

---

*Etceteras . . .*

## Gross Domestic Product—an Index of Economic Welfare or a Meaningless Metric?

ROBERT HIGGS

The blind transfer of the striving for quantitative measurements to a field in which the specific conditions are not present which give it its basic importance in the natural sciences . . . not only leads frequently to the selection for study of the most irrelevant aspects of the phenomena because they happen to be measurable, but also to “measurements” and assignments of numerical values which are absolutely meaningless.

—F. A. Hayek, *The Counter-Revolution of Science*

It is almost impossible to imagine the development of economics since World War II apart from the availability and pervasive employment of the national income and product accounts (NIPA). Before the war, however, such measures remained strictly in the development stage, as economists such as Colin Clark and Simon Kuznets worked for years to produce estimates of the sort that any elementary economics student or newspaper reader now encounters daily in frequently updated form. Estimates of gross domestic product (GDP), along with its variants and detailed components, became an essential part of economic analysis only when the Great Depression, the emergence of Keynesian macroeconomics, and the Western powers' engagement in World War II brought such estimates to the fore in the late 1930s and early 1940s. After the government undertook the preparation of these metrics for use in its management of the wartime command economy, and Paul A. Samuelson and others soon adopted them for presentation in widely used textbooks, it became common for GDP to serve as the almost universally accepted concept of the level of overall economic performance in a national economy. For professional economists,

---

**Robert Higgs** is founding editor and editor at large for *The Independent Review* and senior fellow in political economy at the Independent Institute.

*The Independent Review*, v. 20, n. 1, Summer 2015, ISSN 1086–1653, Copyright © 2015, pp. 153–157.

the estimated value of real GDP simply *is* aggregate output, the main focus of macroeconomics insofar as that branch of economics involves empirical description and analysis.

When Kuznets and others were engaged in pioneering the development of the national accounts, substantial controversy arose regarding a variety of related issues. How should the national product be defined? What criteria should be used in deciding what to include and what to exclude? How should final and intermediate outputs be distinguished? How should various forms of output be valued? Of course, an endless array of practical questions also arose about the sources of data that should be employed in making the estimates; how and how often these sources should be updated; how unavailable data should be estimated, projected, and interpolated; and so forth. Some of the world's leading economists, including Kuznets, Samuelson, John Hicks, Richard Stone, Moses Abramovitz, and others, engaged in debating the relevant theoretical and empirical issues. Kuznets in particular argued strongly against the way in which the Commerce Department had answered the various questions related to the NIPA, especially the question of whether government spending should be included in whole or in part in the GDP total and how unsold government "services" should be valued. The Commerce Department's way of constructing the accounts ultimately prevailed, mainly because by the time the war ended, no single analyst or group of private analysts had the resources to prevail against the huge budget and enormous staff of clerks and statisticians the government could now assign to the job.

Among the many issues involved in these developments, none loomed larger than that of the meaning and purpose of national income and product. Analysts eventually settled, for example, on the definition of GDP as the total value at market prices of all final goods and services produced within a nation's boundaries, usually in a year, although quarterly estimates eventually became routine as well. Once such a number has been produced, however, what does it mean? In particular, is it in some sense a measure of economic welfare? Kuznets, despite many practical compromises in details, had generally sought a measure that in some definite sense could be called a measure of welfare. The parallel developments associated with Richard Stone and others in the United Kingdom, in contrast, forthrightly rejected this interpretation, yet, having done so, they still had to face the question about any given estimate of GDP: What does it mean? Is it simply a hodgepodge of numbers reduced to a single number that in some rough, vague way can be employed to compare one year's economic performance with another year's or to compare one country's performance with another's? When Kuznets's approach lost out to the Commerce Department's, this question was not so much settled as brushed under the rug. Rather than attempt to wrestle with the issue until reaching a definite resolution, economists seemingly tired of debating it and decided simply to forge ahead without worrying much about the meaning of the estimates with which they were working. After all, as they viewed the matter, pressing questions of macroeconomic policy demanded answers, and any

particular way of computing GDP would yield a result that can be treated as if it were the economy's aggregate output—whatever that might mean—for purposes of economic policy making and planning.

