

# How Government Statistics Adjust for Potential Biases from Quality Change and New Goods in an Age of Digital Technologies: A View from the Trenches

Erica L. Groshen, Brian C. Moyer, Ana M. Aizcorbe, Ralph Bradley, and David M. Friedman

*“[W]hen you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind.”*

William Thomson, Lord Kelvin, *Electrical Units of Measurement* (1883)

**T**he US Congress created and funds the US Bureau of Labor Statistics (BLS) and the US Bureau of Economic Analysis (BEA) to produce essential information on economic conditions to inform public and private decisions. A key economic indicator—what our economy produces during a time period—is called “real output” and measured as the nominal dollar value of gross domestic product (GDP) deflated by price indexes to remove the influence of inflation. To get this right, we need to measure accurately both the value of nominal GDP (done by BEA) and key price indexes (done mostly by BLS). Otherwise, real output and related measures like productivity growth statistics will be biased.

■ *Erica L. Groshen was Commissioner, US Bureau of Labor Statistics, Washington, DC, from January 2013 through January 2017. Brian C. Moyer is Director, US Bureau of Economic Analysis, Washington, DC. Ana M. Aizcorbe is Senior Research Economist, US Bureau of Economic Analysis, Washington, DC. Ralph Bradley is Chief of the Division of Price and Index Number Research, US Bureau of Labor Statistics, Washington, DC. David Friedman is Associate Commissioner for Prices and Living Conditions, US Bureau of Labor Statistics, Washington, DC. Their email addresses are [egroshen@yahoo.com](mailto:egroshen@yahoo.com), [brian.moyer@bea.gov](mailto:brian.moyer@bea.gov), [ana.aizcorbe@bea.gov](mailto:ana.aizcorbe@bea.gov), [Bradley.Ralph@bls.gov](mailto:Bradley.Ralph@bls.gov), and [Friedman.David@bls.gov](mailto:Friedman.David@bls.gov).*

† For supplementary materials such as appendices, datasets, and author disclosure statements, see the article page at

<https://doi.org/10.1257/jep.31.2.187>

doi=10.1257/jep.31.2.187

All of us have worked on these measurements while at the Bureau of Labor Statistics and the Bureau of Economic Analysis. In this article, we explore some of the thorny statistical and conceptual issues related to measuring a dynamic economy. An often-stated concern in recent years is that the national economic accounts miss some of the value of some goods and services arising from the growing digital economy. We agree that measurement problems related to quality changes and new goods have likely caused growth of real output and productivity to be understated. Nevertheless, these measurement issues are far from new and, based on the magnitude and timing of recent changes, we conclude that it is unlikely that they can account for the pattern of slower growth in recent years.

We begin by discussing how the Bureau of Labor Statistics currently adjusts price indexes to reduce the bias from quality changes and the introduction of new goods, along with some alternative methods that have been proposed. We then present estimates of the extent of remaining bias in real GDP growth that stem from potential biases in growth of consumption and investment. Based on our analysis of bias estimates performed by experts external to the Bureau of Labor Statistics and the Bureau of Economic Analysis, we find that these influences on existing price indexes may overstate inflation in the categories of “personal consumption expenditures” and “private fixed investment,” leading to a corresponding understatement of real economic growth of less than one-half percentage point per year. Furthermore, we find this to be fairly stable over time. We then also take a look at potential biases that could result from challenges in measuring nominal GDP, including assessing the significance of the argument that the digital economy has created some valuable goods and services (such as smartphone apps and Internet searches) that are not bought or sold, and thus are not counted in GDP. Finally, we review some of the ongoing work at BLS and BEA to reduce potential biases and further improve measurement.

## **Challenges in Measuring Price Indexes**

The Bureau of Labor Statistics publishes monthly indexes of consumer prices, producer prices, and import and export prices. The Consumer Price Index (CPI) measures average changes in the prices paid by urban consumers for a representative set of goods and services. The CPI is often used to make cost-of-living adjustments; indeed, the BLS uses the concept of a Cost of Living Index (as defined by Konüs 1939) as a unifying framework and is the standard by which BLS defines any bias. The Producer Price Index (PPI) measures average change in selling prices received by domestic producers for their output. Nominal value of production can be deflated by the PPI to get an output measure. Finally, Import and Export Price Indexes (MXP) measure average changes in the prices of nonmilitary goods and select services traded between the US economy and the rest of the world. The MXP allows one to estimate real aggregate trade volumes from nominal trade amounts.<sup>1</sup>

<sup>1</sup>For more information on the PPI, see Chapter 14 of the BLS *Handbook of Methods* (2006). For more detail on the CPI, see Abraham (2003) as well as Chapter 17 of the BLS *Handbook of Methods*. For more information about the MXP, see Chapter 15 of the BLS *Handbook of Methods*.

The Bureau of Labor Statistics seeks to produce the best possible monthly price indexes, subject to some very practical and binding constraints. On average, no more than 20 days elapse between the collection of a price and the final publication of the index. This must be done within a rigid budget constraint. Furthermore, the confidentiality of all collected data must be strictly protected at every stage of index construction (more at the BLS website at <https://www.bls.gov/bls/confidentiality.htm>). Respondent participation is entirely voluntary, so the Bureau of Labor Statistics also aims to minimize respondent burden because its field representatives must be able to persuade respondents to participate.<sup>2</sup> These considerations mean that for a methodological improvement to be implemented, it must meet the following criteria: it 1) is feasible within the BLS budget constraint; 2) is computable and reviewable within 20 days; 3) is compatible with the skill set of BLS staff; 4) requires no increase in samples or new surveys (unless there is budget approval for a new survey); 5) does not unduly burden respondents; and 6) is proven to reduce bias in a statistically significant manner.

So how does the Bureau of Labor Statistics treat new and evolving goods and services? To begin with, this problem is hardly a recent development. For example, consider the 1920s. That decade saw a rapid introduction of new goods such as indoor plumbing, electricity, and radios, as well as dramatic quality improvements of existing products such as automobiles and airplanes. Over the past century, technical innovation has continued to improve existing goods and has led to the introduction of myriad new products.

From a price index perspective, the biases caused by technological innovation are distinct from the sometimes-conflated issue of substitution bias. Substitution bias may occur when either a new outlet enters the market and offers existing products at a lower price, or when a foreign country starts producing a lower-price product that was already produced domestically. For example, not accounting for the substitution from US-produced manufacturing inputs to lower-priced foreign inputs, as studied by Houseman, Kurz, Lengermann, and Mandel (2011), does not cause a new goods problem but rather a substitution bias problem. In particular, this input substitution exerts an *upward* bias on US productivity statistics, while not adjusting for quality improvements or new products leads to a *downward* bias on the same statistics.

The decades-long search of academics and statistical agency staff for better ways to adjust price indexes for innovations has generated a vast literature on this topic.<sup>3</sup> These methods fall into six groups: 1) quality adjustment from producers, 2) outside surveys to measure quality changes, 3) hedonic approaches, 4) discrete

<sup>2</sup>Countries vary on whether participation in price surveys is voluntary. For example, participation in Norway's and Canada's price index programs is mandatory (for details, on Statistics Norway, see Johannessen 2016, [https://ec.europa.eu/eurostat/cros/system/files/randi\\_johansenn\\_the\\_use\\_of\\_scanner\\_data\\_in\\_the\\_norwegian\\_cpi.pdf](https://ec.europa.eu/eurostat/cros/system/files/randi_johansenn_the_use_of_scanner_data_in_the_norwegian_cpi.pdf), and for Statistics Canada, see <http://www.statcan.gc.ca/eng/survey/business/2301>).

<sup>3</sup>A partial list of references for studies done by BLS staff appears in <http://www.bls.gov/cpi/cpihqablsbib.pdf>. Those not listed in this link include Armknecht, Lane, and Stewart (1996) and Erickson and Pakes (2011).

Table 1

**Summary of Methods to Account for New and Improved Goods and Services**

<i>Method</i>	<i>Requires demand estimation</i>	<i>Based on characteristics, product, or other</i>	<i>Example of studies</i>	<i>In production</i>	<i>Reason not in production</i>
Quality adjustment from producer	No	Characteristics	Moulton, LaFleur, and Moses (1998)	Yes; PPI, MXP, CPI <sup>a</sup>	
Input from other surveys	No	Characteristics	Murphy et. al (2008)	Yes; primarily PPI	
Explicit hedonic quality adjustment	No	Characteristics	Fixler, Fortuna, Greenlees, and Lane (1999)	Yes; CPI <sup>b</sup> , PPI <sup>c</sup> , MXP <sup>c</sup>	
Time dummy hedonic index	No	Characteristics	Byrne, Oliner, and Sichel (2015); Berndt, Griliches, and Rappaport (1995); Griliches (1961)	No	Restrictive assumptions
Imputed hedonic index	No	Characteristics	Erickson and Pakes (2011)	No	Requires larger sample sizes
Discrete choice	Yes	Characteristics	Berry, Levinsohn, and Pakes (1995); Nevo (2001); Petrin (2002)	No	High computational intensity and cost; poor timeliness
Consumer surplus	Yes	Product	Lee and Pitt (1986); Hausman (1997); Broda and Weinstein (2010)	No	Endogeneity problems (under investigation); high cost
Disease-based price indexes	No	Treated disease	Aizcorbe and Nestoriak (2011); Bradley (2013)	Partial; BEA and BLS experimental indexes	Do not yet adjust for differences in outcomes

<sup>a</sup>For example this is done for new vehicles in the CPI and PPI.

