a situation in which, instead of the government producing cloth and providing it free to suit makers, consumers now buy their own cloth and hire suit makers to convert it into suits. As they did when the cloth was supplied free by the government, the suit makers reduce their prices accordingly. Standard national accounting does not treat such a decrease in suit prices as a genuine price decrease: real output of suits is deemed to have fallen commensurably, and the cloth purchases are included in real GDP. The government case should be handled in exactly the same manner.

> *There is no justification for adopting a technique of deflation which is not employed in the private sphere, and then correcting for the misleading result by adjusting the current dollar value totals.* (p. 208, emphasis added)

VI - A production process different from the one implicitly assumed in the preceding scenarios is now hypothesized. Suit makers both produce suits and deliver them to customers, employing transportation as well as production workers. Suit prices rise to reflect delivery costs.

VII - An adverse change in transportation conditions is now hypothesized, which forces the suit maker to transfer some production workers to the delivery function. This in turn results in both reduced suit production and higher suit prices. Deflation of these current prices will show that total suit production has fallen.

VIII - Revert to the conditions in VI and assume that the government provides free transportation to suit makers. (As before, the outcome is independent of the form of taxation with which the government finances this service.) Suit makers reduce their prices accordingly and produce the same number of suits as in VI.

IX - Assume the adverse event described in VII, but this time it is the government that transfers workers from suit production to the provision of transportation services (increasing taxes to finance the transfers). Suit production falls, although suit prices remain unchanged. Real GDP has fallen, a fact that would be obscured if the intermediate product supplied by the government were included in GDP. That, however, would be considered double counting.

X - Assume a regime under which transportation services are bought by consumers from a transportation firm. Real GDP now consists of both transportation services and suits.

XI - Now assume the adverse transportation event described in VII occurs under the regime in X, and it again causes workers to be transferred from suit production to the provision of transportation services. Suit production naturally falls, although suit prices remain unchanged. Jaszi points out that, logically, there is no more justification for including the transportation services bought by consumers in GDP than there was for including it in the circumstances described in IX, when it was acknowledged that doing so would result in double counting. Here, too, suit production is reduced, exactly as it is in IX, but standard national accounting includes the transportation services in GDP, which it refuses to do in IX. In Jaszi's words:

> ...Counting the service charges of [the transportation firm] obviously leads to duplication: the welfare of the consumer is exaggerated. But in fact this line of argument is not adopted. There is no suggestion to omit payments to [the transportation firm] from the current value figures in order to avoid duplication. *There does not seem to be any reason for adopting a different procedure if the threatened duplication is in the sphere of government*. (p. 813, emphasis added)

APPENDIX III

SOME FURTHER THOUGHTS ON DEFENCE EXPENDITURES

Varied and contentious though it is in many respects, the discussion in the literature of the proper place of defence expenditures in a welfare measure generally assumes that the expenditures serve to protect and preserve (i.e. leave unaltered) the “sum of things,” be it the economy’s productive capacity or that capacity plus various social values, such as peace, liberty, justice, and the like. As was foreshadowed above (in the suggestions that the impact of a given level of expenditures upon national security depends upon the reactions of other countries), this static view of the impact of defence expenditures seems inadequate: defence expenditures generally change the “sum,” either for the better or for the worse. Though not impossible, the mere-preservation impact (i.e. no change) is highly unlikely. Granted, as indicated above, all expenditures are capable of proving wise or unwise, responsible or irresponsible, constructive or destructive, and so on, but these contingencies are so much greater in the case of defence expenditures that they warrant special attention in a discussion of the measurement of welfare. This paper is hardly the place to attempt a proper discussion of the implications of this highly complex situation; but a few basic considerations, some of them prompted by N-T’s assertions, can tentatively, and somewhat speculatively (hence this appendix), be sketched.

