

NBER WORKING PAPER SERIES

UNDERESTIMATING THE REAL GROWTH OF GDP, PERSONAL INCOME AND
PRODUCTIVITY

Martin S. Feldstein

Working Paper 23306
<http://www.nber.org/papers/w23306>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2017

The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2017 by Martin S. Feldstein. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Underestimating the Real Growth of GDP, Personal Income and Productivity
Martin S. Feldstein
NBER Working Paper No. 23306
March 2017
JEL No. E01

ABSTRACT

The problems involved in estimating real output that I discuss in this paper cause the official government statistics to underestimate of the rates of growth of real GDP, real personal income, and productivity. That underestimation is important not just to economists trying to understand where the economy is going but also to the broader public and to the political system.

The understatement of real growth reflects the enormous difficulty of dealing with quality change and the even greater difficulty of measuring the value created by the introduction of new goods and services. Despite the vast amount of attention that has been devoted to this subject in the economic literature and by the government agencies, there remains insufficient understanding of just how imperfect the official estimates actually are. It is important for economists to recognize the limits of our knowledge and to adjust public statements and policies to what we can know.

This paper is not about the recent slowdown in measured productivity but that subject is discussed briefly.

Martin S. Feldstein
President Emeritus
NBER
1050 Massachusetts Avenue
Cambridge, MA 02138-5398
msfeldst@nber.org

Underestimating the Real Growth of GDP, Personal Income, and Productivity

Martin Feldstein*

It ain't what you don't know that gets you into trouble. It's what you know for sure that just ain't so.

Mark Twain

The problems involved in estimating real output that I discuss in this paper cause the official government statistics to underestimate of the rates of growth of real GDP, real personal income, and productivity. That underestimation is important not just to economists trying to understand where the economy is going but also to the broader public and to the political system.

The official figures tell us that real GDP per capita grew at an average rate of just 1.4 percent during the past 20 years. The Economic Report of the President for 2015 (page xxx) stated that, because of changes in the distribution of income, the real income of the median household did not rise at all between 1995 and 2013.

These figures, widely reported in the press and referred to by politicians of both parties, shape the public's perception of the economy's performance. I am struck by the difference between how people judge their own economic condition and their view of the economy as a whole. In the most recent Federal Reserve survey of household attitudes, two-thirds of households reported that they were doing as well or better than they had been five years earlier and that they were either "living comfortably" or "doing ok." (Federal Reserve, 2014). But when asked how the economy is doing, a majority of respondents say the country is doing badly.¹

*Professor of Economics, Harvard University. I am grateful for extensive help with this paper to Anna Stansbury and for comments on earlier drafts from Katherine Abraham, Erica Groshen, Jim Stock, David Weinstein, members of the staffs of the Council of Economic Advisers, the Bureau of Labor Statistics (especially David Friedman) and the Bureau of Economic Analysis, as well participants at a meeting of the Group of Thirty and at a Harvard seminar. A revised version of this paper will appear in the Journal of Economic Perspectives

¹ For example, a majority of respondents to a CNN-ORC poll say that the country is "doing poorly" (CNN international 1/28/2016) or, as CNN summarized

The contrast is important. While people know something about their personal experience, they depend on the official statistics to judge how the economy as a whole is doing. And while the government statisticians are careful to say that GDP doesn't measure how well we are doing, there is the temptation on the part of the press, the politicians, and the public to think that it measures changes in the real standard of living.

The resulting widespread references to slow economic growth reduce the public's faith in the political and economic system. The low measured growth of incomes exacerbates concerns about mobility with people worrying that they and their children are "stuck" at low income levels.² I think it creates a pessimism that contributes to political attitudes that are against free trade and critical of our market economy more generally.

The underestimation of growth also distorts Federal Reserve policy. The perception of slow real growth now supports a Federal Reserve policy of exceptionally low interest rates that is contributing to potential financial instability. Back in 1996, Fed chairman Alan Greenspan persuaded members of the Federal Open Market Committee in 1996 that the official data underestimated productivity growth so that maintaining strong demand would not cause a rise in inflation (Mallaby, 2016).

Although the officially measured rates of output growth have slowed substantially in recent years, the problem of understating real economic growth is not a new one.³ It reflects the enormous difficulty of dealing with quality change and the even greater difficulty of measuring the value created by the introduction of new goods and services. This paper is not about the recent slowdown but I return briefly to that issue in section 4 of this paper.

The government's calculation of real GDP growth begins with the estimation of nominal GDP, i.e. the market value of the millions of goods and services sold in the market to households, firms, governments and foreign buyers. The government

their poll results, "Americans give the economy a "C" grade."(CNN international 4/18/2016)

² A typical example of this is a CNN/ORC poll in which 56 percent of respondents said they think their children will be worse off than they are. (CNN International January 28, 2016)

³ See Fernald, et. al. (2014), Syverson (2016) and Byrne et al (2016) for a discussion of the recent productivity slowdown. They provide convincing analysis that the recent slowdown cannot be explained by the fact that Google and a few other things are omitted from GDP because they are not sold at market prices but are provided without charge and supported by the advertising revenue.

statisticians do a remarkable and prodigious job of collecting data from all of these sources.⁴

But then comes the difficult part of converting the nominal GDP to real GDP. That requires creating an appropriate price index with which to divide the rise in nominal quantities into a real component and an inflation component. The overall GDP price deflator uses components based on the consumer price index (CPI) and the producer price index (PPI), requiring estimates done by the Bureau of Labor Statistics of the Department of Labor (BLS) and by the Bureau of Economic Analysis of the Commerce Department (BEA).

Although there is widespread recognition among economists that measuring changes in the real output of the economy is difficult because of the introduction of new goods and services and of changes in the quality of existing goods and services⁵, there is little understanding of just how limited the existing official estimates actually are. After studying the methods used by the U.S. government statistical agencies as well as the extensive previous academic literature on this subject, I have concluded that, despite the various improvements to statistical methods that have been made through the years, the official data do not adequately measure changes of real output and productivity. The measurement problem has become increasingly difficult with the rising share of services that has grown from about 50 percent of private sector GDP in 1950 to about 70 percent of private GDP

⁴ See “Summary of NIPA Methodologies,” Survey of Current Business,” November 2015, pp 1- 20 for a detailed analysis of the sources used to estimate these purchases. Landefeld et. al (2008) provide a very useful description of how nominal GDP and related measures are estimated from a variety of primary sources. Boskin (2000) shows that these estimates are subject to substantial revisions, with nearly all revisions from 1959 to 1998 in the upward direction and some quite large. As I discuss in the final section of this paper, there are important uses of nominal GDP that do not require conversion to real GDP.