<sup>b</sup>See <http://www.bls.gov/cpi/cpihqablsbib.pdf> for CPI items that are quality adjusted.

<sup>c</sup>PPI and MXP do explicit hedonic quality adjustment for computers.

choice models, 5) explicit measurement of increased consumer surplus from new goods, and 6) the special case of disease-based price indexes. Table 1 lists the various methods and shows which are used in the current production of price indexes by the Bureau of Labor Statistics. The last column explains briefly why some methods are not being used currently. While any method that government statistical agencies use for computing price indexes needs to have a sound theoretical basis, the table illustrates that the operational requirements (like allowing for index computation on a timely basis in a transparent fashion and within budgetary constraints) are also often binding.

### Current Quality and New-Product Adjustment Methods

The *matched model* is the cornerstone of constructing price indexes at the Bureau of Labor Statistics. When products match over time, the characteristics of each product are held constant. Thus, any price change can only be attributed to inflation, and not to changes in characteristics. For example, from December 2013

through November 2014, matches were found for items in the Consumer Price Index 73 percent of the time. Of the remaining 27 percent of items that were not matched, 22 percent reflected temporarily missing items, such as a bathing suit in Milwaukee in December. The other 5 percent represented a permanent disappearance. (Note that similar figures are not available for the Producer Price Index or the Import and Export Price Indexes.)

When a match permanently ends in the Consumer Price Index and the same good cannot be tracked from one period to the next, then (except for housing) the Bureau of Labor Statistics initiates a *quality adjustment procedure* after a replacement good has been established. When the replacement has characteristics very similar to the exiting product, the price of the replacement product is used in place of the exiting product. For example, of the 5 percent of the CPI that represented permanently disappearing items during the period noted above, three-fifths of those items were replaced by a similar good. For the remaining two-fifths, where the characteristics were judged to be insufficiently close, BLS staff made a quality adjustment to the replacement product's price.

In some cases, an item's producer can provide a value for the change in its characteristics. The Bureau of Labor Statistics uses this value to adjust the transaction price before it is entered into the index. This method is referred to as *explicit quality adjustment* and is most prevalent in the Producer Price Index or the Import and Export Price Indexes. It is especially important for automobiles, machinery, and other types of goods that undergo periodic model changes. For example, the price increment due to a new protective coating or electronic feature on a machine part can often be provided by its manufacturer. In general, explicit quality adjustment is easier to do for goods industries than for services industries, because producers are more likely to be willing to estimate a monetary value for the quality of a good, rather than a service. If the item's producer cannot provide a value for a characteristic change, then the *overlap* method is used. This entails having historical prices on both the exiting item and the replacement item at the same time. When computing a price index, the price of the new replacement is reduced by dividing by the ratio of the price of the new good to the exiting good (for a more detailed explanation by the BLS see "Quality Adjustment in the Producer Price Index" <http://www.bls.gov/ppi/qualityadjustment.pdf>).<sup>4</sup>

Quality adjustments *from producers* are generally cost-based. Some observers argue that utility- or welfare-based quality adjustments would be an improvement. Triplett (1982) attempts to find a resolution for this disagreement. He concludes, "In output price indexes (the fixed-weight forms as well as the theoretical ones based on production possibility curves), the quality adjustment required is equal

<sup>4</sup>This method is used by the Billion Prices Project, discussed in this journal by Cavallo and Rigobon (2016). They argue that with the overlap method, "As we increase the number of models included in the index, we more closely approximate the results of the hedonic price index constructed by the BLS." Statistics Canada also uses this approach (see "The Canadian Consumer Price Index Reference Paper," <http://www.statcan.gc.ca/pub/62-553-x/62-553-x2015001-eng.pdf>).

to the resource usage of the characteristics that changed. Only with a resource-cost adjustment does the index price a set of outputs that can be produced with the resources available in the reference period.” Because the measurement objective of real GDP is economic output, rather than welfare, the Bureau of Labor Statistics believes that cost information is appropriate for adjusting prices in an index whose purpose is to deflate nominal industry revenues to measure real output. Other practitioners agree. For example, the International Monetary Fund (2004, chapter 7) views this approach as a best practice for producer price indexes.

Finally, other surveys can be used to adjust for quality. For example, the US Department of Health and Human Services has created a Hospital Compare and a Nursing Home Compare database, which looks at inputs that experts believe can serve as proxies for quality of health care. The Bureau of Labor Statistics uses these data to adjust the hospital and nursing home components of the PPI (Bureau of Labor Statistics 2008 and “Quality Adjustment in the Producer Price Index” at <http://www.bls.gov/ppi/qualityadjustment.pdf>). In addition, the Insurance Services Office (a private firm) creates a database on the risk characteristics of cars, which BLS uses for quality adjustments in auto insurance prices (as reported in Bathgate 2011).

### **Hedonic Adjustments**

Fifty years ago, Lancaster (1966) suggested, “It is the properties or characteristics of the goods from which utility is derived.” For example, we do not consume a car; we consume its horsepower, transmission, size, audio system, and other amenities. This insight helped lead to the hedonic, or the “demand for characteristics,” model, which estimates how each characteristic contributes to the value of a good. The term “hedonics” derives from the Greek root for satisfaction. Court (1939) is often viewed as the first hedonic study that estimated a statistical relationship between prices and characteristics, but Lancaster (1966) and Rosen (1974) provide the major microeconomic insights of the demand for characteristics.

In the Consumer Price Index, about 33 percent of the total expenditures in the underlying basket of goods are eligible for quality adjustment with hedonics when price-determining characteristics change. Housing-related expenditures account for most of this share (as described by BLS at <http://www.bls.gov/cpi/cpihqitem.htm>). Liegey (1993, 1994, 2001a, b, 2003) explains the investigations made by Bureau of Labor Statistics staff on the use of hedonics for various items in the CPI. In addition, the Producer Price Index or the Import and Export Price Indexes use hedonic adjustment procedures for computers.

Hedonic indexes start with a simple regression of the price (or log price) of a good on its observable characteristics (for more detail, see Aizcorbe 2014; Silver and Heravi 2007). The explicit hedonic quality adjustment makes the prices of two differing products comparable by adjusting the price of one good using the coefficients of the regression so that it accounts for the differences in characteristics. While the hedonic index approach needs only to be a measure of the compensating variation required to keep utility constant, the explicit quality adjustment approach

requires that the coefficients on the hedonic regressions be consistent estimators of the characteristic's shadow price.

There are two variations on the hedonic approach that are not currently used by the Bureau of Labor Statistics. One approach uses a time dummy coefficient in the regression to compute the price index. In a related approach, a hedonic imputation index can be constructed by estimating a regression for each time period. The levels of characteristics are then held constant over time and a quality-adjusted price is imputed for each period. However, the assumptions behind the basic time dummy hedonic index are very restrictive (Erickson and Pakes 2011). The alternative hedonic imputation index can in theory give a more accurate measure of the compensating variation than the matched model (Pakes 2003; Erickson and Pakes 2011). Unfortunately, attempts by BLS to implement imputed hedonic price indexes revealed that our sample sizes are too small for this approach. Thus, BLS has so far continued to use explicit hedonic quality adjustment, because it has more general applicability than the time dummy approach and does not require the large sample sizes needed for the hedonic imputation indexes.

Hedonic methods are feasible when adequate sample sizes and information on relevant characteristics are available. Unlike the discrete choice and consumer surplus methods discussed in the next section, hedonic methods do not require estimating demand for a good and can be implemented with data on only prices and characteristics. These data are already collected by the price index programs. The Bureau of Labor Statistics plans to expand use of hedonic methods. However, hedonics must be implemented carefully, case by case, to ensure that key conditions are met: 1) product characteristics must be observable, ruling out features such as enhancing the user's social status; 2) the set of relevant characteristics cannot change, ruling out this approach for goods where stark new attributes are introduced frequently, such as the smartphone; and 3) the market for the product must be competitive, with markups playing only a very limited role (which ensures that a characteristic's coefficient is an unbiased estimate of its shadow price, as discussed in Rosen 1974).

