

# Expanded GDP for Welfare Measurement in the 21<sup>st</sup> Century

Charles Hulten

University of Maryland and NBER

and

Leonard I. Nakamura<sup>1</sup>

Federal Reserve Bank of Philadelphia

April 28, 2019

## Abstract

The digital revolution changed the character of economic activity in ways that are hard to measure using conventional GDP procedures. The information used by consumers to inform their choices has increased dramatically as a result of the Internet and its applications, and new mobile communication devices have greatly increased the speed and reach of its accessibility. The broad scope and rapid uptake of the digital revolution pose a challenge to GDP, but there is a larger issue: many of the benefits of the Internet and its applications by-pass GDP and go directly to consumers. This disconnect introduces a wedge between the growth in real GDP and the growth in consumer well-being, and a slower rate of growth of the former does not necessarily imply a slower rate of the latter. The conceptual framework for this analysis is developed in a previous paper (Hulten and Nakamura (2017, revised), which extended the conventional framework of GDP to include a separate technology for consumer decisions based on Lancaster (1966b), and developed the idea of Expanded GDP (or EGDP). In this paper, we use this framework to critique existing GDP and price measurement procedures and summarize the existing evidence on the size of the wedge between GDP and EGDP.

Keywords: National Accounts, Internet, Information, Inflation, Welfare

JEL No. O4, O3, E01

---

<sup>1</sup> Hulten is Professor of Economics Emeritus at the University of Maryland and Leonard Nakamura is Emeritus Economist at the Federal Reserve Bank of Philadelphia. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. We thank Kyle Brown, Carol Corrado, John Fernald, David Friedman, Erica Groshen, Fatih Guvenen, Nancy Humphrey, Brent Moulton, Jon Samuels, Dan Sichel, Rachel Soloveichik, Hal Varian, and staff members at BLS and BEA for help and advice, Jeanna Kenney provided excellent research assistance.

## I. INTRODUCTION

We are in the midst of a technological revolution of tectonic proportions, centered on the rapid advances in the generation, transmission, use, and storage of information. Schmidt and Rosenberg (2014) have termed it “the Internet Age,” an era in which “the Internet has made information free, copious, and ubiquitous”. However, its reach goes beyond the Internet, per se, to include major advances in health care and higher education and structural changes in finance and banking, and, indeed, nearly all sectors of the economy. Moreover, it is more than just a profusion of new products. The information revolution has led to major changes in the organization of firms, location of production, and the way goods and services are distributed. One result has been an increase in the wellbeing of consumers.

The question addressed in this chapter is whether the procedures currently used to measure GDP adequately capture this increase. There are good reasons to think that it does not.<sup>2</sup> The new information goods do not always play by the same “rules” as those typically counted in GDP, which is an aggregate measure of the goods and services whose value is, for the most part, determined by market transactions. Much of the information available over the Internet is not accompanied by transactions, in effect at a direct price of zero, so there is no monetary yardstick with which to estimate its value to the consumer. Thus, while some of this information does, indeed, involve economic activity supported by transactions that are captured in GDP, the direct consumer welfare value of the information is not counted as GDP.

The statistical system has also struggled with the advent of new or improved goods that deliver superior outcomes per dollar of expenditure. Improvements in the effectiveness of outcomes have occurred in a wide range of goods, from transportation and electronic equipment to health and welfare services. Even before the digital revolution the service sector posed problems for economic measurement; output is often measured in terms of expenditures rather than outcomes and, as Griliches, (1992, 1994) has noted, it is not even clear what actually constitutes output. The digital revolution has, however, increased these problems with innovations like minimally invasive surgery, which brings an enormous increase in patient comfort at a relatively small increase, or even decrease, in resource cost.

The improvement in consumer welfare is the common theme that links the measurement

---

<sup>2</sup> For example, Coyle (2014) remarked that GDP was “a measure of the economy best suited to an earlier era.”

problems associated with the “free” information and the advent of new and better goods and services. One response has been to focus on how current GDP procedures can be adapted to accommodate the range of goods involved, but this approach faces an uphill battle. The essential problem is not just about how efficiently goods and services are produced, but also how effectively they are used in consumption to generate welfare. The basic hypothesis of this chapter is that the two are not the same.

Our recent research has approached this problem by bringing consumer choice into the GDP measurement framework using the standard utility maximization framework of economic theory (Hulten and Nakamura, 2018), extending the “production” approach to GDP by adding a separate technology for the consumption of goods. It follows Lancaster (1966b), who argued that consumer utility is derived from the characteristics of bundles of goods acquired and not from the goods themselves, and that there is a *consumption technology* that transforms goods, measured at production cost, into consumption “activities” or “commodities” that provide utility. This approach allows for an explicit modeling of the wedge that may exist between the acquisition cost of the goods acquired and the resulting outcomes (as with health care), and that outcome may depend on idiosyncratic factors like the existing state of health or education, on which the outcome is contingent.

Once the consumption technology wedge is introduced into the analysis, it is but a short additional step to assume that it may shift over time as the innovations introduced by digital revolution enable consumers to make more efficient use of their incomes. We term this form of innovation “output-saving” since given levels of welfare can be achieved with fewer resources. In effect, this treats the consumption technology in the same conceptual way that Robert Solow (1957) adopted in his analysis of the productivity residual, which measured costless “resource-saving” shift in the production function. The latter describes an increase in the productivity of inputs, the output-saving innovation refers to the “productivity” of the consumption technology.

We then adapt the conventional equivalent and compensating variations of standard economic theory to measure the increase in consumer utility arising from output-saving innovation. This results in a dollar metric for the measuring benefits from innovations that go directly to the consumer in the form of “more bang for the buck.” We add this dollar metric to conventional GDP to obtain an expanded concept of GDP (or “EGDP”). EGDP provides a

natural framework for incorporating the results of empirical research on the information economy into a broader measure of consumer well-being. It allows for the possibility that aggregate economic welfare can increase more rapidly than conventional real GDP during periods of rapid innovation.

The next two sections of this chapter set out the conceptual framework and rationale for EGD<sup>3</sup>. The goal is to decompose the growth rate of EGD into its output saving, resource saving, resource using, and input accumulation components. This is essentially the conventional growth accounting framework with output-saving innovation added and costless product change reclassified as part of the consumption technology. The material that follows is then devoted to an examination of the empirical work that supports each of the sources of growth. The final section pulls together the results to address the question of whether the implied estimate of EGD may have grown faster than real GDP over the last three decades. Our estimates suggest that it did.