⁵ More than fifty years ago the Stigler Commission (1961) wrote: “If a poll were taken of economists and statisticians, in all probability they would designate (and by a wide margin) the failure of price indexes to take full account of quality changes as the most important defect of these indexes.” There is a vast literature that comments on the effect of product innovation on the difficulty of measuring real output, reaching back to Sidgwick (1883), Marshall (1887), Kuznets (1934, 1941) and including among others Griliches (1992), the Boskin Commission (1996), Nordhaus (1997), Hausman (1996, 1999, 2003) and Gordon (2016). The NBER Conference on Research in Income and Wealth has focused attention on this issue for more than eighty years; see Hulten (2015).

now. The official measures provide at best a lower bound on the true real growth rate with no indication of the size of the underestimation.⁶

I begin this essay by discussing the age-old question of why national income should not be considered a measure of wellbeing. Section 2 describes what the government statisticians actually do to in an attempt to measure improvements in the quality of goods and services. Section 3 considers the problem of new products and the various attempts by economists to take new products into account in measuring overall price and output changes. The measurement of productivity and the recent slowdown in the rate of productivity growth are the subjects of section 4. The final section of this paper discusses how the mismeasurement of real output and of prices might be taken into account in considering various questions of economic policy.

1. Measuring Output or Wellbeing?

There is a long-running debate about whether national income estimates should be designed to measure the wellbeing of the population or just of the output of the economy. Although there is a strong temptation among economists and the general public to interpret real GDP or real personal income as a measure of wellbeing, it is clear that these concepts have been intentionally defined in ways that fall far short of measuring even economic wellbeing, let alone the broader wellbeing of individuals influenced by things like the environment and crime.

Ever since the initial work of Kuznets (1934, 1941), national output has been defined to exclude goods and services produced within the home⁷ as well as services that are provided outside the home but not sold.⁸ This has probably become a larger omission over the years with the provision of such services as

⁶ Coyle (2014, page 125) concludes her useful history of GDP by saying that “Gross domestic product is a measure of the economy best suited to an earlier era.” Bean (2016), especially chapters 2 and 3, offers a variety of thoughtful comments about the problems of measuring GDP and related aggregates in the United Kingdom.

⁷ An earlier National Bureau of Economic Research study by Wesley Mitchell and others (1921) offered a “conjectural value of housewives services” equal to about 30 percent of their estimate of the more narrowly defined traditional national income. A more recent estimate can be found in Franzis et al (2011).

⁸ An exception to the rule that only goods and services sold in the market are to be counted in GDP is the inclusion of government services. Since those services are valued in the GDP at their cost, there is no possibility of reflecting changes in government productivity or the value created by the introduction of new government services.

Google and Facebook and the vast expansion of public radio and television programs, all of which are available to users without payment.⁹

Similarly, national income estimates focus on the positive value of the goods and services that households consume and not on the time and effort involved in earning the funds to buy those goods and services. The average workweek has declined but the number of two-earner households has increased. Working conditions have improved as employment has moved from factories and farms to offices. All of this affects economic wellbeing but there is by agreement no attempt to take it into account in our measures of national income.

I mention all of these issues up front not to criticize the official definition of national income but to stress that it is intended by design to be a measure of national output and not as a measure of wellbeing. In practice, however, the official measure of real GDP does not measure national output. As I have already noted and will explain in sections 2, the official measures of changes in output would be more accurately described as changes in input quantities. That stands in contrast to the basic idea that any good or service that is sold to consumers should be valued by the consumers' willingness to pay for that good or service and not by its cost of production.

More generally, however, the public clearly wants a description of changes in wellbeing and inappropriately uses the official measures of real GDP and real personal incomes for that purpose. It might be good to develop a formal array of such well-being indicators and perhaps some form of a summary index. These indicators might include things like the health of the population, the climate in cities, and other things that are not produced in the market economy.¹⁰ Alternatively, more attention might be focused on the Federal Reserve's Survey of the Economic Wellbeing of U.S. Households and its frequency might be increased from an annual survey to quarterly to increase its public saliency.

2. Measuring Quality Change

For each good and service, there are three possibilities when one compares on year with the next: (1) It is the same good or service with the same quality as in the previous period. (2) It is essentially the same good but of a different quality. (3) It is a wholly new good. Each receives a different treatment in the BLS-BEA analysis.

⁹ On the general problems of reflecting the digital economy in national income, including such things as airbnb and uber, see the extensive discussion in Bean (2016) chapter 3.

¹⁰ For a discussion of previous attempts to provide such additional indicators, see Coyle (2014) chapter five.

Fortunately, most goods and services fall in the first category of “no change in quality.” For those products, it is possible to collect the number of physical units sold and the total revenue. The percentage increase in revenue in excess of the percentage increase in physical volume is pure inflation and the rest is the rise in real output. When exactly the same good is not available in the second period, the BLS tries to find a very similar good that does exist in the two successive periods and compares the revenue growth and physical quantity growth for that good. The BLS refers to this procedure as the “matched model method.” (ref?)

Although much of the growth in the real value of economic output reflects quality change and the introduction of wholly new products, the official procedures do not adequately reflect these sources of increased value. For products that experience quality change, the official methods tell us more about the increase in the value of inputs, in other words about the change in the cost of production, and not much about the increased value to the consumer or other ultimate user. And this is true for goods as well as for services, although doing it for services is even more difficult than it is for goods.

The government statisticians divide the period-to-period increase in total spending on each unit of product into a part due to a pure price increase (“inflation”) and a part due to an increase in quality. The part due to the quality increase is considered an increase in the quantity of output although, as I will explain, it is generally a measure of the increase in the quantity of inputs.