When all relevant characteristics for hedonics are not available, Statistics New Zealand and Statistics Netherlands have collaborated to develop a "fixed effects window-splice" price index for scanner data (Krsinich 2016; de Haan and Krsinich 2014). If the characteristics by bar code remain fixed, then this index is equivalent to a hedonic time dummy index. This method shows much promise and could be more efficient than using the traditional time dummy approach as it would require fewer parameters. Also, this approach incorporates the effects of the entire set of characteristics whether or not they are observable. The Bureau of Labor Statistics plans to review this method.

### **Models Not in Production: Discrete Choice Model and Consumer Surplus**

Academic researchers have proposed alternative and more sophisticated procedures that are intriguing. But, as we discuss below, thus far their implementation requires restrictive assumptions and are not feasible given the production constraints at the Bureau of Labor Statistics.

The discrete choice model is a demand-for-characteristics model, like that used to motivate hedonics. However, unlike hedonic methods, discrete choice methods do not assume that the quantity of each good consumed is a continuous variable. McFadden (1978) introduces the discrete choice model, which has several theoretical advantages over hedonics. First, fewer assumptions are required for the coefficients on the characteristics to be an unbiased estimate of the characteristics' shadow prices. Second, the supply side can be explicitly modeled and markups can be allowed. Third, unlike hedonics, this method yields estimates of both consumer demand and aggregate utility based on the prices and characteristics of each product for each time period. Fourth, estimation of these models is able to take advantage of information from volume or sales data, while hedonic methods do not. In the discrete choice model, the quality-adjusted price index is usually computed as the percent change in income that consumers need to keep their expected utility constant over time (for discussion, see Nevo 2003 on how to compute quality-adjusted price indexes from discrete choice models).

However, *in practice* the discrete model approach runs into difficulties. It originally came under criticism for its highly restrictive assumptions, but additional modifications introduced later relaxed these assumptions (as discussed in Berry 1994; Berry, Levinsohn, and Pakes 1995; Nevo 2000). Yet these modifications did not solve key practical difficulties including large computational and personnel cost increases, to the extent that computationally intensive problems may take years to complete (Berry 1994; Berry, Levinsohn, and Pakes 1995; Petrin 2002; Nevo 2001). In addition, the Bureau of Labor Statistics does not have access to the necessary volume sales or share data for estimating the parameters of these models. Even if these volume data were available, to implement these models in the Consumer Price Index would pose serious logistical challenges. The CPI is constructed from over 6,000 item-area subprice indexes. For each price index item that needs quality adjustment, BLS would need to solve a separate discrete choice model for each of that item's areas and/or expenditure categories during the 20 days that BLS has available between price collection and publishing the monthly price index. This could easily require solving over a thousand models each month, which is not currently feasible.

The consumer surplus method is based on the demand for products and not characteristics. Using this method to adjust for changes in quality or new goods is problematic for other reasons. In this method, when a new product is introduced, its demand is estimated econometrically as a function of the new price and prices of incumbent products in a system of structural demand equations that include the other competing products (Hausman 1997; Redding and Weinstein 2016). The parameter estimates of the demand system are then used to solve for the *virtual price* where no one purchases this new product. The price index is computed by plugging in this virtual price for the period when the product had not been introduced. This idea that before a new good is introduced, it should be considered available but at a price at which the quantity demanded is zero, was introduced by Hicks (1940). Generally, a higher estimated virtual price will generate a higher consumer surplus coming from the new good, leading in turn to a lower price index.



Implementing this approach has proven most controversial. For example, the Hausman (1997) study of the cereal market is based on the consumer surplus model. A separate demand equation is estimated for each brand of cereal, such as Special K or Cheerios. This method is more computationally intensive than the hedonic method, where only one reduced form regression of log price on characteristics is estimated. This approach has the potential for bias from specification error, using the wrong instruments, and making incorrect assumptions on how demand shocks affect prices.<sup>5</sup> These sources of econometric bias could produce a biased estimate of the virtual price. The Committee on National Statistics advises against using the Hausman (1997) method, citing the risks of introducing new errors (National Research Council 2002). Moreover, a potentially troubling aspect of this approach is that the bias of the virtual price may be extremely large. For example, Hausman (1997) concludes that the CPI for cereal “may be too high by about 25%.” His average virtual price of the new cereal product is \$7.14 while the observed average price is \$3.78. It is difficult to imagine that introducing a new flavor of Cheerios could have such a large impact on a true cereal price index.

Newer versions of the consumer surplus approach are still being evaluated for production feasibility. As one example, Redding and Weinstein (2016) and Broda and Weinstein (2010) use a constant elasticity of substitution form of preferences. They assume that the elasticity of substitution is bounded below by one. With these restrictions, they only need to estimate the elasticity of a substitution parameter and not additional preference parameters. This method is far more computationally efficient than Hausman’s (1997) approach, and it sidesteps the issue that a “virtual price” where demand is literally zero can be very high. Typically, this method is used with scanner data, and it is possible that when prices bounce over a fixed time period and then return to their original level, the price index does not equal one. This problem is referred to as *chain drift*. Any index that does not satisfy the transitivity and time reversal axioms can be subject to chain drift. Ivancic, Diewert, and Fox (2009) develop a multilateral price index method that corrects for this problem, and Krsinich (2015) shows how this method has been adopted for the New Zealand electronics equipment price index.<sup>6</sup>

<sup>5</sup>Hausman (2003) argues that his consumer surplus method is the solution to the new goods problem in the Consumer Price Index. When Hausman (1997) accounted for the introduction of Apple Cinnamon Cheerios, he had to estimate the demand of all other brands of cereal. Bresnahan (1997) and Bradley (2005) critique the approach used in Hausman (1997). Among other concerns, Hausman’s use of prices in other markets as instruments and his use of unit values could have distorted his estimation of the virtual price. Abraham (2003) provides a thorough explanation of the problems of implementing the Hausman (1997) method.

<sup>6</sup>This index is called a GEKS index, after Gini (1931), Eltető and Köves (1964), and Szulc (1964). While this approach eliminates chain drift, month to month price indexes are functions of prices that are outside the base and comparison months. This is called “loss of characteristicity” (de Haan and van der Grient 2011). Even if the individual bilateral indexes that are plugged into this GEKS index are superlative indexes, there is no guarantee that the GEKS index itself is a superlative index.

### **Disease-Based Price Indexes**

In our view, a substantial current challenge for adjusting for new products lies in the medical sector because of its size, rapid innovation, and unique market features. While many observers focus on technological innovations in the digital economy like the smartphone or other aspects of information technology, the medical sector is still critical because healthcare spending is large: 17.5 percent of GDP in 2014. In the same year, all spending on phones was \$16.6 billion—about 0.1 percent of GDP. Several studies conclude that medical price indexes are upwardly biased because many new treatments for particular diseases replace more expensive approaches (for examples, for heart disease, see Cutler, McClellan, Newhouse, and Remler 1998; for depression, see Berndt et al. 2002; more broadly, see Shapiro, Shapiro, and Wilcox 2001).

Presently, there are two sources of medical price data: providers' billing offices and health insurance claims. Neither source provides characteristics data, such as the remission of disease, length of time for healing, and other indicators for patient wellness. This is a major constraint for quality-adjusting medical prices.

As a partial solution, the Bureau of Labor Statistics and the Bureau of Economic Analysis have created experimental disease-based price indexes that correct for a portion of the new goods bias in medical care, specifically the part that arises when less-expensive goods and services substitute for more-expensive treatments (for discussion, see Aizcorbe and Nestoriak 2011; Bradley 2013).<sup>7</sup> These substitutions often occur as the result of an innovation, such as a new drug that lowers the need to use expensive therapies to treat a disease. For example, the introduction of selected serotonin re-uptake inhibitors (SSRIs) represented a new generation of antidepressants that allowed fewer more-expensive therapy visits. Disease-based price indexes report medical inflation by the treatment of disease, rather than by the good or service that treats this disease. However, they do not at this point account for improved outcomes, such as increases in life expectancy coming from an innovation such as coronary bypass surgery. Disease-based price indexes are still a work in progress, and are not yet ready to be officially published medical price indexes, which supplement the medical practice indexes that are reported on a goods and services basis.