## II. THE THEORY OF AGGREGATE OUTPUT

### A. Gross Domestic Product and Income

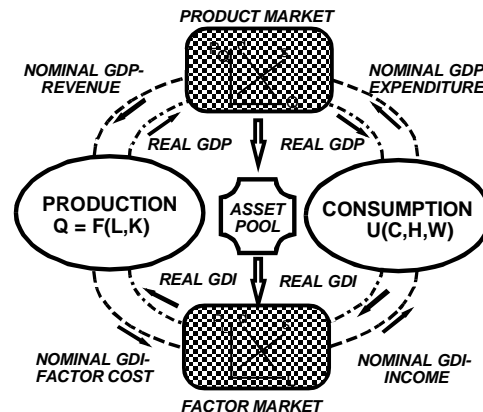
GDP in nominal prices is, with some exceptions, an estimate of the value of goods and services that flow through markets in a given year. GDP in constant prices is a synthetic concept that pulls together the corresponding quantities of goods and services. It is not a good itself, though in growth theory it is often treated as such, but an index of aggregate output whose base year value equals nominal GDP. GDP is linked to Gross Domestic Income (GDI) by the circular flows of inputs and output through product and factor markets.<sup>4</sup> The representation, shown in Figure 1, divides an economy into two basic functions: the production of goods and services, and their consumption by households, which also supply the inputs into production. The linkage between these flows is determined, in the production sector, by a production function  $Q=F(L,K)$  that links the flow of output  $Q$  to the flow of inputs of labor  $L$  and capital  $K$  via the prevailing technology  $F(\bullet)$ ; on the consumption side of the economy, the utility function  $U(C,H;W)$  transforms the output  $C$  into utility, and guides the decision of how much of the available time endowment to allocate to leisure  $H$  and how much to labor  $L$ , as well as the decision about how

---

<sup>3</sup> The technical derivations and assumption can be found in our previous paper, on which the current paper builds.

<sup>4</sup> See Patinkin (1973) for a discussion of the structure and history of the circular flow model.

much consumption should be deferred to future years by building up wealth  $W$ . The outer counterclockwise flow shows the stream of payments into and out of the two sectors as they enter and exit the markets for outputs and inputs. They indicate, in the top part of the diagram, the identity between the amount spent by consumers and the amount received by the



**FIGURE 1**

producer, which together define nominal GDP. At the bottom, the producers' factor cost is the consumers' income, defining GDI. The balancing of supply and demand in the product and factor markets establishes the equalities of the flows. To complete the picture, the revenue that flows into the business side is equal to the factor cost that flows out, and the income that flows to the consumer flows out as expenditure on products. The resulting GDI equals GDP, some \$20 trillion in the U.S. as of mid-2018.

Nominal GDP is measured in the prices prevailing in each year. It sums the product of the price of each good and the corresponding quantity, just as nominal GDI sums the product of the price of each input and its quantity. An estimate of the price change is typically used to deflate the nominal value to arrive at the corresponding relative quantity, which is represented in Figure 1 by the inner clockwise flow that tracks the movement of output and input quantities between producers and consumers. Prices are represented implicitly in Figure 1 by the intersection of the supply and demand functions in the markets for inputs and outputs. They play a central role in regulating the composition of the flow of goods. They also play a key role in efforts to introduce the benefits of new and improved goods into the circular flow representation of the economy.

The aggregate nature of GDP and GDI masks a wealth of detail in the underlying input-output structure of the economy. Thus, GDP is not a measure of the entire production of goods by the constituent sectors of the economy, since sectoral production also includes the intermediate goods delivered to other industries as inputs. The consumption, investment, government, and net exports components of GDP are “final demand” goods available for current or future consumption, domestic and foreign. This is a point that should not be ignored when assessing the impact of innovation on the economy, since when the innovation enters a sector of the economy it may appear very differently than when it impacts final demand (e.g., Hulten, 1978). Thus the advent of broadband allowed goods purchases of CDs and DVDs to be replaced by subscription streaming services. The I-O structure of the economy also implies that GDI is equal to the total value added of labor and capital and not the total cost of production across sectors, which also includes the cost of the intermediate inputs.

Household production deserves special comment given the attention it has received in the literature on the mismeasurement of GDP. One problem with accounting for household production is the conflation of the production of goods within the household with their consumption, since both occur within the home and often involve the same agents. The boundary between the production of a meal in a restaurant and the same meal produced at home by the same chef is not so much a matter of production as the method of distribution.

## B. Capital Formation

GDP and GDI are snapshots of the size of the aggregate economy at a single point (year) in time. The bulk of U.S. GDP goes for the provision of current wants, while the investment component represents the use of current resources to satisfy future consumption. Provision for future wants is, however, not explicitly represented in the traditional circular flow framework, although this need not be the case. Figure 1 shows that the traditional framework can be expanded to include the flows of investment from the product markets to a separate capital account, in which there is the producers’ stock of capital  $K$  on the one hand, and the consumer wealth  $W$  that it implies on the other. This wealth arises from the decision by consumers to defer current consumption by saving, which diverts resources away from the production of consumption goods to the production of capital goods. This investment adds to the existing capital stock, and builds the

future capacity needed to produce consumption goods in the future. The result is shown in the area in the center of Figure 1 labeled “asset pool.”<sup>5</sup>

The pool of the productive capital contains different types of tangible capital (equipment and structures) as well as intangible capital. This category of capital includes R&D, investments in product development and marketing, customer support, and human resources and organizational development. These investments are intended to develop new or better goods, processes, and markets, on the one hand, and to maintain and improve the organization and management of firms, on the other. Until quite recently, expenditures for intangible capital except computer software were treated as intermediate inputs, and thus ignored in the circular flow representation of the aggregate economy. This changed in 2013 with the capitalization by BEA of R&D and expenditures for artistic originals. This move added some 3% to 4% to U.S. GDP that had theretofore gone uncounted, but this amount accounts for less than half of the list of intangibles advocated by Corrado, Hulten, and Sichel (2005, 2009), even after adjustment for software.

### III. GDP AND CONSUMER WELFARE

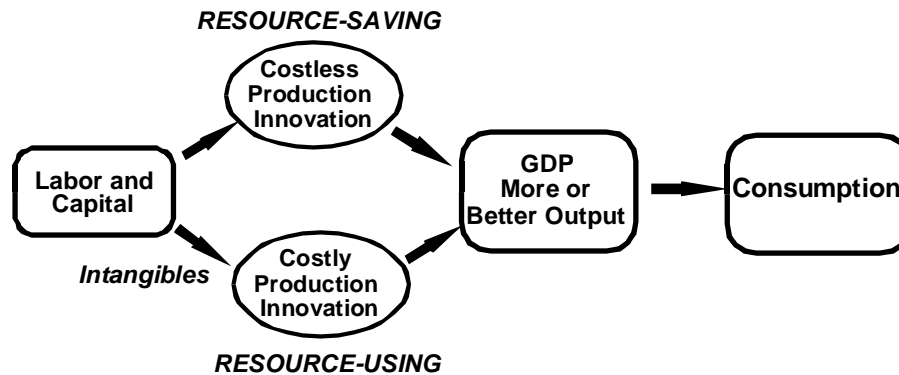
#### A. Diagrammatic Exposition of Innovation and GDP

The circular flow model is a descriptive framework that links the flow of goods and payments in the economy. The role of both the utility and production functions is to transform the flow of inputs and outputs that passes through their segments of the economy. They are treated symmetrical in this process. However, this is emphatically not the way they are treated in standard economic theory, where the maximization of utility is the objective of economic activity and the production technology is a constraint on the achievable outcome. A schematic representation of this optimization exercise is shown in Figure 2, where the first three links show labor and capital being transformed by technology into output (real GDP) via the production function. The output is then transferred to consumer through the product market, in which the volume and price of each good is determined by the interaction of supply and demand. Once the price and quantity of each good is determined, aggregate GDP follows immediately. Under the usual optimization assumptions, the resulting GDP represents the maximum attainable utility. An increase in real output  $Q$  is assumed to increase utility, and a proportional increase in  $Q$  may

---

<sup>5</sup> This figure is based on Figure 2 of Corrado and Hulten (2015).

result in an equal proportional increase in utility (but only if the marginal utility of real income equals one). In this case, a comprehensive measure of real GDP is a sufficient statistic for estimating the increase in well-being (in the sense of the utility function).