At any given time, a country's national security depends upon the degree to which it faces two risks: the risk of becoming involved in a war and the risk of defeat in it. Both risks are sometimes inversely related to the country's defence expenditures, as when the truth of the old adage, if you want peace, prepare for war, is confirmed; and sometimes they are directly related, as when the very preparations for war increase the probability of its occurrence. The awkward fact is that a given change (in either direction) in a country’s defence expenditures will, in some circumstances, influence other countries in a way that increases its national security and, in other circumstances, influence them in a way that decreases it. There is no general case. Taking no cognizance of these considerations, N-T flatly deny that American national security increased substantially during their long interval of observation, 1929-1965, leaving the reader free to infer that it remained fairly stable over these years. Precise and objective assessment of either risk is impossible, so there is no way of determining the overall relationship, during that interval, between American defence expenditures and American national security. Still, the questionable soundness of their categorical judgement becomes apparent when the interval is divided into subintervals during which events of tremendous and world-historical import first loomed and later occurred, events like the Second World War.

N-T's more general reasons for excluding defence expenditures from MEW refer only to the event, war, not to the risks of war that are the true concomitants of their so-called "deterioration in the environment;" and their analysis suffers as a result. As they suggest, national accounting automatically captures the fruits of technological progress, managerial innovation, a new mineral discovery, improved weather, and the like, but a change, say an increase, in the risks of war is not analogously reflected, even though welfare is as reduced by this change as it is increased by the others. Since risks of war are contingencies and not actualities that are sooner-or-later manifested in transactions, welfare, as measured by MEW or any comparable measure (including GDP), reflects neither the true economic cost namely, the present value of the expected economic costs of war, borne by a country exposed to them nor, more pertinently, changes in it. Nor, by the same token, does any such welfare measure capture the true economic gains that accrue when defence expenditures reduce this cost. These unreflected gains are fully as real as the corresponding, and reflected, gains attributable to the other, positive "environmental" changes identified by N-T.

This rudimentary sketch of the effects of defence expenditures upon national security and upon actual (as opposed to "measured") welfare also allows parenthetical clarification of another case mentioned by N-T. In it, in contrast with the preceding case, increases in both measured and actual welfare are associated with reduced defence expenditures. N-T contend that an arms control agreement or a disarmament agreement that frees economic resources (which presumably would otherwise be devoted to national defence) and raises consumption by 10 per cent, increases welfare just as much as do new industrial processes. It should first be noted that an arms control agreement and a disarmament agreement generally do not affect defence expenditures in the same way. The former generally freezes defence expenditures at existing levels, the latter generally reduces them; this makes it the only kind of agreement with the potentiality of increasing resources available for consumption. What is more important, however, is the fact that the increase in welfare is really due to the prior reduction in the risks of war, rather than to the disarmament agreement that reflects and formalizes that reduction, a fact that N-T do not recognize. National accounting has no way of capturing this fact, nor has any welfare measure that ignores contingencies.

To repeat, a given level of defence expenditures is as capable of reducing national security and, therefore, welfare as it is of increasing it. The intangibility of the relevant factors greatly complicates matters from a welfare-measurement perspective, for, as indicated, it is virtually impossible to measure objectively the war risks facing a country. They are functions not only of the relative military capacities of that country and its potential enemies, but also of the enemies' intentions. Estimating the former is usually a fairly manageable task, estimating the latter usually is not. It is also impossible to estimate objectively the present value of the expected costs of war. All this makes it impossible to assign a generally-acceptable monetary value to the impact of a given level of defence expenditures on that year's welfare, a conclusion that applies equally to all the definitions of economic welfare discussed above.

Referring to the increase in American defence expenditures over their 36-year period of observation, N-T also say that:

> ...It is patently...reasonable to assume that the rise in expenditure was due to deterioration in international relations and changes in military technology. The cost of providing a given level of security has risen enormously. If there has been no corresponding gain in security since 1929, the defense cost series is a very misleading indicator of improvements in welfare. (p. 8)