### **Estimated Quality-Adjustment and New Goods Biases and Measured Real GDP Growth**

The Bureau of Economic Analysis uses price indexes to decompose changes in nominal GDP growth into a portion that reflects inflation and a portion that reflects growth in real output. This use of price indexes to deflate nominal spending implies

<sup>7</sup>See <https://www.bls.gov/pir/diseasehome.htm> at the BLS website for more information on BLS experimental disease-based price indexes. See [https://www.bea.gov/national/health\\_care\\_satellite\\_account.htm](https://www.bea.gov/national/health_care_satellite_account.htm) at the BEA website for more information on BEA's treatment of disease-based measures as reflected in Health Satellite Accounts.

that any bias that overstates inflation will result in a downward bias in the attendant measure of real GDP. Bias can also result from challenges in measurement of nominal GDP, which is discussed in the next section.

Here, in order to get some sense of the degree to which real GDP growth may be understated, we apply results from studies that use available empirical evidence to form judgmental assessments of biases in the price indexes. Such studies exist for indexes underlying two of the major expenditure categories of GDP that, together, make up about 85% of the GDP: personal consumption expenditures and private fixed investment. For example, the Boskin Commission concluded that, in 1996, the Consumer Price Index was likely biased by +1.1 percentage points per year, with about half of the bias attributable to problems with accounting for quality change and new goods (Boskin et al. 1997, table 3). Several years later, Lebow and Rudd (2003) estimated that in 2001, CPI growth was biased by +0.87 percentage points per year, with +0.37 percentage points of that stemming from quality change and new goods bias (Lebow and Rudd 2003, table 1). Among the reasons they cited for a lower estimate than had been reported by the Boskin Commission were the use of new empirical studies of biases in the underlying CPI components and improvements in methods implemented by the Bureau of Labor Statistics after the Boskin Commission convened. Beyond the CPI, Byrne, Fernald, and Reinsdorf (2016) used price indexes developed by Byrne and Corrado (2017) to provide a similar analysis for the effect of bias in price indexes for information and communications technology equipment on real growth of private fixed investment.

These careful studies inherently involve an element of judgment, and the investigators clearly emphasize the uncertainty surrounding their estimates. For example, Lebow and Rudd (2003) assigned a subjective 90 percent confidence interval to their point estimate of total bias in the Consumer Price Index, yielding an estimated bias that ranged from 0.3 to 1.4 percentage points. Likewise, Byrne, Fernald, and Reinsdorf (2016) provide two sets of estimates, which they label “conservative” and “liberal.” Despite the uncertainty, we believe these estimates are of great value to help direct improvement efforts, inform users of data limitations, and (in the case of Byrne et al.) rule out certain hypotheses, such as a recent large increase in bias.

To consider the effect of imperfect adjustments for quality change and new goods purchased by households, we use the Lebow and Rudd (2003) assessments of biases in the individual components of the Consumer Price Index, representing the most recent comprehensive set of estimates available, many of which are used to deflate the components of personal consumption expenditures. We make adjustments to their estimates for medical care spending and internet services to reflect results available from more recent studies. For medical care, Cutler, Rosen, and Vijan (2006) studied the impact of innovations in medical care on life expectancy and found that improved mortality over the 1960–2000 period was concentrated in five conditions (see their table 2): cardiovascular disease (4.9 years), death in infancy (1.4 years), death from external causes (4 months), pneumonia and influenza (3 months), and malignant cancers (2.5 months). According to the Agency for Healthcare Research and Quality (2017), spending on these conditions totaled \$363

Table 2

**Impact of Estimated Biases to Personal Consumption Expenditures (PCE) Deflators on Measured Real GDP Growth, 2000–2015**

<i>Expenditure Category</i>	<i>Share of GDP</i>				<i>Lebow–Rudd (2003) bias</i>
	<i>2000</i>	<i>2005</i>	<i>2010</i>	<i>2015</i>	<i>2003</i>
<i>Medical care:</i>					
Prescription drugs	1.3%	1.6%	1.9%	2.3%	1.20%
Nonprescription drugs	0.2%	0.2%	0.3%	0.3%	0.50%
Medical care services <sup>a</sup>	9.8%	10.9%	12.2%	12.5%	0.76%
<i>PC services (including internet)<sup>b</sup></i>	0.2%	0.2%	0.4%	0.6%	6.50%
<i>Contributions to real GDP growth (percentage points per year)</i>					
<i>Medical care:</i>					
Prescription drugs	–0.02	–0.02	–0.02	–0.03	
Nonprescription drugs	0.00	0.00	0.00	0.00	
Medical care services	–0.07	–0.08	–0.09	–0.09	
<i>PC services (including internet)</i>	–0.01	–0.01	–0.03	–0.04	
<i>All other PCE categories</i>	–0.10	–0.10	–0.10	–0.09	
<b>All personal consumption expenditure categories</b>	<b>–0.20</b>	<b>–0.22</b>	<b>–0.24</b>	<b>–0.26</b>	

*Note:* Table 2 presents our translation of the impact of bias in each segment of personal consumption expenditures on measured real GDP growth, calculated by multiplying each Lebow and Rudd (2003) bias estimate—the percentage point overstatement of price change in the individual price index per year—by the segment’s share of GDP (with the adjustments noted in the text). Total for All personal consumption expenditure categories may not add up exactly to the subcomponents shown in the columns due to rounding.

<sup>a</sup>Bias estimate for medical care services has been adjusted based on data from AHRQ (2017).

<sup>b</sup>Bias estimate for PC services (including internet) is based on Greenstein and McDevitt (2011).

billion, or about 20 percent of all spending on medical care. To fold this analysis into an estimated bias, we follow Lebow and Rudd in reducing the 4.5 percentage point bias reported in Cutler et al. (1998) by 0.7 percentage point, to account for improvements to the methods used by the Bureau of Labor Statistics. However, given the new information in Cutler (2006), we apply the resulting bias (4.5 – 0.7 = 3.8) to a smaller share of spending (1/5 versus the 2/3 used in Lebow–Rudd). This translates into a +0.76 bias for medical care services overall.

Similarly, Greenstein and McDevitt (2011) estimate that price indexes for broadband access overstate price growth by 3 to 10 percentage points, and we apply the midpoint of this range to the personal consumption services (including internet) category in overall personal consumption expenditures. For the remaining categories of personal consumption expenditures, we use the Lebow and Rudd (2003) estimates, suitably reweighted.

Table 2 presents our translation of the impact of bias in each segment of personal consumption expenditures on measured real GDP growth, calculated by multiplying each Lebow and Rudd (2003) bias estimate—the percentage point overstatement of price change in the individual price index per year—by the segment’s

Table 3

**Impact of Estimated Biases to Private Fixed Investment (PFI)  
Deflators on Measured Real GDP Growth, 2000–2015**

<i>Equipment type</i>	<i>Share of GDP</i>				<i>Byrne, Fernald, and Reinsdorf (2016) estimated bias</i>	
	<i>2000</i>	<i>2005</i>	<i>2010</i>	<i>2015</i>	<i>1995–2004</i>	<i>2004–2014</i>
Communication equipment	1.2%	0.7%	0.6%	0.6%	5.8%	7.6%
Computers and peripherals	1.0%	0.6%	0.5%	0.4%	8.0%	12.0%
Other information systems equipment	0.7%	0.7%	0.7%	0.8%	8.3%	5.4%
Software	1.8%	1.7%	1.7%	1.8%	1.4%	0.9%
	<i>Contributions to real GDP growth (percentage points per year)</i>					
Communication equipment	-0.07	-0.04	-0.03	-0.03		
Computers and peripherals	-0.08	-0.05	-0.04	-0.03		
Other information systems equipment	-0.05	-0.06	-0.06	-0.06		
Software	-0.03	-0.02	-0.02	-0.03		
<b>All private fixed investment categories</b>	<b>-0.23</b>	<b>-0.17</b>	<b>-0.16</b>	<b>-0.15</b>		

*Note and Source:* To assess the importance of quality change and new goods bias in investment goods, we use the Byrne, Fernald, and Reinsdorf (2016, table 1) estimates for biases in the deflators that the Bureau of Economic Analysis uses for information and communications technology products. The contributions to GDP growth for 2000 and 2005 are calculated using the bias estimates for 1995–2004; the contributions for 2010 and 2015 use the bias estimates for 2004–2014. Total for All private fixed investment categories may not add up exactly to the subcomponents shown in the columns due to rounding.

share of GDP (with the adjustments noted above). Note that a *positive* bias in the Consumer Price Index leads to a *negative* bias in GDP growth. To assess whether growth of problematic sectors has increased the potential bias to GDP growth, we report the resulting contributions using GDP shares for 2000, 2005, 2010, and 2015.

The estimated overall impact of the biases of consumption deflators on real GDP appears in the bottom line: measured growth in real GDP was reduced by -0.20 percentage point in 2000, with this downward bias growing only modestly over time to -0.26 percentage point in 2015. Of course, because the bias estimates that we apply are point-in-time estimates, we cannot assess here how the biases to the individual components may have changed over time, including the impact of continued improvements in measurement of the Consumer Price Index; hence, the increase in overall bias reported here only reflects the effects of changes over time in GDP shares.