To the extent that the welfare-economics debate over national income and product had reached an answer, it ran along these lines. One cannot directly add up in a meaningful way the apples, oranges, and countless other goods and services produced in any actual economy; the physical units are incommensurable, so if one were to add them, the resulting sum would be senseless. One can add their values, however, provided that one has measures of their relative values to use as weights. Market prices appeal to most economists as a proper weighting instrument. So, for example, an economy that produces nothing but 20 apples, each valued in the market at \$0.50, and 30 oranges, each valued in the market at \$1.00, has a GDP of \$40. But what underlying logic justifies the use of market prices as the weighting instrument? Why is a market-price-weighted aggregate any more meaningful than an aggregate constructed by using random numbers instead of market prices as the weights?

The answer supplied by economic theory is that *if* the economy is in full competitive equilibrium, the margins of production and consumption for every pair of goods will have been adjusted so as to bring about simultaneously the threefold equality of (1) the two goods' relative market prices, (2) the marginal rate of substitution in consumption for each consumer of these two goods, and (3) the marginal rate of technical substitution in production for each producer of these two goods. In short, price equals relative valuation equals marginal cost across the board. Observing a good's prevailing relative price, we are also observing its relative valuation—that is, how many apples people are willing to trade for an orange at the existing margins of production and consumption of apples and oranges. Thus, in my example where each orange has a market price twice as great as each apple, each consumer of these two goods has adjusted his consumption so that, given the amounts of each that he consumes, he is willing to exchange one orange for two apples and vice versa; and each producer of these two goods has adjusted his production so that, given the amounts of each that he produces, he can produce an additional orange at the opportunity cost of two apples and vice versa. In view of these equalities, it makes economic sense to weight each orange twice as heavily as each apple when we construct the value aggregate we call “gross domestic product.”

Given such equalities, we may fairly say that the GDP we compute by using observed market prices to weight the national product's component goods and services is indeed a measure of economic welfare. As Abramovitz observed more than half a century ago, national product “is taken to be the objective, measurable counterpart of economic welfare, and that is why sustained change in national product is commonly accepted as the basic index of economic growth” (1959, 3). Of course, no actual economy ever settles—indeed, it seems beyond the realm of possibility that any actual economy *can* settle—into a full competitive equilibrium. So, despite the

appealing theoretical basis for using market prices as weights, the meaning we attach to such a procedure holds up only to the extent that actual economic conditions approximate the general competitive equilibrium *sufficiently* closely. In practice, judgments about how close is close enough seem to be entirely arbitrary and absent any form of theoretical or empirical justification. In any event, no more justifiable weight than the actual market prices seems to exist, notwithstanding the gulf that may exist between actual market conditions and full competitive equilibrium conditions.

These matters were brought to my attention most recently as I read Diane Coyle's book *GDP: A Brief but Affectionate History* (2014), a rather sketchy survey of national income and product accounting aimed at lay readers rather than at experts in the field (of whom, truth be told, there are precious few). One point that Coyle makes repeatedly in her survey is that "GDP is not a measure of welfare. . . . GDP measures output; it does not measure well-being" (40; see also pp. 91, 113, and 136). I grant that Coyle may be correctly rendering here the understanding of nearly all economists and other analysts who employ NIPA concepts and estimates in their analysis. More than forty years ago, William Nordhaus and James Tobin declared in almost identical language, "GNP is not a measure of economic welfare. . . . [I]t is an index of production, not consumption" (1972, 4).

Yet there is no escaping the question: If GDP does not measure welfare, at least in the economic-theoretic sense I have described here, what is the meaning of what it does measure? It will not do simply to say, as Coyle and others keep insisting, that GDP measures production or "output" because the economy produces outputs of many different, directly incommensurable goods and services, and the meaning of "output" in the aggregate is only as coherent as the means employed to render the components' values comparable and hence subject to meaningful addition. Where actual market prices are used for this purpose, the calculation procedure rests, as already explained, on very shaky ground, given that no actual economy can be said to mimic the theoretical conditions that prevail in full competitive equilibrium.

In reality, matters are even more desperate because for a large set of goods and services included in the GDP, no market prices exist: either the outputs are not traded in price-revealing private-property exchanges (e.g., imputed values of owner-occupied housing), or the outputs are produced by governments and distributed without direct charge to some or all residents of the country. In practice, unsold government services are absurdly valued by the amount the government happens to pay its employees who provide them. A more arbitrary and ill-justified basis for these services' inclusion in the national product can scarcely be imagined. This procedure implies, for example, that if the government simply raises the wages and salaries it pays its employees, their contribution to GDP increases even though the specific "services" they render remain identical to those they rendered previously. Even more preposterous, perhaps, is the raw assumption that all of these services have any value in the first place, given that many of them consist of actions by which the government harasses, harms, or otherwise makes itself obnoxious to the general population.