**FIGURE 2**

Innovation affects output in two ways in this setup. The production function can shift upward for a given combination of labor and capital, causing the inputs to be more productive. This is the situation envisioned by Robert Solow in his 1957 formulation of the TFP residual, in which the shift is treated as an autonomous process that is costless in terms of the need for resources (it falls as “Manna from Heaven”). It includes innovation due to inspiration and tinkering, but mainly represents knowledge spillovers, which Nordhaus (2005) argues is the primary source of macro-innovation. It is labeled “resource-saving” in the figure, due to the costless improvements in productivity it enables. The second source of innovation shown in the figure is involves systematic investment is innovation. This involves the intangible capital noted in the preceding section. Because it implies a systematic commitment of resources, it is labeled “resource-using” in the figure.

There is a further distinction between innovation that increases the quantity of output and that which increase the quality of existing goods or introduces new goods that is implicit in Figure 2. The former is typically called “process-oriented” technical change, while the latter is “product-oriented”. This is the rationale for distinguishing between “more or better output” in the GDP part of the figure, reflecting the convention that a “better” is typically expressed as “more” output for purposes of measurement, to the extent that an adjustment is actually made.



## B. GDP Expanded to Allow for Direct Consumption Benefits

Most thinking about GDP has focused on Figures 1 and 2. Indeed, Figure 2 illustrates the point at which the conventional measurement framework leaves off. However, an increase in the consumption efficiency and the increase in well-being it enables, does not fit easily in the conventional framework. To address this problem, we have proposed expanding the figure above to include a separate technology for consuming the goods obtained from producers. It follows Lancaster's 1966 "New Approach" to consumer theory in which consumer utility is derived from the characteristics of the goods acquired and not from the goods themselves, and that there is a consumption technology that transforms goods, measured at production cost, into consumption "activities" or "commodities" that provide utility.

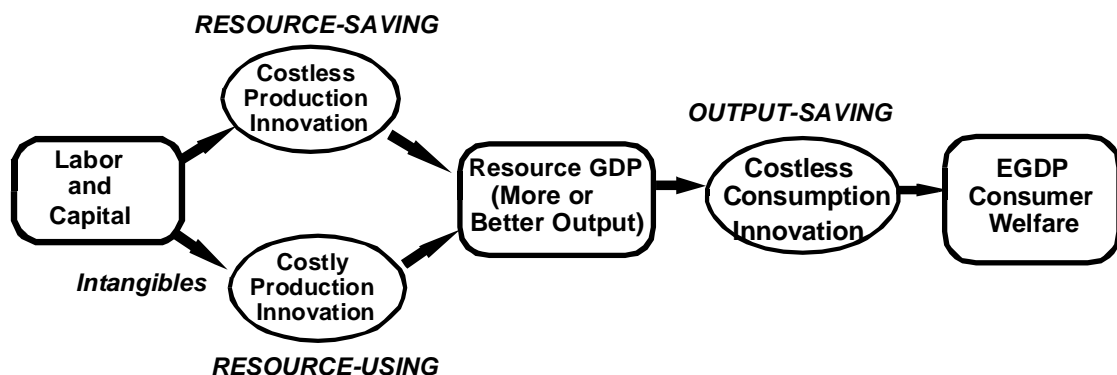
This is relevant for the issues at hand, since once the idea of a separate technology for consumption is introduced, the distinction between output and outcomes has a natural theoretical basis.<sup>6</sup> Moreover, it is reasonable to expect that the technology might change over time in ways that make consumer choice more efficient, as, for example, when an increase in information allows consumers to derive more utility from the amount of money or time expended. We term this form of innovation "output-saving" since it enables an increase in consumer welfare for the same amount of resources, or, equivalently, the prior level of welfare can be achieved with fewer resources. To fix ideas, consider a free social media app that steers drivers away from traffic jams, enabling them to reach their destinations more swiftly with less expenditure on gasoline. The app lets consumers make better driving decisions but there is no visible transaction. Without the expansion of GDP that we propose, the app shows up in GDP as a decline in output.

Figure 3 adds a consumption technology to the schema set out in Figure 2. The concept of GDP shown in the middle of the figure is now real output measured at *resource cost*. This is the output acquired at its marginal cost of production, and is the output that is transformed by the consumer into the Lancaster commodities that yield utility. Output-saving innovation operates in the link between resource output and commodity utility and is the source of the wedge between GDP growth and the increase in well-being. The size of this wedge is also affected by *costless* improvements in the quality of the resource-output transferred to the consumer. The costless feature of quality change means that the marginal resource cost of a higher quality version of a

---

<sup>6</sup> The importance of the interaction between producer and consumer is also emphasized by Peter Hill (1999).

good is zero and the benefit in terms of increased utility goes directly to the consumer, as opposed to the conventional practice of treating it as simply more of the older version product. In other words, where the conventional approach implicitly treats costless improvements in the product quality as a shift in the production function (resource-saving technical change), whereas we propose to treat it as a shift in the consumption technology (output-saving technical change).



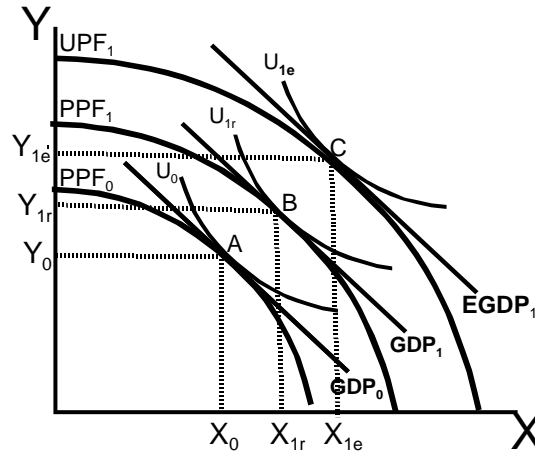
**FIGURE 3**

The expansion of conventional output from Figure 2 to Figure 3 can be formalized as a change in the utility function from  $U(Q_t)$  to  $U(c(Q_t, t))$ . The consumption technology  $c(Q_t, t)$  replaces  $Q_t$  and the time-shifter  $t$  is present in the consumption technology to allow the transformation of resource-based goods into Lancaster commodities to become more efficient over time, yielding more utility per unit output. It parallels the productivity-enhancing Manna-from-Heaven role played by the  $t$ -shifter in the Solow production function. The consumption technology  $c(Q_t, t)$  models the wedge between the two sides of the economy and introduces a conceptual richness that GDP alone cannot achieve. In addition, it can be extended to accommodate additional state variables, as in Section VIII where we discuss state contingency in health and education.

### C. The Consumption Technology and Expanded GDP

What exactly does a separate consumer technology mean for the measurement of GDP? Is there a dollar metric of the size of the output-outcome wedge? The problem is that the right hand side of Figure 3 links output in constant dollar prices to utility whose natural units are unobservable utils. However, this is a familiar problem in economic theory. The standard solution is to appeal to the compensating and equivalent variations (the *CV* and *EV*) associated with the utility

maximization problem as monetary metrics of the distance between to indifference curves on the utility function. The *CV* and *EV* are measures of the willingness to pay for moving from a lower to a higher indifference curve, thereby converting a change in utility into a monetary value whose units are commensurable with of GDP.<sup>7</sup> Figure 4 shows how this might work.