To assess the importance of quality change and new goods bias in investment goods, we use the Byrne, Fernald, and Reinsdorf (2016, table 1) estimates for biases in the deflators that the Bureau of Economic Analysis uses for information and communications technology products. As shown in Table 3, they report separate estimates for biases in the pre- and post-slowdown periods, which show increases in the biases for communications equipment and computers and peripherals but declines in the biases for other information systems equipment. All told, the implied

impact to measured real GDP growth from biases in these information and communications technology components of private fixed investment is small: less than 1/4 percentage point in all four years, with a decline to  $-0.15$  percentage point in 2015.<sup>8</sup>

Taken together, the reduction in measured real GDP growth from biases in both personal consumption expenditures and private fixed investment would be about  $-0.4$  percentage point in all four periods.

## Challenges in Measuring Nominal GDP in the Digital Age

In addition to the concern that price indexes do not adequately capture quality changes and new goods, there is a separate concern that estimates of GDP by the Bureau of Economic Analysis ignore valuable new goods and services that aren't sold, such as internet searches or encyclopedia services that are provided essentially free to the user. In addition, changes in the way that households and firms obtain goods and services, such as an evolving ability of firms to outsource goods and services previously provided in-house, have also raised questions about whether and how these phenomena are measured in GDP. Overall, these concerns seem overstated. For most cases mentioned, either the official government statistics have always excluded the value of products similar to these because they are outside the scope of what GDP aims to measure, or the official economic statistics do actually include their value, although it is embedded in other measured market activity.

As economists tell students in every introductory economics class, GDP measures the *market* value of the goods, services, and structures produced by the nation's economy in a particular period. GDP is not designed to measure well-being or welfare: for example, it does not account for rates of poverty, crime, or literacy. Nor does GDP attempt to comprehensively measure nonmarket activity, such as household production or "free" services, digital or otherwise. Whether GDP should be measured more expansively has long been debated by economists. Indeed, Kuznets (1934), an early architect of the national economic accounts, noted the shortcomings of focusing exclusively on market activities and of excluding nonmarket activities and assets that have productive value.

Since then, a tremendous amount of research has sought to develop methods to better address the need to include nonmarket or near-market activities and to better measure economic welfare. Given the conceptual and practical difficulties of such efforts, however, several National Academy of Sciences studies on the environment (Nordhaus and Kokkelenberg 1999) and nonmarket production (Abraham and Mackie 2005), as well as the *System of National Accounts 2008* (European Commission

<sup>8</sup>Hatzius (2015) finds that the estimated biases in price indexes related to information and communications technology are sufficiently high and grow enough in 2005 to explain a significant part of the slowdown in measured productivity that happened at about that time. However, as pointed out in Sverson (2016) and Gordon (2015a), the assumed biases that require this result seem implausibly high.

et al. 2009) guidelines for measuring GDP, all recommend that the core GDP account concepts remain as is, while encouraging the creation of supplemental (or satellite) accounts to address other issues. For example, BEA began publishing a satellite account for household production in 2012 that provides estimates for GDP that incorporate the value of home production by households, the largest component of which is production of nonmarket services like cooking, gardening, or housework (Bridgman et al. 2012). In a similar spirit, the Stiglitz–Sen–Fitoussi commission (2009) report on expanded welfare measures suggests ways that “‘classical GDP issues’ can be addressed within existing GDP accounts or through an extension and improvement of measures included in existing accounts” (Jorgenson, Landefeld, and Nordhaus 2006; Landefeld, Moulton, Platt, and Villones 2010). We agree with Ahmad and Schreyer (2016) that while building alternative welfare measures can certainly be useful and informative, such measures have a different purpose than GDP statistics.

To evaluate how emerging economic phenomena are treated in the national accounts, we look separately and systematically at how the Bureau of Economic Analysis assembles three of the four expenditure components of GDP. (We omit government for brevity, since it has not been the focus of concern about productivity mismeasurement.) Specifically, we focus on some central examples of goods and services related to the digital economy, intellectual property, and globalization, and discuss how they are included, or not, in these components of GDP.

To begin, we note that consumption spending in GDP statistics consists of *purchases* of goods and services by households (families and unrelated individuals) and nonprofit institutions serving households (such as Goodwill Industries International). The Bureau of Economic Analysis derives these estimates from statistical reports and surveys, primarily from the Census Bureau but also from other government agencies, administrative and regulatory agency reports, and reports from private organizations, such as trade associations (BEA 2016a). Most new goods and services that involve market transactions are likely properly recorded in GDP. For example, purchases made on the Internet have been reported to the Census Bureau’s retail trade surveys since 1999. However, there may be lags between the time when the new good is first introduced and when it is represented in the source data.

Some observers are concerned that consumption as measured by the GDP misses “free” digital services used by households. For example, Internet services such as Google search or Facebook are provided without any direct charge to users, which might suggest that they aren’t included in GDP. However, most providers of these “free” Internet services charge advertisers. In GDP, ad-supported content—like television—is treated as an intermediate input, which means that its value is reflected in the value of the goods or services that are sold through advertising. Nakamura and Soloveichik (2015) find that an alternative treatment that regards ad-supported content as part of consumption in GDP would likely have only a small impact on GDP: US spending on advertising has been about 1.3 percent of nominal GDP for decades. However, content and services provided online without advertising such as Wikipedia, blogs, photo archives, and so on, will not be directly recorded in GDP, because there is no associated market transaction.

Others have raised questions about what happens when economic activity shifts between market and household production. For example, the Internet and its powerful search engines have lowered the cost of finding information. This has moved some activities into the home, like the services that used to be provided by paid travel agencies, for example. While estimates of these kinds of changes would be useful, perhaps in the household satellite account, such activities are appropriately excluded from GDP when they become a nonmarket activity; after all, if an activity does not provide income to some party, it is not part of the market activity that makes up GDP.

Next, consider business sector purchases. The business sector comprises all for-profit corporate and noncorporate private entities, a broad category that includes mutual financial institutions, private noninsured pension funds, cooperatives, nonprofit organizations that primarily serve businesses, Federal Reserve Banks, federally sponsored credit agencies, and government enterprises. From the standpoint of calculating GDP, this sector is responsible for gross private domestic investment and consists of purchases of fixed assets (structures, equipment, and intellectual property) by private businesses that contribute to production and have a useful life of more than one year, purchases of homes by households, and investment in inventories. Of course, purchases of other goods and services used by businesses for production of final goods are counted in GDP as intermediate (not final) goods.

The digital economy poses several challenges for measuring investment, particularly investment in intellectual property. That category of investment currently includes software, scientific research and development, and artistic originals. Because the intangible nature of these assets makes this type of investment especially hard to measure (Corrado, Hulten, and Sichel 2009), some notable examples of intellectual property that are currently not counted as investment in national accounts include research and development spending on social sciences and humanities and certain economic competencies (for example, as embodied in advertising, marketing, or organizational structure).

Some observers wonder what happens to GDP when services previously conducted internally by firms are moved to the Internet cloud. For example, whereas firms previously invested in their own servers and software, now they increasingly pay for cloud services from a central firm. To the extent that investments in servers and other equipment are shifted from firms to cloud service providers, this development would not cause mismeasurement of GDP. Servers are still purchased and recorded as investment, and purchased cloud services are recorded as intermediate goods.

While we think that goods and services that involve market transactions are likely properly included in total GDP spending, there are cases where the transactions can be misallocated across GDP categories. Consider, for example, the introduction of ride-sharing applications like Uber and Lyft and alternates to hotel services like Airbnb. To the extent that individuals earn income from selling their services through these platforms, income and output measures will accurately



reflect the contribution of these new products to consumer spending in GDP final expenditures. But such changes still raise measurement issues. One is that to the extent that (unincorporated) individuals are using assets to provide transportation services, for example, then spending on such assets should be recorded as business investment, not consumption. That said, any distortions from this are likely small. In the case of Uber, for example, Bean (2016) estimates that this kind of misclassification for vehicles could perhaps amount to no more than 1.5 percent of business investment.

Since many US-made products are not consumed here and others used here are made abroad, GDP must include exports and subtract imports. Measurement issues also extend to this adjustment. In a globalized economy, many goods are US-designed (an investment) but manufactured abroad. This practice creates particular problems in accounting for the domestic value of intellectual property. For example, consider a smartphone that is designed in the United States, produced in an Asian country, and then purchased and imported by the US firm for final sale. The Bureau of Economic Analysis counts the wholesale value of the phone, which may include the value of the US firm's intellectual property, as an import and in final sales. Ideally, BEA would also capture the export of the intellectual property to the foreign producer on its surveys of international trade in services. However, under certain contract manufacturing arrangements, there may be no separate transaction for exports of design/software to the Asian manufacturer, thus understating exports in the national accounts. Ongoing work at the BEA and elsewhere continue to explore the potential magnitude of potential omissions like this (for example, Houseman and Mandel 2015).