The BLS is responsible for creating the producer price indexes and the consumer price indexes.¹¹The BEA uses those indexes and other data to create the measures of real output.¹²

The key question is how the BLS estimates the change in price when there is a change in the quality of the good or service. Hedonic regressions are used for about one-third of the goods and services in the consumer price index, particularly

¹¹ For the best description of the BLS methods, see [BLS Handbook of Methods](#), chapter 14 for the PPI indexes (available at <http://www.bls.gov/opub/hom/pdf/homch14.pdf>) and chapter 17 for the CPI indexes (available at <http://www.bls.gov/opub/hom/pdf/homch17.pdf>.) See also “Summary of NIPA Methodologies,” Survey of Current Business,” November 2015, pp 1- 20. In contrast to the various CPI component indexes, many of the PPI indexes are used primarily to deflate the prices of intermediate products rather than to deflate output for final demand.

¹² These estimates are also used for measuring the output of the nonfarm business sector and are used by the Department of Commerce to calculate the GDP deflator and real GDP. These figures are also used to calculate the deflator for consumer expenditures that the Federal Reserve uses for its price stability target.

for housing.¹³ When the hedonic regression is not used, the BLS uses information about the cost of production supplied by the producer. The BLS calls this method of “quality adjustment” the “resource cost method.”

For the “resource cost method” the BLS asks the producer of each good or service whether there has been a change in the product made by that producer.¹⁴ If there has been a change in the product, the BLS asks about the “marginal cost of new input requirements that are directly tied to changes in product quality.” The rationale for this input cost method for defining the “quality adjustment” or, equivalently, the measure of the increased output, is described in Triplett (1982) when Triplett was the Assistant Commissioner of the BLS.

According to the resource cost method, the BLS concludes that there has been a quality improvement if and only if there is such an increase in the cost of making the product or service.

The government statisticians then use the marginal cost of the quality improvement, measured as a percentage of the initial cost of the product, to calculate the share of the price rise that is due to a quality improvement and that is therefore deemed to be an increase in the output of the product.¹⁵ The rest is regarded as pure inflation.

This is a remarkably narrow and faulty definition of product or service improvement. It explicitly implies that if it doesn’t cost more to produce a product or service this year than it did last year, there has been no improvement. According to this method, a pure technical innovation that makes the product or service better for the consumer doesn’t count as a product improvement unless it involves an increased cost of production!

In reality, product improvements generally occur because of new ideas about how to redesign or modify an existing product or service. Those changes need not involve an increased cost of production. The resource cost method can also provide a way to deal with any decline in the quality of goods and services, treating a decline of production cost as evidence of a decline in quality.

When the government statisticians deal with services, they generally focus on measuring the prices of inputs. The change in the “real output” of any narrowly distinguished type of service is defined by the BEA by dividing the total expenditure

¹³ I return below to the problem of interpreting the hedonic regression for this purpose.

¹⁴ If there has been no change in the product, any change in its price is considered to be pure inflation as called for in the “matched model method.”

¹⁵ As some critics have noted, this method of valuing output is similar to the old-fashioned labor theory of value.

on that service by an input price index. The true output of the service that consumers actually value is ignored.

An important example is the health care industry which is responsible for more than 15 percent of U.S. GDP. The official GDP statistics for the health care industry focus on costs, ignoring the effect of the health products and services on the health of the patient.

For example, the “output” measure for hospitals has recently shifted from a day of inpatient care to an episode of hospital treatment for a particular condition. Changes in the cost per episode of treatment is used to measure the quality change. Triplett (2012, p 17), a careful analyst of the statistical health debate, concluded that there is a “very large error in measuring output generated in the medical care sector.”

More generally, as Triplett and Bosworth (2004) note, the official data imply that productivity in the health industry (i.e., the ratio of output to the number of employee hours involved in production) declined year after year between 1987 and 2001. They conclude (page 265) that such a decline in true productivity is unlikely but that officially measured productivity declines because “the traditional price index procedures for handling product and service improvements do not work for most medical improvements.” More recent data show that health sector productivity has been declining since 2001. The key point is that none of these measure of productivity attempt to value the improved patient outcomes.

When Triplett and Bosworth (2004) wrote about the remarkable improvement in treating cataracts – from more than a week as an immobilized hospital inpatient to a quick outpatient procedure -- they questioned whether reflecting medical improvements like that would cross over the traditional “production boundary in national accounts” and asked whether “the increased value to the patient of improvement in surgery ... belongs in national accounts if no additional charges are made.” (page 335)¹⁶

For another example of a service, consider mutual fund management. The BEA notes that there has been a substantial expansion over time in the types of funds that are available (including ETFs, fund-of-funds, long-short funds, a large number of emerging market funds, etc.) but ignores this increase in diversity of products and focuses only on the charges based on a percentage of all assets, concluding that “Under the current methodology, no special procedures are

¹⁶ The Department of Commerce is experimenting with health sector “satellite accounts” that calculate the cost of treating a patient with a particular diagnosis for a calendar year, including the cost of hospital care, of physicians and pharmaceuticals. Some other studies attempt to measure the effect of the treatment on such health outcomes as Quality Adjusted Life Years (QALYs) or Disability Adjusted Life years (DALYs).

necessary for adjusting for the changes in the quality of portfolio management transactions.” (NAICS 523920, page 13)

The treatment of internet services is yet another example of the misleading nature of the official procedure. The official estimates of real personal consumption expenditures on internet access are reported to have declined between 2011 and 2012 (reference ?? -- FRED) despite the 45 percent increase in the number of smart phones between those two years, the 75 percent rise in the number of active Apple apps (to 625,000), and a more than doubling of the amount of data transmitted via mobile connections. Fortunately, since 2012 the official estimate of real personal consumption expenditures on internet access has been corrected to show a positive increase in previous years.

Although the “resource cost method” may be the most common approach, it is not the only one. For a several industries, the BLS measures real output at the industry level by the physical quantity of services provided. These physical quantity measures are used for measuring the growth of output and productivity for these individual industries.¹⁷

For example, for passenger air travel, output of the industry is simply the number of passenger miles and productivity is defined as just passenger miles per employee hour. The analysis of output “does not account for changes in service quality such as flight delays and route circuitry...” (Duke and Torres, 2005)

From time to time the BLS staff reexamines its approach to a particular industry. When the productivity program re-examined its measure of the commercial banking industry in 2012, it revised the activities of commercial banks and raised the estimated annual output growth from 1987 to 2010 by 58 percent, from 2.4 percent a year to 3.8 percent a year (BLS Monthly Labor Review, July 2012, page 7).