**FIGURE 4**

The production possibility frontier  $PPF_0$  for two goods,  $X$  and  $Y$ , is shown in this figure at an initial point in time ( $t=0$ ). It represents the maximal combinations of  $X$  and  $Y$  that can be produced from the labor and capital available in that year, given the prevailing technologies for producing the goods (the first three stages of Figure 3).  $U_0$  is the highest attainable indifference curve of the representative consumer, and the tangency between this indifference curve and the  $PPF_0$  constraint is located at the point  $A$ , associated with the optimal  $X_0$  and  $Y_0$ . The tangency defines the equilibrium prices,  $P^X_0$  and  $P^Y_0$ , and the line  $P^X_0X_0 + P^Y_0Y_0$ , defines  $GDP_0$ . The slope of the GDP line at  $A$  can therefore be interpreted as the ratio of the marginal costs of producing  $X$  and  $Y$ , but also as the ratio of the marginal utilities of consuming these goods.

The growth of labor and capital, plus resource-saving and resource-using technical change, causes the  $PPF_0$  to shift upward to  $PPF_1$  between periods  $t=0$  and  $t=1$ . An equilibrium is established at the point  $B$ , at a higher indifference curve  $U_{1r}$  with an amount of real  $GDP_{1r} = P^X_0X_{1r} + P^Y_0Y_{1r}$ . The subscript  $r$  is used here to denote that the quantities of  $X$  and  $Y$  are

<sup>7</sup>Since our objective is to obtain a dollar metric of output-saving innovation that can be incorporated into the conventional GDP framework, the question of how much happier the consumer feels is not a concern in this chapter. How much the consumer is willing to pay for the change in utility, is.

measured in resource units. The dollar value of the real growth occurring between the two period equals  $GDP_{1r} - GDP_0$ , and the rate of growth is  $(GDP_{1r} - GDP_0)/GDP_0$ . The allocation of this rate among the growth in the inputs and technology can be estimated using the Solow (1957) residual method.  $GDP_{1r} - GDP_0$  in this diagram is also the change in the amount of real consumption expenditure.

This is where the usual “theory” of GDP leaves off, as in Figure 3. When the Lancaster consumption technology is included in the analysis, a second source of value comes into play. An increase in the amount of information freely available for consumer choice or a costless improvement in product quality causes the utility function to shift outward to  $U_{1e}$  in Figure 4, even though output in resource units  $(X_{1r}, Y_{1r})$  remains unchanged, as do real  $GDP_1$  and prices  $(P^X_0, P^Y_0)$ . At these prices, the tangency between  $U_{1e}$  occurs at the point  $C$ . This tangency implicitly defines a new frontier labeled  $UPF_1$  to emphasize that is the utility possibility frontier associated with the production possibilities frontier  $PPF_1$ . A pair of virtual outputs  $(X_{1e}, Y_{1e})$  are defined in which the outputs are now denominated in efficiency units (hence the subscripts  $e$ ). This convention transforms the units of  $X$  and  $Y$  from the cost of the resources they embody into the units of the utility they convey. If the transformation proceeds according to the same proportion  $\theta$  for both goods, and in Figure 4, the result is  $X_{1e} = (1 + \theta)X_{1r}$ , and  $Y_{1e} = (1 + \theta)Y_{1r}$ . This is the phenomenon we have called *output-saving technical change*: an increase in utility for the same amount of resource-based output (occurring in this example at the rate  $\theta$ ).

A little algebra establishes that the shift from in utility from  $U_{1r}$  at  $B$  to  $U_{1e}$  at  $C$  is related to  $\theta$  in the following way:  $U_{1e} = (1 + \theta)U_{1r}$ , under the simplifying assumptions of Figure 4, so that  $\theta = [(U_{1e}) - U_{1r}]/U_{1r} = \Delta U/U$ . In other words, the rate of change of output-saving technical change is associated with the rate of change in utility between points  $B$  and  $C$  in Figure 4. This is hardly surprising in view of the way we have defined output-saving technical change. A more important result emerges from the fact that the line tangent to  $U_{1e}$  at  $C$  can be used to define what we have termed expanded gross domestic product:  $EGDP_1 = P^X_0 X_{1e} + P^Y_0 Y_{1e}$ . It then follows that  $EGDP_1 = (1 + \theta)(P^X_0 X_{1r} + P^Y_0 Y_{1r}) = (1 + \theta)GDP_1$ . In other words, output-saving technical change leads to a grossed-up form of real GDP as conventionally defined.

Here is where the  $CV$  and  $EV$  measures of the willingness to pay enter the analysis. Since relative prices are assumed not to change during the move from  $B$  to  $C$ , we denote the  $CV/EV$  by

$V$  and note that it is the monetary “distance” between the lines  $EGDP_I$  and  $GDP_I$ . In other words,  $V = EGDP_I - GDP_I = (1+\theta)GDP_I - GDP_I$ , from which it follows that  $V = \theta GDP_I$  and that  $\theta = V/GDP_I$ . This result is significant for the issues at hand because it shows that the unobservable rate of output-saving technical change,  $\theta$ , is potentially observable through willingness to pay for  $\theta$ , denominated in monetary units.

#### D. Information and Product Quality Change as Sources of Output-Saving Innovation

The rationale for output-saving innovation has thus far been presented largely in terms of the benefits of increased information for efficient consumer choice, and the associated  $V$  as a monetary metric of those benefits. However, the output-saving effect is more general in its scope. Two types of the output-saving technical changes can be distinguished: first, product-*disembodied* innovation,  $\mu$ , which includes the benefits of increased information, but also includes costless improvements in outcomes in the provisions of many services (e.g., improvements in convenience, the diffusion of best-practice techniques in the service sectors); and, second, product-*embodied* innovation in consumption goods,  $\beta$ , occurs as superior varieties of existing goods enter the market with superior characteristics, or new goods appear. The sum of the two equal the rate of output-saving innovation, i.e.,  $\theta = \mu + \beta$ .

The shift in utility from  $B$  to  $C$  in Figure 4 can occur as a result of either  $\mu$  or  $\beta$ . In both cases, the same output bundle is measured in resource units,  $(X_{Ir}, Y_{Ir})$  to pack more utility per unit of real income, leading to the efficiency-unit concept of output  $(X_{Ie}, Y_{Ie})$ . Both give rise to a monetary metric  $V$  based on the willingness to pay for the enhanced utility per unit of resource-based output. There is, however, a significant difference in the way the measurement of  $V$  can be approached. In the product-*disembodied* case, there is no observable transaction price with which to value the increment in utility brought about by  $\mu$ . In the product-*embodied* case, however, the transaction price of a superior new variety of a product is observable and, in some cases, the price of the older inferior variety can be also be observed and used as the basis for estimating the extent of the improvement, for example, using hedonic regression. We return to this issue in section VI below.