All told, we believe that concerns about a downward bias on output are overstated because for most cases mentioned, either the value of these products is outside the scope of GDP or is embedded in other measured market activity.

## **Improving Measurement of Prices and Output**

In this section, we review some of the projects underway at the Bureau of Labor Statistics and the Bureau of Economic Analysis that will improve the ability to measure changes in real output.

### **Initiatives at the Bureau of Labor Statistics**

The Bureau of Labor Statistics continually looks at possible new data sources and for ways to expand the use of quality adjustment methods like hedonic analysis. The most frequent issue is that these alternative sources of data on prices usually lack data on characteristics of goods or on the arrival of new goods. These deficiencies can make it challenging to address the issues of adjusting for quality change and new goods that are the focus of this article.

For instance, the Billion Prices Project (discussed in this journal in Cavallo and Rigobon 2016) scrapes prices from the internet while the Adobe Digital Economy

Project (ADEP) uses price and quantity information stored by its cloud service customers. Currently, if respondents do not send electronic records and do not permit use of their application programming interface, then web-scraping is an option for which the Bureau of Labor Statistics is developing expertise. Note that many transactions prices are not available online, so it is impossible to generate an “all items” index with scraped data or with ADEP’s cloud. For example, transactions-level medical, college tuition, new vehicle, and utility prices cannot be scraped, nor are they stored on a marketing cloud. Furthermore, many price-determining characteristics are missing, and highly sophisticated programming is required to identify new goods. (However, the “fixed effects window-splice” approach discussed in the price index section could provide the same results as a traditional time dummy approach.) For prices where web-scraping is viable, it can be contracted out. Indeed, Statistics New Zealand has hired PriceStats, the commercial arm of the Billion Prices Project, to do web-scraping. The other route is to develop skills in-house for a method that will address the confidentiality and informed consent responsibilities of the BLS.

A few respondents for the Consumer Price Index have offered their electronic transactions data to the Bureau of Labor Statistics. To be useful as building blocks for price indexes that adjust for quality changes, these records must contain both transaction prices and price-determining characteristics. To date, BLS receives electronically transferred data from two major national retailers and one market research firm. The data from one of the retailers is now in production. BLS continues to work with the other two respondents to obtain an adequate set of price-determining characteristics.

The Bureau of Labor Statistics began investigating use of scanner data over 20 years ago and has used it over the years to validate and diagnose current Consumer Price Index samples, but not directly for monthly production. Purchasing real-time data is expensive. Recently, BLS obtained an estimate for purchasing real-time scanner data for grocery stores from a third party, and the estimated cost would exceed CPI’s current grocery store data collection costs by as much as 72 percent. It also can be challenging to adjust for new goods in this data, although computationally efficient new goods adjustment methods using scanner data are also being investigated (for example, Broda and Weinstein 2010; Redding and Weinstein 2016). Testing of these kinds of models is underway to see if the methods can be implemented under the BLS time and budget constraints. BLS has also observed other countries’ uses of scanner data. Statistics New Zealand, Statistics Netherlands, and Statistics Norway have already incorporated scanner data in their indexes (Krsinich 2015; van der Grient and de Haan 2010, 2011; Johannessen 2016). However, there is a large difference between the way these countries and the United States can get their scanner data. Both New Zealand and Norway receive such data directly from the outlets, while the BLS would have to purchase scanner data from a private vendor. Hence the higher cost to BLS mentioned above.

When it comes to quality changes and new goods, the medical care sector poses particular challenges. Thus, the Bureau of Economic Analysis and the Bureau of

Labor Statistics are working on quality-adjusting the medical price indexes on a disease-by-disease basis discussed earlier. For each disease, BLS plans to use Medicare claims data to follow both the illnesses and the treatments of the same patients over time. The output measure is the number of days that a patient survives treatment without being readmitted to an inpatient hospital or contracting additional illnesses (like infections). This method closely follows the Romley, Goldman, and Sood (2015) measure of the output of inpatient hospital services, although BLS will instead focus on the output of treatment bundles, similar to the approach used in Berndt et al. (2002).

The Bureau of Labor Statistics continues to expand use of hedonics. In the Producer Price Index, for example, hedonic methods are now used to quality-adjust broadband services. Further hedonic adjustments are under investigation for women's dresses and network switches. While the PPI currently uses hedonic adjustment for computers, improvements are being investigated for microprocessors where new cross-validation methods will allow a more parsimonious use of parameters. The added method of cross-validation is important because the sample sizes used to compute the hedonic regression are usually small. In addition, possible approaches to implementing imputed hedonic price index methods are under consideration (for example, Pakes 2003; Erickson and Pakes 2011). BLS will investigate whether new cross-validation and machine-learning techniques will allow estimation of hedonic price indexes with current sample sizes.

### **Initiatives at the Bureau of Economic Analysis**

The Bureau of Economic Analysis has a number of initiatives underway to address challenges related to measuring GDP in an economy characterized by digital technologies and commerce, as well as by global value chains.

Regarding consumer spending, the Bureau of Economic Analysis will continue its research on ways to improve the treatment of advertising-supported media, acknowledging that the Internet has fundamentally changed the way households consume entertainment services. BEA will seek to improve consumer spending estimates to reflect the growing importance of e-commerce. As online purchases represent an increasing share of consumer spending, BEA will update its data sources and methods to capture more information about these purchases.

With respect to digital technologies and commerce, the Bureau of Economic Analysis plans to publish a roadmap that outlines the research needed to more accurately measure the contribution of information technology to economic growth and to improve the measurement of digitally enabled commerce. In addition, BEA has begun publishing an annual report that examines trends in services for which digital technologies likely play an important role in facilitating trade, such as telecommunications, insurance, and financial services (Grimm 2016; Bureau of Economic Analysis 2016b).

Concerning issues raised by globalization and trade, the Bureau of Economic Analysis will be expanding coverage to include intellectual property transactions, which will provide a more complete picture of foreign trade in computer, audio-visual media, and research and development services. Also, BEA plans to take several

steps to better measure the value that is created at various steps in such production chains: for example, it will continue to update and refine its supply-use tables, including the production of extended supply-use tables that introduced firm-level characteristics such as ownership type and multinational status, allowing for a more refined analysis of global value chains.

Finally, the Bureau of Economic Analysis has embarked on several initiatives with statistical agency partners to leverage alternative data sources to improve the measurement of high-tech goods and services prices. For example, with regard to software, which accounts for half of investment in information and communications technology goods, BEA has purchased three types of data for potential construction of price indexes: scanner data for consumer titles, administrative data for commercial applications, and data to shed light on custom software development. Regarding cell phones and wireless plans, which pose thorny measurement challenges, BEA has obtained survey data that can potentially be used to construct price indexes for both phone and wireless services; the data contain information on households' annual payments, whether the payment includes the phone, specific features of the cell phones, or features of the plans. In addition, BEA is also supporting academic researchers who are exploring how best to measure prices for cloud computing services.

The task of calculating price indexes and output in the 21st century, and doing so in a way that provides timely monthly data within budget constraints, is not for the rigid or the fainthearted. The Bureau of Labor Statistics and Bureau of Economic Analysis agree that price index mismeasurement continues to lead to understated growth in real output over time, perhaps especially in healthcare but also possibly in areas related to information and communications technology. Although rapid innovation and globalization present numerous measurement challenges, we are optimistic that they can be addressed. Government statistical agencies and academic researchers have successfully, if often incrementally, overcome many previous problems that once seemed intractable, whether in nominal GDP, price indexes, or elsewhere. Statistics are always estimates; they will never be perfect. Yet official economic statistics particularly possess a unique combination of accuracy, objectivity, relevance, timeliness, and accessibility that serve as infrastructure in support of efficient markets. They are essential to help policymakers and citizens form opinions and make decisions.