A third approach to measuring quality change is the “hedonic regression” method originally developed by Griliches (1961). The basic idea, which was used extensively for personal computers, is to regress the prices of computers in year t on a variety of the computers’ capacity and performance measures. This gives an implicit price for each of these features (if the linearity assumption of the model is correct). Applying these features to a computer model in year $t+1$ generates a price (say $p^*(t+1)$) that would apply for that computer if the values of the individual features at time t had continued to prevail. Since the actual price at time $t+1$ is less than $p^*(t+1)$, the price decline can be used to value the quality change.

¹⁷ The BEA also uses what they call the “quantity extrapolation method” and the “direct valuation method” for a few types of output. For example, the real quantity of bank services is derived from volume data on consumers’ deposits and loans. See Survey of Current Business, November 2015.

A variety of econometric studies (Chow, 1967; Baily and Gordon, 1989; Triplett, 1989) showed that the true price of mainframe computers assessed in this way declined at an annual rate of more than 20 percent per year during the period from 1950 to 1980. Berndt et al (1995) found a 28 percent annual rate of decline for personal computers during a more recent period. Hedonic regressions may have produced the answer to Robert Solow's comment in 1987 that computers are seen everywhere except in the productivity statistics.

The Bureau of Economic Analysis of the Commerce Department uses hedonic price indexes to deflate nominal output for about 20 percent of GDP. This includes such things as apparel, software, and housing. The use of hedonics is no doubt very difficult to apply for these products and services for which, unlike computers, there is not a clear list of measured technical product attributes. There is also a problem of assuming that the attributes affect willingness to pay in a linear or log-linear way. According to the government, extensions of hedonics to even more products and services is limited by the lack of detailed data and staff resources required to build and maintain the hedonic models.¹⁸

My own judgment is that, for most goods and services, the official estimate of quality change contains very little information about the value of the output to consumers and other final purchasers. That implies that the corresponding official measures of total real output are underestimates and that there is a substantial but unknown upward bias in the measure of price inflation. We don't know what the true values are and we don't know how wide a margin of error there is around the official estimates.

As the methods that I have described indicate, the official estimates of "real output" for many of the goods and services may be closer to the value of the "real inputs", i.e., essentially the value of labor costs deflated by wage rates, rather than as measures of the value to consumers.

3. Dealing with New Products

The resource cost method and other government procedures for valuing changes in quality do not provide an approach to dealing with the value to consumers of new goods and services. Instead, the sales of new products become part of nominal GDP but the extra value to consumers created by the innovation is not reflected in the output measure or the price index. New products and services are not reflected in the price indexes until they represent a significant level of expenditures. They are then introduced into the price index calculations and subsequent changes in their price are taken into account in the usual way. It is only in that way that the new product affects changes in real output.

¹⁸ See also Hausman (2003) for a discussion of the limitations of hedonic pricing.

An example will clarify how this works in practice. Consider statins, the remarkable class of drugs that lowers cholesterol and reduces deaths from heart attacks. By 2003, statins were the best-selling pharmaceutical product in history. By then it was in the CPI price index. When patents on early versions of statins expired and generic forms became available, their prices fell and the BLS recorded those price declines, implying a rise in real incomes. But the BLS never estimated any value for the improvement in health that came about as a result of the introduction of statins.

To see the magnitude of omitting the value of that single health care innovation here is a quick history of the impact of the statins. In 1994, researchers published a five-year study of 4000-plus patients. They found that taking a statin caused a 35 percent reduction in cholesterol and a 42 percent reduction in the probability of dying of a heart attack. It didn't take long for statins to become a best-selling product with dramatic effects on cholesterol and heart attacks. According to the National Institute of Health (ref?), between 2000 and 2007, the percentage of men 65 and older taking a statin doubled to about 50 percent of men over age 65. High cholesterol levels declined by more than half among men and women over age 75 and the death rate from heart disease among those over 65 fell by one third.

Using the estimated statistical value of a human life (Schelling, 1968; Viscusi and Aldy, 2003), Grabowski et al (2012) calculated that the combination of reduced mortality and lower hospital costs associated with heart attacks and strokes in the year 2008 alone was some \$400 billion. None of this value is included in the government's estimate of increased real income.

This example of how statins have been treated in the national income statistics indicates how all new products and services are treated. The value of to consumers of a new good or service is ignored when the new product is introduced. Its price is added to the consumer price index when spending on that good or service is large enough to warrant inclusion. Subsequent declines in the price of the product are treated as real income gains while price increases are part of inflationary real income losses. In short, the value to the consumer of the innovation is completely ignored.

Although there has been a substantial amount of interesting academic research on the valuation of new products, this work has produced no practical measure for reflecting the value of new goods and services in the measures of real income and output.

Ignoring the introduction of new products is therefore a serious further source of understating the real growth of output, incomes, and productivity. New products and services are potentially valuable in themselves and are also valued by consumers because they add to the variety of available options. In an economy in which new goods and services are continually created, their omission in the current method of valuing aggregate real output makes the existing measure of real output

even more deficient and more of a continually increasing underestimate of true output. Hulten (2015) summarizes decades of research on dealing with new products done by the Conference on Research in Income and Wealth with the conclusion that “the current practice for incorporating new goods are complicated but may miss much of the value of these innovations.” (page 2)

The introduction of new products into the official price indexes has historically also been subject to remarkably long delays. The Boskin Commission noted that at the time of their report in 1996 there were 36 million cellular phones in the United States but their existence had not yet been included in the CPI. Decade long delays were also noted for things like room air conditioners. Autos were only introduced to the CPI in 1940 and refrigerators in 1934 (Stigler Commission, 1961). More recently, the BLS has introduced new products more quickly but only after they have achieved substantial scale in spending. These delays cause the price index to miss the declines in prices that happen early in product cycles.

But these delays in the introduction of new products to the price indexes are not the key problem. Much more important is the fact that the official statistics ignore the very substantial direct benefit to consumers when new products and services become available, causing an underestimate of the rate of increase in real output and an overestimate of the corresponding price index.