### E. A Closer Look at the Role of Prices

In both the product-embodied and the product-disembodied cases, because the change in utility is assumed to be costless, the amount of money spent to purchase  $(X_{lr}, Y_{lr})$  is the same as the amount associated with  $(X_{le}, Y_{le})$ , so that  $P^X_0 X_{lr} + P^Y_0 Y_{lr} = P^{Xe}_0 X_{le} + P^{Ye}_0 Y_{le}$ . The  $P^{Xe}$  and  $P^{Ye}$  are the shadow prices of  $X_{le}$  and  $Y_{le}$ , and are also denominated in efficiency units (we assume, here, that there is no pure price inflation so the accounting can be done in base-year prices). Since the amount of expenditure  $P^X_0 X_{lr}$  is equal to the notional  $P^{Xe}_0 X_{le}$ ,  $P^X_0 X_{lr} = P^{Xe}_0 X_{le}$ , it thus follows that  $P^X_0 / P^{Xe}_0$  is equal to the  $X_{le} / X_{lr}$ , which in turn equals  $(1 + \theta)$ . Thus as utility increases by the factor  $\theta$ , the cost of acquiring this utility falls. We will revisit these equations in the sections that discuss the various methods that have been applied to the problem of measuring the embodied quality change  $\beta$  component of  $\theta$ .

One further point is important here. Because output-saving technical change means that each dollar spent on either good “buys” more utility, this increase would normally imply that more of the good subject to technical change would be demanded, and because the resource cost of the good has not changed, the quantity demanded would increase to the point at which the gap between the marginal utility and price would be extinguished. However, the opportunity for this arbitrage does not exist in all cases. When a superior pharmaceutical drug arrives in the market place, the individual consumer does not respond by buying more of the drug, instead by replacing the older less effective alternative. There may be some increase in demand in cases when a larger number of people can now be treated, but people normally do not take pleasure in consuming most drugs (with obvious exceptions like opioids). A medical shock, like an accident or health crisis, leads to a demand for a medical intervention and if a superior alternative becomes available, it will usually replace the old procedure. As for information, it too tends to be case specific, as, for example, when looking for the answer to a specific question, or looking for a superior alternative to current consumption choices. Nor do people necessarily usually purchase more personal computers as their efficiency increases. The key point is that there are many situations in which the market mechanism does not arbitrage the benefits of innovation, and in this case, there will be a gap between the goods measured at cost of acquisition and the corresponding benefits received, and this gap may persist giving rise to output-saving innovation.

## F. Quality Change in Capital and Intermediate Goods

The product quality modeled in Figures 3 and 4 refers to consumer goods that enter the utility function. However, these are not the only products subject to quality improvement. Quality change in the intermediate goods used in production may occur over time, as, for example, in the advent of superior manufacturing and construction materials. They appear as more input when an adjustment for quality is made or as part of the shift in technology when it is not, because their use makes production more efficient. In either case, the result is more output and it appears as shift the PPF in Figure 4 outward from  $GDP_0$  to  $GDP_1$ .

Product quality change also occurs in capital goods, where it is called “capital embodied technical change.” Capital is an input that delivers productive services over time, and thus functions as an intertemporal intermediate good (Hulten, 1978). Like conventional contemporaneous intermediate goods, capital embodied technical change in one year thus increases the output, but the increase occurs in multiple years in the future. It is modeled in Figure 1 only as an increase in the stock of capital whose value also constitutes an increase in wealth. The current level of production is affected by the existing stock of capital built up over past investments. A change in the quality of investment-good output is thus reflected only in the total of effective capital that enters subsequent production functions. Modeling this adequately requires an intertemporal production model that goes beyond Figure 4. However, the simplified model set out in Figures 3 and 4 is intended to illustrate how the conventional accounting framework portrayed in Figures 1 and 2 can be expanded to include a Lancaster consumption technology and the associated output-saving technical. A richer intertemporal treatment would include these extensions.

## IV. ESTIMATES OF GDP AND WELFARE: AN OVERVIEW

The Industrial revolution and its aftermath have resulted in a dramatic increase in income. Angus Maddison’s 2007 estimates of world GDP since 1700 suggest that real world GDP per capita increased by almost nine-fold over the period 1700-1998, with most of the increase coming during the later stages of the Industrial Revolution. Moreover, the increase from 1700 to 1998 was by far the largest in the countries that led that revolution. The increase in the countries of Western Europe was nearly 18-fold, and that in the United States over the shorter 1820-1998 period was estimated to be 22-fold, leaving the rest of the world far behind. Moreover, estimates

of real GDP per capita in the National Accounts (Table 7.1) show that real GDP per capita has increased by over 250% from 1950 to 2017, and by around 50 percent from the inception of the Internet in the early 1990s to 2017.

The centuries since the start of the Industrial Revolution also witnessed extraordinary improvements in the wellbeing of individuals. The world of 1820 lacked effective medical treatments for many, if not most, afflictions. The discovery of the germ theory of infection by Lister was a major step forward, ultimately persuading surgeons they should wash their hands prior to surgery. The development of effective forms of anesthesia was also a huge advance in medical treatment (it is hard to imagine, today, surgery without it). Antibiotics in the twentieth century allowed routine infections to be treated that previously led to many deaths. Similarly, the development of vaccines brought fearsome diseases like smallpox, diphtheria, tetanus, yellow fever, and polio more or less under control, with enormous increases in human wellbeing. The medical revolution proceeds apace with important breakthroughs in surgery (non-intrusive, robotic, and nano). Diagnostic procedures have evolved from the simple X-ray (a breakthrough in its day) to CT scans and MRIs. These innovations have had a major impact on life expectancy, which increased from 48 years to 78 years over the course of the 20<sup>th</sup> century. How much GDP would society be willing to sacrifice in order to protect these gains?

Significant increases in welfare also occurred in other areas. The first half of the 19<sup>th</sup> century was a period without electricity, flush toilets, central heating, telecommunications, and automobiles and aircraft. The growth in labor-saving home appliances, like automatic washing machines and refrigeration, brought large and direct gains in the wellbeing of families, as did residential air conditioning. Many advances have come since the mid-20<sup>th</sup> Century. As recently as 1950, a quarter of America's homes had no flush toilet, according the U.S. Census Bureau housing data. In 1990, only one percent of our homes lacked complete plumbing facilities, but in 1940, nearly half lacked complete plumbing. Improvements in sanitation were also important in increasing public health. In 1960, about 1-in-5 households had no telephone available. Wood was used as a major heating fuel in 1940 (23%), but virtually disappeared by 1970 (only 1.3%). Robert Gordon (2016) has chronicled the gains in welfare that arose from many of these innovations.



The rapid uptake of digital goods is significant in this regard. According to Census estimates, the portion of U.S. households with a computer at home rose from about three-quarters in 1993 to almost four-in-five in 2012, and the fraction with Internet use at home went from one-in-five in 1997 to nearly three-quarters in 2012. Moreover, estimates by the PEW Research Center show that the percentage of adults who use at least one social media site increased from less than one-in-ten in 2005 to two-thirds in 2015, and other PEW surveys found that the market penetration of smart phones more than doubled from 2011 to 2016, from 35 percent to 77 percent.<sup>8</sup> The rapid uptake was matched with a dramatic increase in speed and capacity. In 1988, Internet speeds on dial-up modems were 9.6 Kb, while 2G cellular speeds were about the same. Now broadband speeds up to one Gigabit are available in a few locations, and 100 Mb and higher speeds are widely available. And 4GLTE cellular speeds are 100 Mb, and these too are in wide use. Over a 27 year period, from 1988 to 2015, speeds have gone up some ten thousand times, or a 40 percent annual rate.