One can agree with these words and still insist that they do not add up to an argument for excluding defence expenditures from the measurement of welfare. It is undeniable that military technology developed spectacularly during these years, as did its cost, but what matters is that the technology's destructive capacity, expressed in terms of economic damage, also increased, if anything, even more spectacularly. Hence, even if war risks were the same in 1965 as in 1929 (they were not), the present value of the expected economic cost of war was far higher in 1965 than in 1929. Nuclear weapons existed in 1965. This conclusion would probably hold even if it were argued that the balance of terror that existed in 1965 between the democratic and communist worlds implied the existence of lesser war risks than before, so great was the increase in the destructive capabilities of weaponry. To put it in more general terms, even when war risks hold constant the economic stakes of war grow as the destructive power of weaponry grows, which means that national security decreases. The only way to maintain a constant level of national security is somehow to accomplish a reduction in war risks that is proportionate to the increase in the destructive power of the weaponry available to potential combatants. Although it is trivially true that a defence-cost series is, in itself, a very misleading indicator of improvements in welfare, as N-T maintain, their supporting arguments fail to illuminate many factors that must be reckoned with. Among these highly-complex factors are two related ones: first, the development of military technology has proven, at least for the major powers, to be a rich source of technological innovation in nonmilitary fields (e.g. health care); second, that many military facilities (e.g. airfields, shipyards, tank factories, etc.) are dual purpose. They can as readily be used for civilian as for military purposes, and often are.

How, then, in the face of these formidable problems, should defence expenditures be reported in a welfare measure, whether welfare is conceived broadly, in terms of general welfare, or narrowly, in terms of economic welfare (however the latter is defined), given that they are not to be excluded? Since the soundest approach conceptually (to proceed as recommended by Hicks, Kuznets and Little with respect to GDP and income distribution, and complement defence expenditure data with data on the associated change in national security) is not feasible, the question really is one of choosing the least unsatisfactory alternative. To adopt the approach recommended by Jaszi and (with minor reservations) Kuznets is in effect to assume that whatever the actual level of defence expenditures, it has been well calculated to maintain the existing level of national security. The trouble with this approach is that, all too often in this century alone, it became apparent to everyone retrospectively (and was apparent to some at the time) that the expenditures were badly calculated for that purpose. Consideration might therefore be given to complementing the expenditure data with the best available data on the defence expenditures of potential enemies and allies. This would be better than nothing, although there should be no illusion about the fact that without (unavailable) data on the associated change in national security, the picture is incomplete and, in some situations, misleading. But perhaps it is now more complete and less misleading than it would otherwise be.

APPENDIX 1V

OSBERG: IEW

Level of effective per capita consumption. *(+)*

(a)effective per capita consumption of marketed goods and services; and

Wt.:a1 (b)effective per capita consumption of Household production, leisure and other unmarketed goods and services.

Net societal accumulation of productive resources over policy period *(±)*

(a)net accumulation of tangible capital, housing stocks and consumer durables;

Wt.:a2 (b)net accumulation of training and R&D investment;

(c)net changes in the value of natural resource stocks;

(d)environmental costs; and

(e)net foreign debt.

Poverty and economic inequality *.(±)*

(a)the shares of the poorest 20 percent, richest 20 percent and middle income quintiles in the lifetime incomes of their cohorts;

Wt.:a3 (b)the fraction of the population having an average annual income below a poverty line of one-half the median income; and

(c)the average amount of income it would take to raise all poor households to a poverty-line income.

Security of individual year-to-year income flows. *(±)*

(a)the level and rate of increase of unemployment;

Wt.:a4 (b)the percentage of the labour force that can expect large (20 percent plus) variations in annual real earnings; and

(c)the change in annual inflation rates.

COBB AND COBB: ISEW

Personal consumption	
Distributional inequality	
Weighted personal consumption (1/2)	*(+)*
Services of household labour	*(+)*
Services of consumer durables	*(+)*
Services of streets and highways	*(+)*
Public expenditures on health and education	*(+)*
Expenditures on consumer durables	*(+)*
Defensive private expenditures health and Education	*(-)*
Expenditures on national advertising	*(-)*
Costs of commuting	*(-)*
Costs of urbanization	*(-)*
Costs of auto accidents	*(-)*
Costs of water pollution	*(-)*
Costs of noise pollution	*(-)*
Loss of wetlands	*(-)*
Loss of farm land	*(-)*
Depletion of nonrenewable resources	*(-)*
Long term environmental damage	*(-)*
Net capital growth	*(±)*
Change in net international position	*(±)*

NORDHAUS AND TOBIN: MEW

Personal consumption	*(+)*
Private instrumental expenditures	*(-)*
Durable goods purchases	*(-)*
Other household investment	*(-)*
Services of consumer capital imputation	*(+)*
Imputation for leisure	*(+)*
Imputation for nonmarket activities	*(+)*
Disamenity correction	*(-)*
Government consumption	*(+)*
Services of government capital imputation	*(+)*
MEW net investment	*(±)*

REFERENCES

Adler, H.J. 1982, "Selected Problems of Welfare and Production in the National Accounts," *Review of Income and Wealth*, June, 1982, 121-132.