■ *The authors thank Alyssa Holdren for excellent research assistance; Dee Bathgate, Rob Cage, Jeff Hill, Crystal Konny, John Layng, Bonnie Murphy, and Bill Thompson for information regarding the handling of quality adjustment and the introduction of new goods in BLS price indexes; and Abe Dunn, Lucy Eldridge, David Lebow, Jeremy Rudd, Jon Samuels, and Jay Stewart for very useful comments.*

## References

- Abraham, Katharine G.** 2003. "Toward a Cost-of-Living Index: Progress and Prospects." *Journal of Economic Perspectives* 17(1): 45–58.
- Abraham, Katharine G., and Christopher Mackie, eds.** 2005. *Beyond the Market: Designing Nonmarket Accounts for the United States*. Washington, DC: National Academies Press.
- Agency for Healthcare Research and Quality (AHRQ)** 2017. Total Expenses and Percent Distribution for Selected Conditions by Type of Service: United States, 2014. Medical Expenditure Panel Survey Household Component Data. Generated interactively on March 21, 2017.
- Ahmad, Nadim, and Paul Schreyer.** 2016. "Measuring GDP in a Digitalised Economy." OECD Statistics Working Papers 2016/07.
- Aizcorbe, Ana.** 2014. *A Practical Guide to Price Index and Hedonic Techniques*. Oxford University Press.
- Aizcorbe, Ana, and Nicole Nestoriak.** 2011. "Changing Mix of Medical Care Services: Stylized Facts and Implications for Price Indexes." *Journal of Health Economics* 30(3): 568–74.
- Armkecht, Paul A., Walter Lane, and Ken Stewart.** 1996. "New Products and the U.S. Consumer Price Index." Chap. 9 in *The Economics of New Goods*, 373–96. National Bureau of Economic Research.
- Bathgate, Deanna.** 2011. "Mini-Presentation by David Friedman: The 26th Voorburg Group Meeting on Service Statistics, Newport, U.K., 19–23 September 2011." <http://www4.statcan.ca/english/voorburg/Documents/2011%20Newport/Papers/2011%20-%2036.pdf>.
- Bean, Charles.** 2016. *Independent Review of UK Economic Statistics*. Cabinet Office, United Kingdom.
- Berndt, Ernst R., Anupa Bir, Susan H. Busch, Richard G. Frank, and Sharon-Lise Normand.** 2002. "The Medical Treatment of Depression, 1991–1996: Productive Inefficiency, Expected Outcome Variations, and Price Indexes." *Journal of Health Economics* 21(3): 373–96.
- Berndt, Ernst R., Zvi Griliches, and Neil J. Rappaport.** 1995. "Econometric Estimates of Price Indexes for Personal Computers in the 1990's." *Journal of Econometrics* 68(1): 243–68.
- Berry, Steven T.** 1994. "Estimating Discrete-Choice Models of Product Differentiation." *RAND Journal of Economics* 25(2): 242–62.
- Berry, Steven, James Levinsohn, and Ariel Pakes.** 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 63(4): 841–90.
- Boskin, Michael, Ellen R. Dulberger, Robert J. Gordon, Zvi Griliches, and Dale W. Jorgenson.** 1997. "The CPI Commission: Findings and Recommendations." *American Economic Review* 87(2): 78–83.
- Bradley, Ralph.** 2005. "Pitfalls of Using Unit Values as a Price Measure or Price Index." *Journal of Economic and Social Measurement* 30(1): 39–61.
- Bradley, Ralph.** 2013. "Feasible Methods to Estimate Disease Based Price Indexes." *Journal of Health Economics* 32(3): 504–514.
- Bresnahan, Timothy.** 1997. Comment on "Valuation of New Goods under Perfect and Imperfect Competition," by Jerry A. Hausman. In chap. 5 of *The Economics of New Goods*, edited by Timothy F. Bresnahan and Robert J. Gordon. University of Chicago Press.
- Bridgmen, Benjamin, Andrew Dugan, Mikhael Lal, Matthew Osborne, and Shaunda Villones.** 2012. "Accounting for Household Production in the National Accounts, 1965–2010." *Survey of Current Business*, May.
- 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. "Measuring the Economy: A Primer on GDP and the National Income and Product Accounts." December. Available at: <https://bea.gov/methodologies/index.htm>.
- Bureau of Economic Analysis.** 2016a. "Concepts and Methods of the U.S. National Income and Product Accounts." <https://www.bea.gov/national/pdf/allchapters.pdf>.
- Bureau of Economic Analysis.** 2016b. "U.S. International Services: Trade in Services in 2015 and Services Supplied Through Affiliates in 2014." *Survey of Current Business*, December.
- Bureau of Labor Statistics (BLS).** 2008. "Proposal for Adjusting the General Hospital Producer Price Index for Quality Change." February 15. <http://conference.nber.org/confer/2008/si2008/PRCR/murphy2.pdf>.
- Bureau of Labor Statistics (BLS).** 2016. *Handbook of Methods*. Division of BLS Publishing, Office of Publications and Special Studies.
- Byrne, David M., and Carol Corrado.** 2017. "ICT Prices and ICT Services: What Do They Tell Us about Productivity and Technology?" Finance and Economics Discussion Paper 2017-015, Federal Reserve Board, May.
- Byrne, David M., John G. Fernald, and Marshall B. Reinsdorf.** 2016. "Does the United States have a Productivity Slowdown or a Measurement

Problem?" *Brookings Paper on Economic Activity*, Spring. <https://www.brookings.edu/bpea-articles/does-the-united-states-have-a-productivity-slow-down-or-a-measurement-problem/>.

**Byrne, David M., Stephen D. Oliner, and Daniel E. Sichel.** 2015. "How Fast are Semiconductor Prices Falling?" NBER Working Paper 21074.

**Cavallo, Alberto, and Roberto Rigobon.** 2016. "The Billion Prices Project: Using Online Prices for Measurement and Research." *Journal of Economic Perspectives* 30(2): 151–78.

**Corrado, Carol, Charles Hulten, and Daniel Sichel.** 2009. "Intangible Capital and U.S. Economic Growth." *Review of Income and Wealth* 55(3): 661–85.

**Court, Andrew T.** 1939. "Hedonic Price Indexes with Automotive Examples. In *The Dynamics of Automotive Demand*, edited by Charles F. Roos, 99–117. New York: General Motors Corporation.

**Cutler, David M., Mark McClellan, Joseph P. Newhouse, and Dahlia Remler.** 1998. "Are Medical Prices Declining? Evidence from Heart Attack Treatments." *Quarterly Journal of Economics* 113(4): 991–1024.

**Cutler, David M., Allison B. Rosen, and Sandeep Vijan.** 2006. "The Value of Medical Spending in the United States, 1960–2000." *New England Journal of Medicine* 355(9): 920–27.

**de Haan, Jan, and Haymerik A. van der Grient.** 2011. "Eliminating Chain Drift in Price Indexes Based on Scanner Data." *Journal of Econometrics* 161(1): 36–46.

**de Haan, Jan, and Frances Krsinich.** 2014. "Scanner Data and the Treatment of Quality Change in Non-Revisable Price Indexes." *Journal of Business & Economic Statistics* 32(3): 341–58.

**Eltető, O., and P. Köves.** 1964. "On a Problem of Index Number Computation Relating to International Comparisons." *Statisztikai Szemle* 42: 507–518 (in Hungarian).

**Erickson, Timothy, and Ariel Pakes.** 2011. "An Experimental Component Index for the CPI: From Annual Computer Data to Monthly Data on Other Goods." *American Economic Review* 101(5): 1707–38.

**European Commission, International Monetary Fund, the Organisation for Economic Co-operation and Development, the United Nations, and the World Bank.** 2009. *System of National Accounts, 2008*. <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>.

**Fixler, Dennis, Charles Fortuna, John Greenlees, and Walter Lane.** 1999. "The Use of Hedonic Regressions to Handle Quality Change: The Experience in the U.S. CPI." US Bureau of Labor Statistics, Presented at the Fifth Meeting of the International Working Group on Price Indices,

Reykjavik, Iceland.

**Gini, C.** 1931. "On the Circular Test of Index Numbers." *Metron* 9(2): 3–24.

**Gordon, Robert J.** 2015a. "Productivity, Prices, and Measurement." Presentation at BoskinFest, Stanford University.

**Gordon, Robert J.** 2015b. "The Sources of Slowing Growth in Productivity Growth and Potential Output." Presentation at the Philadelphia Fed Policy Forum, December 4.

**Greenstein, Shane, and Ryan McDevitt.** 2011. "Evidence of a Modest Price Decline in US Broadband Services." *Information Economics and Policy* 23(2): 200–211.

**Griliches, Zvi.** 1961. "Hedonic Prices for Automobiles: An Econometric Analysis of Quality Change." *The Price Statistics of the Federal Government, General Series No. 73*, pp. 137–96. Columbia University Press for the National Bureau of Economic Research, New York.

**Grimm, Alexis N.** 2016. "Trends in U.S. Trade in Information and Communications Technology (ICT) Services and in ICT-Enabled Services." *Survey of Current Business*, May, 96(5).

**Hatzius, Jan.** 2015. "Productivity Paradox 2.0." *Top of Mind*, Issue 39, p. 6–7. Goldman Sachs. <http://www.goldmansachs.com/our-thinking/pages/macroeconomic-insights-folder/the-productivity-paradox/report.pdf>.