The question raised in this paper is whether the gains in individual wellbeing are fully valued by the corresponding gains in income per capita. The average individual is now able to purchase not only more goods but also much better goods than two hundred years ago. The large increase in real GDP reflects the increase in welfare from goods that consumers actually purchased “out-of-pocket.” The question addressed in this paper is whether there was an additional willingness to pay to these goods beyond the transaction prices that were paid, an additional willingness to pay in order to avoid being deprived of these goods. In other words, is there gap between GDP and EGDP in Figure 4, and if there is a gap, what is its magnitude?

The remaining sections of this chapter provide a more detailed look at the link between the growth in real GDP per capital and the growth in consumer wellbeing and EGDP - in other words, a look at the various segments of Figure 3 with a view to assessing their potential magnitude and the implied biases *vis a vis* current statistical practice. We focus on the role played by the different sources of innovation identified in Figure 3: those that are resource saving (which, for empirical purposes, we assume proceeds at a rate  $\lambda$ ), those that are resource

---

<sup>8</sup> U.S. Census Bureau (2014), Andrew Perrin (2015), Pew Research Center (2017), Monica Anderson (2015).

using (at the rate of intangible capital growth), and those that are output saving (with its product-disembodied,  $\mu$ , and embodied,  $\beta$ , components).

These various components come together to form the basic framework linking the growth in welfare per capita,  $u - \ell$ , to the growth in output be worker and the parameters of output saving innovation:

$$(1) \quad u - \ell = \mu + \beta + (q - \ell) .$$

This indicates that the representative person's welfare depends on both the amount of income they have and how well they use it. The goal is to look at the empirical evidence on the magnitude of each of these three components in turn.

## V. THE SUPPLY-SIDE CONTRIBUTION TO OVERALL GROWTH

### A. The Sources of Growth Model

Following Solow (1957), the growth in the  $(q - \ell)$  component of equation (1) can be further decomposed into its causal factors, which are the contributions of labor, capital, and innovation. The theoretical model underlying this decomposition is based on a production function whose technology uses labor and capital of different types. Under certain assumptions (constant returns to scale, competitive pricing), the growth rate of aggregate output per unit of labor input (labor productivity) is the sum of its sources:

$$(2) \quad q^r - \ell = \lambda + v_K (k - \ell) + v_N (n - \ell) .$$

Small letters denote rates of growth of output ( $q$ ), labor ( $\ell$ ), productive tangible capital ( $k$ ), intangible capital ( $n$ ); the variables  $v_E$  and  $v_N$  are the input shares in total income that proxy for the corresponding elasticities of output in the standard Solow-Jorgenson-Griliches sources-of-growth framework. The  $\lambda$  measures the resource-saving technical change of Section III, while  $v_N (n - \ell)$  is a measure of resource using innovation.

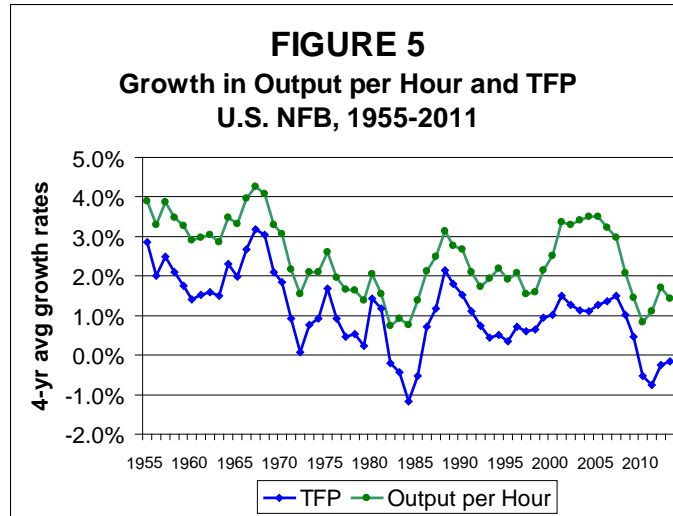
In conventional national accounting practice, real output is adjusted for the rate of change of product quality,  $\beta$ , leading to the growth rate of efficiency-adjusted output  $q^e = \beta + q^r$ , the latter denoting output growth in units of resource cost. The  $q^e$  is the efficiency-adjusted output of section IVD, and in this alternation formulation, the sources of growth equation becomes

$$(3) \quad q^e - \ell = [\beta + \lambda] + v_K (k - \ell) + v_N (n - \ell) .$$

It is important to note that, in this form, the innovation is no longer just the resource-saving innovation term  $\lambda$ , but now includes the costless product innovation term  $\beta$ . It essentially treats efficiency-output as though it were equivalent to having more resource-based output and thus regards supply-side innovation as encompassing both. All the elements of this equation are observable except the composite innovation term  $[\beta + \lambda]$ . This composite term can, however, be estimated as a residual once the observable elements are assembled (the TFP residual). In the case of equation (2), the residual measures only  $\lambda$ .

The sources-of-growth results for the U.S. *private business economy*, based on (2), are shown in Table 1 for the period 1948 to 2007. A version of the sources-of-growth model is presented in this table, derived from studies of Corrado and Hulten (2010, 2015), where it is shown that the annual growth rate of private business efficiency-output per unit of labor over the period 1948 to 2007 averaged 2.4%. The sources of this growth are reported in the rows of Table 1, which correspond to the elements on the right-hand side of (2) (with the addition of a term that corrects for changes in the composition of the labor force, due largely to increased educational attainment). For the period as a whole, this decomposition reveals that the deepening of tangible capital accounted for 27% of the 2.4% output growth, of which 10% came from ICT equipment per worker hour. Intangible capital contributed another 17%, of which only 4% from formal R&D. Changes in the composition of the workforce added another 8%, while the TFP residual explained by the other sources made the largest contribution at 47%.

These estimates refer to the period as a whole. A look at the sub periods reveals some important within-period trends. It is significant for the taxonomy of innovation presented in Section III that the long-term trend in TFP moved downward since the 1960s: TFP grew at an average annual rate of 1.8% over the period 1948-65 and explained almost half of the growth rate of output per worker hour; the growth rate fell to 1.2% in the most recent period, 1995 to 2007, and its contribution to output growth fell from 60% to just over 40%. The declining trend in TFP is also evident in Figure 5, which plots the time trend in the four-year moving average over the slightly longer period up to 2011 (because of the moving average, the initial year shown is 1955). The growing gap between TFP and output per worker hour indicated a declining relative contribution of TFP to the latter.



However, while the trend in TFP is downward, the contribution of intangible capital deepening,  $v_N(n-\ell)$ , shown in Table 1, followed a generally upward trend. An important implication of these contrasting trends is that there has been a shift away from costless resource-saving innovation (augmented by the product-quality part of what we have terms output-saving innovation) toward costly resource-using innovation, as represent by  $v_N(n-\ell)$ . The sum of the two has not changed all that much, but the welfare implications have. Resource-saving innovation is a “free lunch” in terms of the direct increase in welfare, while resource-using innovation represents a sacrifice in consumption. The free lunch is the better alternative from the welfare standpoint, but it is really not a choice variable. On the other hand, it is no great surprise that as technologically complexity arose, innovation requires more than serendipity to be sustained, hence the increased importance of systematic and focused investments in innovation and the associated equipment and learning.