Jerry Hausman (1996, 2003) showed how the value to consumers of a single new product could be measured by estimating the value of introducing a single new brand of breakfast cereal (Apple-Cinnamon Cheerios). His approach, following the theory presented by Hicks (1940), was to estimate the “virtual price,” i.e., the price that would prevail when the good is just introduced at zero quantity and the increased consumer surplus when the new good is introduced. The consumer gains an amount of real income when the good is introduced implied by the decline in its price from the virtual price to the actual market price. He concluded that the CPI component for cereals may be overstated by about 20 percent because of its neglect of new cereal brands.

The Hausman procedure is not a practical solution to the general problem of valuing the vast array of new goods and particularly new services that are created each year. But if the Hausman estimate is even roughly indicative of the overstatement of the CPI because of the failure to reflect the introduction of new varieties of cereal brands, think how much more is the overstatement of the CPI and the understatement of real income that result from failing to take into account new products like antibiotics and the drugs that now successfully treat leukemia and other cancers.

Broda and Weinstein (2010) and Redding and Weinstein (2016) extend the Hausman approach and present a new method to value new products as well as the value to consumers of changes in product quality. They analyze a very large set of data on bar-coded package goods for which prices and quantities are available over

time. By studying these data in the framework of a demand system based on constant elasticity of substitution utility functions, they find that the conventional price indexes overstate inflation by as much as five percentage points because the conventional measure ignores quality and new goods biases. While this method is very interesting, its use is limited to goods and services for which a large amount of bar coded price and quantity data are available. It also requires accepting a specific theoretical demand specification like the constant elasticity of substitution for all of the relevant products. It is hard to see how it could be extended to most services where comparable price and quantity information is not available and where the substitution assumptions may not hold.

The Boskin Commission (1996) was charged by the U.S. Senate with calculating the bias in the consumer price index that was used for adjusting Social Security benefits for changes in the cost of living. The Commission considered several sources of bias in the existing CPI including the bias caused by changes in quality and by the omission of new products. The Commission provided estimates of each type of bias in the CPI.

Since the Commission was not able to do new research on the issue of quality change and innovation bias, they drew on existing research and on their personal perceptions. For example, for “food and beverage,” which accounts for 15 percent of the CPI, the Commission members asked themselves how much a consumer would be willing to pay “for the privilege of choosing from the variety of items available in today’s supermarket instead of being constrained to the much more limited variety available 30 years ago.” They concluded, based on pure introspection, that “a conservative estimate ... might be 10 percent for food consumed at home other than produce, 20 percent for produce where the increased variety in winter (as well as summer farmers’ markets) has been so notable, and 5 percent for alcoholic beverages...” They used these numbers for 30 years and converted them to annual average rates of change for the 30-year period. This may or may not be plausible but there is no real basis for believing that any of these estimates is even vaguely accurate.

Housing is the most heavily weighted component of the CPI with a weight of 34 percent. The Boskin Commission concluded that “a conservative estimate is that the total increase in apartment quality per square foot, including the rental value of all appliances, central air conditioning, and improved bathroom plumbing, and other amenities amounted to 10 percent over the past 40 years, or 0.25 percent per year.” Maybe that is right and maybe a better estimate would be one percent per year. There is nothing in the Commission’s report to decide.

In the end, the Commission concluded that the weighted average of these individual biases implies a total bias from product innovation and quality change in the annual CPI inflation rate for 1996 of 0.6 percentage points. I have no idea how much margin of error should be attached to that estimate. It served the purpose of

providing a politically acceptable basis for reducing the rate of increase of Social Security benefits.

The creation of new products also means the creation of an increased variety of choice, a possible form of quality improvement in itself, as Hausman (1996) noted. The value to consumers of this access to a variety that allows individuals to make choices that conform to their personal taste can be substantial. Coyle (2015) noted that in the 30 years after 1970 the number of TV channels went from five to 185 and the number of soft drink brands went from 20 to 87.

The failure to take new products into account in a way that reflects their value to consumers may be an even greater distortion in the estimate of real growth than the failure to reflect changes in the quality of goods and services. There is no way to know.

4. Productivity Change and Its Recent Slowdown

Labor productivity is defined as the ratio of real output to the number of hours worked by all employed persons. The BLS estimates labor productivity for the nonfarm business sector as well as for some parts of that sector using output estimates provided by the BEA.¹⁹

The key problem in measuring labor productivity is in the numerator, i.e., in measuring output. The failure to measure quality changes adequately and to incorporate the value of new products means that true output has grown faster than measured output and therefore that the pace of productivity growth has been underestimated.

This is a particularly difficult problem in the service industries. Triplett and Bosworth (2000 p 6 and 2004, p 331) noted that the official data imply that productivity declined in several of the major service industries – including health care, hotels, education, entertainment and recreation – and concluded that this apparent decline was “unlikely” and probably reflected measurement problems. Greenspan also concluded that the unlikely productivity declines in the service industries as officially measured was the reason for the low estimate of overall productivity growth that was being used in 1996 as a rationale to raise interest rates.

¹⁹ Multifactor productivity is the ratio of real output to a combination of labor and capital input services. It is intended to measure the increase in output that is not attributable to either labor inputs or capital inputs. A good deal of research has been devoted to the very difficult problem of measuring the input of capital services and to the correct way to combine labor and capital inputs.

But while the understatement of productivity growth is a chronic problem, what is new is the very sharp decline in the officially measured rate of productivity growth in the past few years. That sharp decline remains a puzzle that is yet to be resolved. The papers by Fernald et. al. (2014), Byrne et. al. (2016) and Syverson (2016) show that the recent productivity slowdown cannot be attributed to the effects of the recession of 2008-09, to changes in the labor force demographics in recent years, or to the growth of unmeasured internet services.

One possible explanation of the recent downturn in productivity growth may be that the unusually rapid *increase* in the rate of productivity growth in the prior few years was an anomaly and the recent productivity decline is just a return to earlier productivity patterns. Of course, this slowdown in the official measure of productivity growth has to be understood in the context of the overall mismeasurement of the official estimates of output and productivity.

Looking more closely shows that the recent decline in the measure of overall labor productivity growth in the nonfarm business sector reflects an enormous diversity of changes of productivity in specific industry groups. For the nonfarm business sector as a whole, the rate of productivity growth fell from 3.2 percent a year in the decade from 1995 to 2004 to just 1.5 percent in the decade from 2004 to 2013.