Chipman, J.S., and Moore, J.C. 1971, "Why an Improvement in GNP Need Not Imply an Improvement in Potential Welfare," *Kyklos*, Vol. 29, 391-418.

Colm, G., 1954, "The Government Sector in National Income Accounts: A Re-examination of a Few Controversial Issues," paper delivered at Conference on Research in Income and Wealth, New York, October 29.

Cobb, C., Halstead, T., and Rowe, J. 1995, "If the GDP is Up, Why is America Down?" *The Atlantic Monthly*, October.

Denison, E.F. 1971, "Welfare Measurement and the GNP," *Survey of Current Business*, January, U.S. Department of Commerce, Washington, 13-16.

Encyclopedia of Philosophy, Vol. 4, New York, Macmillan, 1967.

Gilbert, M. Jaszi, G., Denison, E.F., and Schwartz, C.F. 1948, "Objectives of National Income Measurement, *Review of Economics and Statistics*, August, 179-195.

Gilder, G. 1981, *Wealth and Poverty*, Toronto, Bantam.

Hicks, J.R. 1948, "The Evaluation of the Social Income -- A Comment on Professor Kuznets' Reflections," *Economica*, August, 163-172.

_______, 1940, "The Valuation of the Social Income," *Economica*, May, 105-124.

_______, and U.K. 1939, "Public Finance in the National Income," *Review of Economic Studies*, February.

Jaszi, G. 1958, "The Conceptual Basis of the Accounts: A Re-examination," *Studies in Income and Wealth*, Vol. 22, National Bureau of Economic Research, Princeton, Princeton University Press, 18-126.

_______, 1946, *The Concept of National Income and National Product with Special Reference to Government Transactions*, Unpublished Ph.D. Dissertation, Cambridge, Harvard University.

Kuznets, S. 1951, "Government Product and National Income," *Income and Wealth*, Series I, International Association for Research on Income and Wealth, Cambridge, Bowes and Bowes, 178-244.

_______, 1948, "On the Valuation of Social Income -- Reflections on Professor Hicks' Article, Part I," *Economica*, February, 1-16.

_______, 1941, *National Income and Its Composition, 1919-1938*, Vol. I, New York, National Bureau of Economic Research.

_______, 1933, "National Income," *Encyclopaedia of the Social Sciences*, Vol XI, 205-224, reprinted in Fellner, W. and Haley B.F. (eds) (1946), *Readings in the Theory of Income Distribution*, Philadelphia, Blakiston.

Laidler, D. 1985, "Economic Ideas and Social Welfare: An Overview," in Laidler, D. (ed), *Approaches to Economic Well-Being*, Toronto, University of Toronto Press, 1-46.

Leckachman, R. 1971, "The Income Accounts of Tomorrow," *Survey of Current Business*, July, U.S. Department of Commerce, Washington, 119-123.

Little, I.M.D. 1957, *A Critique of Welfare Economics*, 2nd. Edition, Oxford, Clarendon Press.

Nozick, R.E., 1971, *Anarchy, State and Utopia*, New York, Basic Books.

Nordhaus, W., and Tobin, J. 1972, "Is Growth Obsolete?" *Economic Growth*, Anniversary Colloquium V, New York, National Bureau of Economic Research, 1972, 1-92.

Okun, A.M. 1971, "Social Welfare Has No Price Tag," *Survey of Current Business*, July, U.S. Department of Commerce, Washington, 129-133.

Osberg, L. 1985, "The Measurement of Economic Well-Being," in Laidler (ed), 1985, Approaches to Economic Well-Being, Toronto, University of Toronto Press, 49-87.

Rawls, J. 1971, *A Theory of Justice*, Cambridge, Harvard University Press.

Samuelson, P.A. 1950, "Evaluation of Real National Income," *Oxford Economic Papers*, N.S., January, 1-29.

Studenski, P. 1958, *The Income of Nations: Theory, Measurement, and Analysis: Past and Present*, New York, New York University Press.