It thus probably makes more sense to classify government services in general as “anti-output” rather than as “output.” Less stridently, one might exclude them on the grounds that regardless of how valuable they may be, they are intermediate, not final services. In any case, adding the alleged value of government “services” to GDP serves ideological and propaganda purposes far more surely than it serves scientific purposes (Higgs 1998, 150–1520).

In sum, it is not a defensible dodge for analysts to declare that although GDP does not measure welfare, it does measure output, because the latter measurement requires a justification fully as much as the former, and in practice the way that actual market prices and other valuations are assigned to specific outputs scarcely gives rise to the computation of a clearly meaningful measure of aggregate output. If GDP is to make any sense at all, it *must* do so in relation to some concept of economic welfare. Economics is not a science of the production of hammers and nails but of wealth—the goods and services that people actually value, as they demonstrate by voluntarily forgoing alternative goods and services in the process of acquiring those they choose to acquire. It is unfortunate that once we recognize this reality, we are left to conclude only that, notwithstanding the disclaimers of nearly all current economists and other analysts who use the NIPA concepts and estimates, GDP is indeed a measure of economic welfare. It is simply an exceedingly poor such measure. In any event, it is difficult to believe that this statistical measure is the sort of raw material with which a defensible science can be conducted. As one looks upon what passes for empirical analysis in macroeconomics, the first impression that comes to mind is not the loveliness of GDP, but the ugliness of GIGO—garbage in, garbage out.

## References

- Abramovitz, Moses. 1959. The Welfare Interpretation of Secular Trends in National Income and Product. In Moses Abramovitz and others, *The Allocation of Economic Resources*, 1–22. Stanford, Calif.: Stanford University Press.
- Coyle, Diane. 2014. *GDP: A Brief but Affectionate History*. Princeton, N.J.: Princeton University Press.
- Hayek, F. A. 1952. *The Counter-Revolution of Science: Studies on the Abuse of Reason*. New York: Free Press.
- Higgs, Robert. 1998. Official Economic Statistics: The Emperor’s Clothes Are Dirty. *The Independent Review* 3 (Summer): 147–53.
- Nordhaus, William, and James Tobin. 1972. Is Growth Obsolete? In National Bureau of Economic Research, *Economic Growth: Fiftieth Anniversary Colloquium V*, 1–80. New York: Columbia University Press.

# SUBSCRIBE NOW AND RECEIVE A FREE BOOK!



“*The Independent Review* does not accept pronouncements of government officials nor the conventional wisdom at face value.”

—**JOHN R. MACARTHUR**, Publisher, *Harper’s*

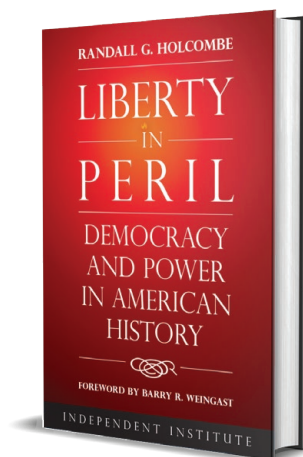
“*The Independent Review* is excellent.”

—**GARY BECKER**, Nobel Laureate in Economic Sciences

Subscribe to [The Independent Review](#) and receive a free book of your choice such as *Liberty in Peril: Democracy and Power in American History*, by Randall G. Holcombe.

Thought-provoking and educational, [The Independent Review](#) is blazing the way toward informed debate. This quarterly journal offers leading-edge insights on today’s most critical issues in economics, healthcare, education, the environment, energy, defense, law, history, political science, philosophy, and sociology.

Student? Educator? Journalist? Business or civic leader? Engaged citizen? This journal is for YOU!



Order today for more **FREE** book options

**SUBSCRIBE**

*The Independent Review* is now available digitally on mobile devices and tablets via the Apple/Android App Stores and Magzter. Subscriptions and single issues start at \$2.99. [Learn More.](#)