This decline of 1.7 percentage points in the overall productivity change reflects an enormous range of declines in various industries.²⁰ For example, even if attention is limited to the relatively aggregate 3-digit level, the official productivity data show that productivity in apparel manufacturing went from growth at 1 percent in the earlier decade to a decline at 5 percent in the later period, a drop of 6 percentage points. For manufacturing of computers and electronic products, productivity growth fell from 15 percent to 4 percent, a fall of 11 percentage points. Some industries experienced faster productivity growth, with productivity in the manufacturing of wood products increasing from a 2 percent annual rise in the early period to a 2.4 percent rise in the later period. The differences are even greater at a more disaggregated level. For example, at the four digit level the productivity growth increased by 5 percentage points for radio and TV broadcasting but declined by 18 percent for semiconductors and electronic components.²¹

²⁰ For the real productivity data by industry, see Bureau of Labor Statistics, "Industry Productivity: 1987-2015". The overall productivity measure is not calculated by combining the individual industry numbers but is estimated separately based on a measure of real value added.

²¹ The deflation of output for disaggregated industries is even harder than for the economy as a whole because nominal outputs must be deflated by quality adjusted prices for the disaggregated industries. See Dennison (1989)

Any attempt to explain the recent decline in the estimated productivity growth rate must attempt to understand not just the aggregate behavior for the nonfarm business sector as a whole but also what happened at the disaggregated level. Unfortunately, that is made difficult, perhaps impossibly difficult, by the problems of dealing with product change and the introduction of new goods and services.

5. Using Our Imperfect Data

What can be learned from the imperfect measures of real output and from the corresponding changes overstatement of price inflation? How should our understanding of the mismeasurement affect the making of monetary and fiscal policies?

Assessing Cyclical Economic Conditions

Consider first the assessment of short-run business cycle conditions. Policy makers and financial markets often look at short-term fluctuations of real GDP as an indication of the state of the business cycle. Although measuring fluctuations of real GDP is flawed by the difficulty of dealing with new products and quality changes, the official measure of real GDP fluctuations can in principle capture the relevant short-run changes in the pace of economic activity. To understand the state of the business cycle, analysts can focus on the variations in estimated real economic activity relative to the misestimated underlying trend. It is important however to recognize the substantial uncertainty about the estimated short-run fluctuations in GDP and the subsequent revisions.²²

The Business Cycle Dating Committee of the National Bureau of Economic Research has traditionally looked at employment, industrial production, wholesale-retail sales, and real income.²³ The first three of these measures not only avoid the

²² The Federal Reserve banks of New York and Atlanta have recently begun using official data to produce preliminary estimates of changes in real GDP even before the corresponding quarter is over. As of the time of writing the initial draft of this paper (April 2016) the New York Federal Reserve estimated that real GDP increased by 1.1 percent in the recently completed first quarter of 2016 while the Atlanta Federal Reserve estimated that the increase in the same quarter was only 0.1 percent.

²³ In recent years, the NBER Committee has also looked at monthly GDP when Macro Advisers began creating monthly estimates of GDP. The NBER never used two quarters of decline in real GDP to define a recession.

measurement problems associated with changes in real GDP but deal with economic activity without the large amounts of imputation (like the services of owner occupied homes) that are involved in the estimate of GDP.

All data involve problems of interpretation in judging the state of economic activity, but employment, industrial production, and nominal sales are relatively free from the problem of quality adjustment and price measurement that affect measures of real GDP. Employment data are available monthly with substantial detail based on a large survey of employers. Industrial production is estimated by the Federal Reserve based primarily based on data on physical production (such as tons of steel and barrels of oil) obtained from trade associations and government agencies, supplemented when necessary with data on production-worker hours and for some high tech products by using nominal output and a price index.²⁴

Assessing Longer Term Growth and Inflation

For the longer term, the official measures of changes in real output are misleading because they essentially ignore the value created by the introduction of new goods and services and underestimate changes in the quality of these products. It follows therefore that “true” real output is growing faster than the official estimates imply and that the corresponding “true” GDP price index is rising more slowly than the official one or is actually declining.

These official real growth estimates are clearly false and provide too pessimistic a picture of the path of economic growth and the development of real incomes. The economics profession should educate the general public and the policy officials that “true” real incomes are rising faster than these official data imply. We can also provide reassurance that the real incomes of future generations are very unlikely to be lower than the real incomes of today.²⁵

One can only speculate about whether the bias in the measured pace of real output change is greater now than in the past. One reason to think that the gap between true output growth and measured growth is greater now than in the past is that services are a bigger part of the economy and the degree of underestimation of quality change and product innovation may be greater for services. In 1950, services represented about 50 percent of private value added while today it is about 70 percent. Within services, health occupies a larger share and improvements there may be greater than in other parts of the service sector. The internet and services

²⁴ See Federal Reserve Board release G-17 for details.

²⁵ Even if the future will not see the “epochal innovations” of the type that Kuznets referred to or such fundamental changes as electricity and indoor plumbing that caused jumps in living standards (as emphasized by Gordon, 2016), the current and future generations can continue to experience rising real incomes due to technological changes, improvements in education, and increases in health care.

through the internet which have become much more important are harder to measure.

Trends in the overestimation of inflation and therefore in the underestimation of real incomes may vary among demographic groups and income groups because of differences in the mix of goods and services consumed by these different groups. Are the goods and services bought by older people improving relatively faster than the goods and services bought by younger households? Health care is an obvious example, although most of the consumption of health care is financed by government transfers.

Implications for Fiscal and Monetary Policy

When we turn to the policy implications of the mismeasurement of output, growth, and inflation, an important point is that none of the problems associated with converting nominal output to real output affects the usefulness of the nominal magnitudes. So policy issues that depend on nominal measures are unaffected by the problems discussed in this essay.

The most obvious of these is the ratio of debt to GDP since both the numerator and the denominator are nominal values. Similarly, the rate of change of the debt to GDP ratio depends only on the nominal value of the annual deficit and the annual rate of nominal GDP growth. The rate of increase of the debt to GDP ratio is equal to the difference between the nominal deficit ratio and the product of the nominal GDP growth rate and the debt to GDP ratio. If the debt to GDP ratio is not on an explosive path, its long-run equilibrium value is equal to the annual nominal deficit ratio divided by the rate of nominal GDP growth.