The resource-saving and resource-using technical change are not the only source of innovation. ICT equipment has been an important coinvestment of intangible capital during the digital revolution, as has the increase in the composition of the labor force toward more educated and highly skilled workers. When the growth in the contribution of human capital is combined with the ICT term, and then added to the intangible capital term, the result shows a substantial change from the period 1948-1973 to 1995-2007, from a 0.56% in the earlier period (19% of overall growth), to 1.31% (a 47% contribution). Thus, although the relative contribution of TFP has declined, innovation and its correlates have not, although the composition has changed.

## B. CRITIQUE OF THE GROWTH ACCOUNTING RESULTS

Growth accounting produces estimates that by far the most secure results in the empirical chain linking resources and technology to EGDP in Figure 3. They are supported by national accounting data are assembled by the BLS in its official productivity estimates. They are, however, inevitably not without problems. Indeed, Abramovitz (1956) famously noted that the TFP residual is, in a sense, a “measure of our ignorance” since it sweeps together all the factors that affect output growth that cannot be measured explicitly. These include not only the effects of costless advances in technology but also non-technological factors like the omitted variables (infrastructure capital, the intangible capital categories not included in BEA estimates), non-market but resource-using output from household production, and chronic biases in the estimation of service sector output. And, even if the TFP residual were accurately measured, there is still the identification problem of sorting out the separate magnitudes of  $\beta$  and  $\lambda$ .

There is also a troublesome identification problem arising from the failure to account adequately for the effect of fluctuations in aggregate demand on the intensity of use of labor and capital. Capital is measured as a stock of accumulated past investment (adjusted for depreciation) rather than as flow of actual services emanating from the stock. The stock itself does not change much during fluctuation in demand, but the flow of productive services does and the degree of capital utilization changes over the business cycle. As a result, the gap between the stocks and flows is forced into the residual measure of TFP, causing the pro-cyclicality of TFP seen in Figure 5. It is for this reason that the time period covered in Table 1 stops at 2007, the year before the Great Recession. Thereafter, TFP growth dropped significantly and, indeed, turned negative, indicating a contraction in the level of productive efficiency.

A negative growth rate of TFP is plausible during sharp downturns in economic activity, but it is hard to reconcile with its conventional interpretation as an indicator of technical change over longer periods of time. However, this is precisely what happens in some individual industries, notably those engaged in the production of services. Another part of the BLS productivity program presents growth accounting estimates for individual industries in the U.S. economy based on a variant of (3) in which output, gross of deliveries to other industries, is decomposed into the share-weighted contribution of the inputs, now expanded to intermediate inputs obtained from other industries. The concept of  $\lambda$  at the industry level, and the estimate of

residual TFP, reflect changes in the efficiency with which gross output is produced. The resulting TFP growth is found to be zero for the service sector (NAICS industries 54 through 81) over the period 1987 to 2015. It is actually negative for the shorter period 1987-2007. Moreover, the TFP annual growth rate is negative for the entire 1987-2015 estimates for some service subsectors: Educational Services (-0.5%), Ambulatory Health Care (-0.4%), Hospitals, Nursing, and Residential Care (-0.9%), Management of Companies and Enterprises (-0.4%), Legal Services (-0.3%).

It is possible that lower productivity is inherent in the production of services, and they possibly suffer from Baumol's Cost Disease, although this is controversial and, in any event, refers to labor productivity (output per unit labor) and does not envision negative productivity change.<sup>9</sup> Indeed, negative productivity growth over three decades is highly implausible, and all the more so when it is recognized that these decades span the Digital Revolution. To emphasize this point, if the level of TFP in education were indexed to 100 in 1987, the index would fall to 87 in 2015. For Hospitals, Nursing, and Residential Care, the index in 2015 would fall to 77. This indicates a drop in TFP in education and health care of a large magnitude that would certainly have been noticed "on the ground" had it actually occurred.

While the Baumol explanation may play a role, the dominant factor explaining a prolonged period of negative TFP is most likely output mismeasurement. The mismeasurement explanation was discussed by Zvi Griliches (1994), who observed that "The conceptual problem arises because in many services sectors it is not exactly clear what is being transacted, what is the output, and what services correspond to the payments made to their providers (p.7)." He thus labeled the industries we are discussing as hard-to-measure industries. A consequence is that here is no agreement as to the units of measurement that underlie output of some services, and current procedures may not even be getting the resource-based  $Q'$  right, much less the efficiency-

---

<sup>9</sup> The Baumol Disease explanation of the productivity was challenged by events after the first productivity slowdown. Triplett and Bosworth (2004) found the services were not that much a drag on overall growth. Looking at a longer period than Griliches, they report a speed-up in services relative to the goods producing sectors: labor productivity in the former rose from an average annual growth rate of 0.7% percent during the 1987-1995 period to 2.6% the years 1995-2001; for the latter, the corresponding numbers were 1.8% and 2.3%, respectively. They also find that 80% of the increase in the overall growth in output per unit labor after 1995 was due to ICT's contribution to the service sectors, contrary to the hypothesis that services were inherently resistant to productivity change. However Sichel (1997) argues that only a limited amount of the productivity slowdown can be attributed to the change in industrial composition per se.

based  $Q^e$ . However, price deflators are also part of the problem, for, as he observed in 1992, there are a “number of service industries series ... deflated by makeshift deflators.”

The Griliches statement touches on one of the key ideas modeled in our framework: that consumer outcomes are different from produced output, and output is different from the expenditures. These measurement issues are echoed in Cutler and Berndt (2001), who point to what they have called the “output movement” in health economics, which attempts to measure the impact of medical care on health outcomes rather than the amount of resources expended. In the case of output and productivity of the education sector, Triplett and Bosworth (2004) summarized the proceedings of their April 2000 Brookings-sponsored workshop observed that “there was very little agreement on how to develop strong quantifiable measures of either output or productivity. Particular concerns were expressed about how to adjust for variations in education quality (page 286)”.

In defense of the BLS program, the BLS website that presents the non-manufacturing industry productivity estimates contains the following disclaimer: “Output and the corresponding inputs for nonmanufacturing industries are often difficult to measure and can produce productivity measures of inconsistent quality. Customers should be cautious when interpreting the data”.<sup>10</sup>

### C. Problems with Measuring Intangible Capital

We have thus far focused on problems with the estimates of TFP, but there are also problems associated with the intangible capital terms in (3) and Table 1. The intangible capital term,  $v_N(n-\ell)$  is a proxy for resource-using innovation, but it too is subject to measurement error. Intangible capital tends to be produced within an enterprise on an own-account basis, and thus does not generate a market transaction from which a value, and thus a market price and quantity, might be inferred. Instead, much of our information about this kind of capital is obtained from general surveys, or from imputations with a large scope for error. The BEA moved in 2013 to stop treating R&D and artistic originals as within-firm intermediate goods which do not find

---

<sup>10</sup> Multifactor Productivity and Related KLEMS Measures from the NIPA Industry Database, 1987 to 2016 (<https://www.bls.gov/mfp/mprdownload.htm>).

their way into GDP and to start capitalizing intangible expenditures and thus treat them as investment goods that are part of GDP.