**Hausman, Jerry.** 1997. "Valuation of New Goods under Perfect and Imperfect Competition." Chapter 4 in *The Economics of New Goods*, edited by Timothy F. Bresnahan and Robert J. Gordon.

**Hausman, Jerry.** 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.

**Houseman, Susan, Christopher Kurz, Paul Lengermann, and Benjamin Mandel.** 2011. "Offshoring Bias in U.S. Manufacturing." *Journal of Economic Perspectives* 25(2): 111–32.

**Houseman, Susan N., and Michael Mandel, eds.** 2015. *Measuring Globalization: Better Trade Statistics for Better Policy*, vols. 1 and 2. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

**International Monetary Fund.** 2004. *Producer Price Index Manual: Theory and Practice*.

**Ivancic, Lorraine, Walter Erwin Diewert, and Kevin J. Fox.** 2009. "Scanner Data, Time Aggregation and the Construction of Price Indexes." Discussion Paper 09-09, Department of Economics, University of British Columbia, Vancouver, Canada.

**Johannessen, Randi.** 2016. "Scanner Data in CPI/HICP." Presentation at ESS Modernization Workshop in Bucharest, Romania, March 16–17, 2016. <https://ec.europa.eu/eurostat/>

cross/system/files/randi\_johansenn\_the\_use\_of\_scanner\_data\_in\_the\_norwegian\_cpi.pdf.

**Jorgenson, Dale W., Steven Landefeld, and William Nordhaus.** 2006. *A New Architecture for the U.S. National Accounts*. University of Chicago Press.

**Konüs, Alexander.** 1939. "The Problem of the True Index of the Cost of Living." *Econometrica* 7(1): 10–29.

**Krsinich, Frances.** 2015. "Implementation of Consumer Electronics Scanner Data in the New Zealand CPI." Paper presented at the Ottawa Group on Price Indices, May 20–22, Tokyo, Japan.

**Krsinich, Frances.** 2016. "The FEWS Index: Fixed Effects with a Window Splice." *Journal of Official Statistics* 32(2): 375–404.

**Kuznets, Simon.** 1934. "National Income 1929–1932." Senate Document No. 124, 73rd Congress, 2nd Session. Washington, DC: US Government Printing Office.

**Lancaster, Kelvin J.** 1966. "A New Approach to Consumer Theory." *Journal of Political Economy* 74(2): 132–57.

**Landefeld, J. Steven, Brent R. Moulton, Joel D. Platt, and Shaunda M. Villones.** 2010. "GDP and Beyond: Measuring Economic Progress and Sustainability." *Survey of Current Business*, April.

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

**Lebow, David, and Jeremy B. Rudd.** 2003. "Measurement Error in the Consumer Price Index: Where Do We Stand?" *Journal of Economic Literature* 41(1): 159–201.

**Lee, Lung-Fei, and Mark M. Pitt.** 1986. "Microeconomic Demand System with Binding Nonnegativity Constraints: The Dual Approach." *Econometrica* 54(5): 1237–42.

**Liegey, Paul R., Jr.** 2001a. "Developing a Hedonic Regression Model for DVD Players in the U.S. CPI." <http://www.bls.gov/cpi/cpidvd.htm>.

**Liegey, Paul R., Jr.** 2001b. "Hedonic Quality Adjustment Methods for Microwave Ovens in the U.S. CPI." <http://www.bls.gov/cpi/cpimwo.htm>.

**Liegey, Paul R., Jr.** 2003. "Hedonic Quality Adjustment Methods for Clothes Dryers in the U.S. CPI." <http://www.bls.gov/cpi/cpidryer.htm>.

**Liegey, Paul R., Jr.** 1993. "Adjusting Apparel Indexes in the Consumer Price Index for Quality Differences." Chap. 6 in *Price Measurements and Their Uses*, edited by Murray F. Foss, Marilyn Manser, and Allan H. Young. University of Chicago Press.

**Liegey, Paul R., Jr.** 1994. "Apparel Price Indexes: Effects of Hedonic Adjustment." *Monthly Labor Review* 117(5): 38–45. <http://www.bls.gov/pub/mlr/1994/05/art6full.pdf>.

**McFadden, Daniel.** 1978. "Modelling the Choice of Residential Location." In *Spatial Interaction Theory and Planning Models*, edited by A. Karlqvist et al., pp. 75–96. Amsterdam: North Holland.

**Moulton, Brent R., Timothy J. LaFleur, and Karin E. Moses.** 1998. "Research on Improved Quality Adjustment in The CPI: The Case of Televisions." Bureau of Labor Statistics Working Paper. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.506.3183&rep=rep1&type=pdf>.

**Murphy Bonnie H., Michael Holdway, John L. Lucier, Jason Carnival, Elizabeth Garabis, and Elaine Cardenas.** 2008. "Proposal for Adjusting the General Hospital Producer Price Index for Quality Change." <http://conference.nber.org/confer/2008/si2008/PRCR/murphy2.pdf>.

**Nakamura, Leonard I., and Rachel H. Soloveichik.** 2015. "Valuing 'Free' Media across Countries in GDP." Available at SSRN: <http://papers.ssrn.com/abstract=2631621>.

**National Research Council.** 2002. *At What Price? Conceptualizing and Measuring Cost-of-Living and Price Indexes*. Panel on Conceptual, Measurement, and Other Statistical Issues in Developing Cost-of-Living Indexes, Charles L. Schultze and Christopher Mackie, Editors. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

**Nevo, Aviv.** 2000. "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand." *Journal of Economics & Management Strategy* 9(4): 513–548.

**Nevo, Aviv.** 2001. "Measuring Market Power in the Ready-to-Eat Cereal Industry." *Econometrica* 69(2): 307–42.

**Nevo, Aviv.** 2003. "New Products, Quality Changes, and Welfare Measures Computed from Estimated Demand Systems." *Review of Economics and Statistics* 85(2): 266–75.

**Nordhaus, William D., and Edward C. Kockelenberg, eds.** 1999. *Nature's Numbers: Expanding the National Economic Accounts to Include the Environment*. National Academies Press.

**Pakes, Ariel.** 2003. "A Reconsideration of Hedonic Price Indexes with an Application to PC's." *American Economic Review* 93(5): 1578–96.

**Petrin, Amil.** 2002. "Quantifying the Benefits of New Products: The Case of the Minivan." *Journal of Political Economy* 110(4): 705–29.

**Redding, Stephen J., and David E. Weinstein.** 2016. "A Unified Approach to Estimating Demand and Welfare." NBER Working Paper 22479.

**Romley, John A., Dana P. Goldman, and Neeraj Sood.** 2015. "US Hospitals Experienced Substantial Productivity Growth during 2002–11." *Health Affairs* 34(3): 511–18.

**Rosen, Sherwin.** 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy* 82(1): 34–55.

**Shapiro, Irving, Matthew D. Shapiro, and David W. Wilcox.** 2001. "Measuring the Value of Cataract Surgery." In *Medical Care Output and Productivity*, edited by David M. Cutler and Ernst R. Berndt, 411–437. *Studies in Income and Wealth*, vol. 62. University of Chicago Press.

**Silver, Mick, and Heravi, Saeed.** 2007. "The Difference between Hedonic Imputation Indexes and Time Dummy Hedonic Indexes." *Journal of Business and Economic Statistics* 25(2): 239–46.

**Stiglitz, Joseph E., Amartya Sen, and Jean-Paul Fitoussi.** 2009. *Report by the Commission on the Measurement of Economic Performance and Social Progress*. United Nations Press.

**Syverson, Chad.** 2016. "Challenges to Mismeasurement Explanations for the U.S. Productivity

Slowdown." NBER Working Paper 21974.

**Szulc, B.** 1964. "Indices for Multiregional Comparisons." *Przegląd Statystyczny* 3: 239–254 (in Polish).

**Triplet, Jack E.** 1982. "Concepts of Quality in Input and Output Price Measures: A Resolution of the User-Value Resource-Cost Debate." In *The U.S. National Income and Product Accounts: Selected Topics*, edited by Murray F. Foss, 269–312. University of Chicago Press.

**van der Grient, Heymerik A., and Jan de Haan.** 2010. "The Use of Supermarket Scanner Data in the Dutch CPI." Paper presented at the Joint ECE/ILO Workshop on Scanner Data, May 10, 2010. Geneva.

**van der Grient, Heymerik A., and Jan de Haan.** 2011. "Scanner Data Price Indexes: The 'Dutch Method' versus RYGEKS." Paper presented at the Twelfth Meeting of the International Working Group on Price Indices (Ottawa Group).