The evidence that the true inflation rate is less than the measured inflation rate may imply that the true inflation rate is now less than zero. Fortunately, this does not imply that the U.S. is experiencing the traditional problem of debt deflation (Fisher, 1933) that occurs when a declining price level reduces aggregate demand by increasing the value of household debt relative to current incomes. The traditional problem of debt deflation does not occur now because the nominal value of monthly wage income is not declining and the real monthly wage is rising more rapidly.

A negative true rate of inflation does imply that the real rate of interest is higher than the conventionally measured rate. If households recognize that their dollars will buy relatively more in the future, this could alter the household saving rate – either increasing saving in response to the greater reward for saving or decreasing saving because a given volume of assets will buy more in the future. Because many things affect the household saving rate, it is not clear which of these effects now dominates.

The uncertainty about the true rate of inflation should affect the optimal monetary policy. There seems little point in having a precise inflation target when the true rate of inflation is measured with a great deal of uncertainty. The goal of price stability also takes on a new meaning if true inflation is substantially negative while measured inflation is low but positive. Would it be better to have a target range for measured inflation? Or to restate the inflation goal of monetary policy as reacting when there is a rapid movement in measured inflation either up or down?

A great deal of effort and talent has been applied over the past decades to improve the measurement of real output and the corresponding price index. The inability to measure either output or inflation with any accuracy shows how difficult are the problems of quality adjustment and the evaluation of new products. It is important to recognize the limits of our knowledge and to adjust public statements and policies to what we can know.

Cambridge, MA
October 2016

References

Baily, Martin, and Robert J. Gordon. 1989. "Measurement Issues, the Productivity Slowdown and the Explosion of Computer Power". CEPR Discussion Papers No. 305.

Bean, Charles. 2016. "Independent Review of UK Economic Statistics".

Berndt, Ernst R., Zvi Griliches, and Neal J. Rappaport. 1995. "Econometric Estimates of Price Indexes for Personal Computers in the 1990s". *Journal of Econometrics*. 68 (1): 243-268.

Bivens, Josh, and Lawrence Mishel. 2015. "Understanding the Historic Divergence Between Productivity and a Typical Worker's Pay: Why It Matters and Why It's Real". *Economic Policy Institute Briefing Paper #406*.

Boskin, Michael J., Ellen R. Dulberger, Zvi Griliches, Robert J. Gordon, and Dale Jorgensen. 1996. "Toward a more accurate measure of the cost of living". Final report to the Senate Finance Committee for the Advisory Commission to Study the Consumer Price Index.

Boskin, Michael J. 2000. "Economic Measurement: Progress and Challenges". *American Economic Review*. 90 (2): 247-252.

Broda, Christian, and David E. Weinstein. 2010. "Product Creation and Destruction: Evidence and Price Implications". *American Economic Review*. 100 (3): 691-723.

Bureau of Economic Analysis. 2015. "Updated Summary of NIPA Methodologies". *Survey of Current Business*. 95 (11): 1-20.

Bureau of Labor Statistics. 1987-2015. "Major Sector Productivity and Costs". United States Department of Labor. <http://download.bls.gov/pub/time.series/pr/>. (accessed April 3, 2016).

Bureau of Labor Statistics. 1987-2015. "Industry Productivity". United States Department of Labor. <http://download.bls.gov/pub/time.series/ip/>. (accessed April 3, 2016).

Bureau of Labor Statistics. "Chapter 14. Producer Prices". In *Handbook of Methods*. <http://www.bls.gov/opub/hom/pdf/homch14.pdf>

Bureau of Labor Statistics. "Chapter 17. The Consumer Price Index". In *Handbook of Methods*. <http://www.bls.gov/opub/hom/pdf/homch17.pdf>

Bureau of Labor Statistics. 2014. "Quality Adjustment in the Producer Price Index". Last modified August 2014. <http://www.bls.gov/ppi/qualityadjust.pdf>

Bureau of Labor Statistics. 2015. "Hedonic Models in the Producer Price Index (PPI)". Last modified October 21 2015. <http://www.bls.gov/ppi/ppicomqa.htm>

Bureau of Labor Statistics. 2015. "Synopsis: NAICS 523920. Portfolio Management". Producer Price Index Program. Revised March 2 2015.

Bureau of Labor Statistics. 2015. "Synopsis: NAICS 541110. Offices of Lawyers". Producer Price Index Program. Revised September 6 2013.

Byrne, David M., John G. Fernald, and Marshall B. Reinsdorf. 2016. "Does the United States have a productivity slowdown or a measurement problem?". *Brookings Papers on Economic Activity Conference Draft*, March 10-11 2016.

Chow, Gregory C. 1967. "Technological change and the demand for computers". *The American Economic Review*. 57 (5): 1117-1130.

Coyle, Diane. 2015. *GDP: A brief but affectionate history*. Princeton: Princeton University Press.

Dennison, Edward. 1989. *Estimates of Productivity Change by Industry: An Evaluation and an Alternative*. Washington, DC: Brookings Institution Press.

Duke, John, and Victor Torres. 2005. "Multifactor Productivity Change in the Air Transportation Industry." *Bureau of Labor Statistics Monthly Labor Review*. March 2005: 32-45.

Federal Reserve Board. 2014. Report on the Economic Wellbeing of U.S. Households in 2013.

Federal Reserve Board. "Industrial Production and Capacity Utilization – G.17". <http://www.federalreserve.gov/releases/g17/current/>

Federal Reserve Bank of Atlanta. "GDP Now". <https://www.frbatlanta.org/cqer/research/gdpnow.aspx?panel=1>. (accessed April 2016).

Federal Reserve Bank of New York. 2016. "Nowcasting Report: April 8, 2016". https://www.newyorkfed.org/medialibrary/media/research/policy/nowcast/nowcast_2016_0408.pdf?la=en

Fernald, John G. 2014. "Productivity and Potential Output Before, During and After the Great Recession". National Bureau of Economic Research Working Paper No. 20248.

Fisher, Irving. 1933. "The Debt-Deflation Theory of Great Depressions". *Econometrica*. 1 (4): 337-357.

Fogel, Robert William, Enid M Fogel, Mark Gulgielmo, and Nathaniel Grotte. 2013. *Political arithmetic: Simon Kuznets and the empirical tradition in economics*. Chicago: University of Chicago Press.

Franzis, Henry and Jay Stewart (2011) "How does nonmarket production affect measured earnings inequality," *Journal of Population Economics* 24(1): 3-22

Gordon, Robert J. 2016. *The Rise and Fall of American Growth: The US Standard of Living since the Civil War*. Princeton: Princeton University Press.

Grabowski, David et. al. "The Large Social Value Resulting from Use of Statins Warrants Steps to Improve Adherence and Broader Treatment," *Health Affairs* 10: 276-85

Griliches, Zvi. 1961. "Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change." In *The Price Statistics of the Federal Government*, General Series 73. New York: National Bureau of Economic Research.

Griliches, Zvi, ed. 1992. *Output Measurement in the Service Sectors*. National Bureau of Economic Research Studies in Income and Wealth. Chicago: University of Chicago Press.

Hausman, Jerry A. 1996. "Valuation of new goods under perfect and imperfect competition." In *The economics of new goods*, ed. Timothy F. Bresnahan and Robert J. Gordon, 207-248. Chicago: University of Chicago Press.

Hausman, Jerry A. 1999. "Cellular telephone, new products, and the CPI". *Journal of business & economic statistics*. 17 (2): 188-194.

Hausman, Jerry A. 2003. "Sources of Bias and Solutions to Bias in the Consumer Price Index," *Journal of Economic perspectives*, 17 (1) 23-44

Hicks, John R. 1940. "The Valuation of the Social Income." *Economica*.7 (26): 105-124.

Hulten, Charles R. 2015. "Measuring the economy of the 21st century". *NBER Reporter* 2015 Number 4: 1-7.

Kuznets, Simon. 1934. *National Income, 1929-1932*. New York: National Bureau of Economic Research.

Kuznets, Simon, Lillian Epstein, and Elizabeth Jenks. 1941. *National Income and Its Composition, 1919-1938, Volume I*. New York: National Bureau of Economic Research.

Kuznets, Simon. 1971. "Modern Economic Growth: Findings and Reflections". Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel, Prize Lecture, December 11 1971.

Landefeld, J. Steven, Eugene P. Seskin, and Barbara L. Fraumeni. 2008. "Taking the Pulse of the Economy: Measuring GDP". *Journal of Economic Perspectives* 22 (2): 193-216.

Mallaby, Sebastian. 2016 *The Man Who Knew*. Bloomsbury

Mandel, Michael. 2012. "Beyond Goods and Services: The (Unmeasured) Rise of the Data-Driven Economy". Progressive Policy Institute Policy Memo, November 2012. <http://www.progressivepolicy.org/wp-content/uploads/2012/10/10.2012-Mandel-Beyond-Goods-and-Services-The-Unmeasured-Rise-of-the-Data-Driven-Economy.pdf>

Marshall, Alfred. 1987. "Remedies for Fluctuations of General Prices," in *Memorials of Alfred Marshall*. A.C. Pigou (ed.) 1887

Mitchell, Wesley Clair, Wilford Isbell King, and Frederick R. Macaulay. 1921. *Income in the United States: Its Amount and Distribution, 1909-1919, Volume 1: Summary*. New York: National Bureau of Economic Research.

Nordhaus, William D. 1997. "Traditional Productivity Estimates Are Asleep at the (Technological) Switch". *The Economic Journal*. 107 (444): 1548-1559.

Redding, Stephen J., and David E. Weinstein. 2016. "A Unified Approach to Estimating Demand and Welfare".
<http://www.columbia.edu/~dew35/Papers/Estimating-Demand-and-Welfare.pdf>

Royster, Sara E. 2012. "Improved Measures of Commercial Banking Output and Productivity". *Bureau of Labor Statistics Monthly Labor Review*. July 2012: 3-12.

Schelling, Thomas. 1968. "The Life You Save May Be Your Own". In *Problems in Public Expenditure Analysis*, ed. Samuel B. Chase, 143-144. Washington, D.C.: Brookings Institution.

Sidgwick, Henry. *The Principles of Political Economy.*, 1883

Solow, Robert. 2015. "The Future of Work: Why Wages Aren't Keeping Up." *Pacific Standard*, August 11 2015. <https://psmag.com/the-future-of-work-why-wages-aren-t-keeping-up-6fcfac468e4#.m7mg6zsz9>

Stigler, George, ed. 1961. "The Price Statistics of the Federal Government." Report to the Office of Statistical Standards, Bureau of the Budget. New York: National Bureau of Economic Research.

Syverson, Chad. 2016. "Challenges to Mismeasurement Explanations for the U.S. Productivity Slowdown". National Bureau of Economic Research Working Paper No. 21974.

Triplett, Jack E. (1982) "Concepts of Quality in Input and Output Price Measures: A Resolution of the User-Value Resource-Cost Debate," in Murray Foss (ed) The U.S. National Income and Product Accounts: Selected Topics. University of Chicago Press.

Triplett, Jack E. 1989. "Price and Technological Change in a Capital Good: A survey of research on computers." In *Technology and Capital Formation*, ed. Dale W. Jorgenson and Ralph Landau, 127-213. Cambridge, MA: MIT Press.

Triplett, Jack E. 2012. "Health System Productivity". In *The Oxford Handbook of Health Economics*, ed. Sherry Glied and Peter C. Smith, Chapter 30. Oxford: Oxford University Press.
<http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199238828.001.0001/oxfordhb-9780199238828>

Triplett, Jack E., and Barry Bosworth. 2000. "Numbers Matter: The US Statistical System and a Rapidly Changing Economy". *Brookings Policy Brief* #63.

Triplett, Jack E., and Barry Bosworth. 2004. *Productivity in the US Services Sector: New Sources of Economic Growth*. Washington, D.C.: Brookings Institution Press.

United States. 2015. *Economic Report of the President Transmitted to the Congress*. Washington: G.P.O.

U.S. Department of Housing and Urban Development. 2016. "New Residential Construction in March 2016".

<http://www.census.gov/construction/nrc/pdf/newresconst.pdf>

Viscusi, W. Kip, and Joseph E. Aldy. 2003. "The Value of a Statistical Life: A Critical Review of Market Estimates throughout the World." *Journal of Risk and Uncertainty*. 27 (1): 5-76.