Software had been represented in the national accounts since 1999, but even the list of intangibles included by the BEA, and presented in the BLS productivity estimates, falls short by about one-half of the longer list in the taxonomy developed by in Corrado, Hulten, Sichel (2005, 2009). The estimates in Table 1 are based on an updated version the Corrado-Hulten-Sichel framework, and thus differ from those presented in the BLS productivity tables which include only a partial list. One consequence of capitalized intangibles is that the relative importance of TFP as a source of growth falls from 50% to 39% when moving from the BLS TFP estimates to the fuller list (Corrado and Hulten (2014), Table 3). Another consequence is that the results investment is added to GDP, which increase it size but has only a modest effect on the growth rate of output.

## VI. PRODUCT QUALITY CHANGE AND NEW GOODS

### A. Output-saving technical Change and the Quality Measurement Problem

We have thus far concentrated in this chapter on the supply-side problems of measuring the sources of output growth. In the following sections, we turn to consumption technology and the problem of measuring output-saving technical change. We have distinguished between two components of this technical change: one that is embodied in new products and represented by  $\beta$ , and another that is disembodied and as denoted with  $\mu$ . The former appears as a new good in the market place or a superior variety of an older good; the latter refers to an innovation that allows existing goods to be used more efficiently. Both are assumed to occur without a resource cost, as with the Manna from Heaven assumption of TFP. The distinction between  $\beta$  and  $\mu$  is not fundamental to the underlying idea of a consumption technology –both have the same effect on this technology - but the distinction is important for measurement, since much of the relevant empirical work has concentrated on  $\beta$ .

The nature of the dichotomy is illustrated by the following two examples. The incredible increase in the computing power following Moore’s Law meant that each successive generation of computers was more effective in its use than its predecessor, and occurred at a minimal increase unit cost. Each new model thus packed more “bang for the buck,” or put differently, the implicit price per unit of computing,  $P^e$ , fell while the price of the actual computer box,  $P$ , did



not. The growing wedge between the two prices was driven by the rate at which product quality changed (formally,  $P^e = P/(1+\beta)$ , or  $\beta = (P-P^e)/P^e$ ). However, one consequence of the computer revolution (and, more generally, the digital revolution) also enabled an enormous increase in the information available to consumers over the Internet, making their choice processes more efficient and allowing their income to produce more bang for the buck. This is the province of parameter disembodied output-saving parameter  $\mu$ . Part of the effectiveness of increasing computer power is thus linked to its use, and this may be hard to disentangle.

The advent of minimally-invasive major surgery is another example of the problem. This advance offers many patients a significant increase in physical comfort and reduced recovery times, but these patient outcomes are not priced into the cost of the procedures. The unpriced welfare gain thus represents an increase in  $\mu$ , but could also be seen as a costless improvement in the quality of surgery. However, despite ambiguity involved in both examples, we will discuss the measurement of  $\beta$  and  $\mu$  separately in the sections that follow. We do this in part because the two involve different literatures and measurement techniques.

## B. The Quality Measurement Problem

The problem of measuring product quality change is one of the most heavily studied issues of measurement statistics, with three blue-chip panels presenting assessments of the degree of product quality bias in official price indexes and recommending solutions: the 1961 Stigler Commission, the 1996 Boskin Commission, and the 2002 Schultze Commission. Major assessments of the procedures used by BLS and BEA have been published by members of those agencies (Moulton and Moses (1997), and Groshen, Moyer, Aizcorbe, Bradley, and Friedman (2017).) There is, in addition, a large academic literature. The overall thrust of these efforts is a consensus (though perhaps a weak one) that price statistics have been, and still are, subject to a variety of measurement biases, and the main question is about the magnitude of the biases.

The fact that biases have lingered over many decades is a testament to just how difficult the problems are. Indeed, Shapiro and Wilcox (1996) called quality change “the house-to-house combat of price measurement”, and argued that “There is no simple formula that one can apply to deduce a magnitude of the problem, nor any simple solution. Unfortunately, there is no substitute for the equivalent of a ground war: an eclectic case-by-case assessment of individual

products” (p. 124). When quality change is rapid and takes novel economic forms, new challenges emerge and may grow faster than solutions.

The measurement problems are challenging, but the nature of the problem is conceptually straightforward, as the theoretical discussion of product quality in Section III showed. The problem is to estimate the  $\beta$  implicit in  $P^e = P/(1+\beta)$ . Unfortunately, this is no easy matter, since only the observed transaction price  $P$  is observed directly over time, leaving the other two variables to be identified from the one equation. The following examples reveal the nature and extent of the problem. Consider, first, an improvement in product quality that implies that \$1 spent on a particular type of good now buys the equivalent of  $(1+\beta)$  units of the good, whereas in the past only one unit was obtained. In other words, the number of units measured with respect to their efficiency increased by  $(1+\beta)$  and the efficiency price  $P^e$  decreased by the same amount. Assuming that there is no pure price inflation and the marginal cost of the actual unit transacted has not changed, the observed price of the new unit remains constant at  $P$ . Observing that price over time therefore does not reveal the quality change that has occurred. Other procedures must be used to get at the  $\beta$ .

The overlap method is one possibility, which arises when the older inferior variety of a good lingers in the market. The previous year’s model, the one that does not embody the  $(1+\beta)$  improvement, can linger in a competitive market only if its price is discounted by the factor  $(1+\beta)$ . A comparison of this discounted price with the price of the improved variety, the unchanged  $P$ , will identify the  $(1+\beta)$ . Another approach is to use the hedonic price technique, which treats a good as a composite of its various characteristics. An improvement in the overall quality of the product is then associated with an improvement in one or more of the characteristics. If a price has previously been computed for each characteristic using hedonic regressions analysis, the results can be used value the new set of quality-enhanced characteristics, thus resulting in an estimate of the overall efficiency price and quantity.

Mention must also be made of product innovation that occurs in the form of goods that are new to the market place. A problem arises when attempting to measure the value of the new good because there are no prior versions on which to base a price estimate. The new good must thus be estimated without this information, and the conventional solution to this problem has been to use an explicit demand function to estimate the price of the new good that would have induced consumers to demand a zero quantity (the Hicks-Rothbart “reservation” price). Goods

distributed free of direct charge, such as many Internet applications, pose an additional problem since the price before and after their introduction is zero.

### C. The BLS Price Measurement Program

The BLS is the government agency charged with the bulk of the Shapiro-Wilcox house-to-house combat in the price measurement battle. It is tasked with the duty of preparing a monthly report on the prices consumers pay for a sample “basket” of goods, with the general objective of determining how much the cost of living has increased due to monetary price inflation. Price inflation erodes the “bang for the buck” of each dollar of income, and the Consumer Price Index indicates (in principle) how much additional income is required to maintain the average consumer at the previous period’s level of utility if nominal income were not to change. The CPI can thus serve as a cost-of-living adjustment for wage and other contracts and government benefit programs, but it also measures the general rate of price inflation in consumer goods and the erosion in purchasing power that implies. However, an improvement in product quality provides more “bang for the buck” of each dollar spent and offsets the inflationary erosion, and must be taken into account.

One implication of product innovation is that the same basket of goods cannot be priced repeatedly over a period of time when new, and sometimes superior, goods enter the market place and find their way into the basket, and other are driven out of the market by innovation. The agents assigned to go out each month to price these goods in a retail outlet are often confronted with the problem of finding alternative items to price. The procedures they follow are described in Chapter 17 of the BLS Handbook of Methods.

The prescribed procedures are complicated and not easy to summarize. Fortunately, the survey by Groshen et al. gives an excellent and up-to-date overview of the program. When an item that was priced in the preceding month goes missing, the agents look for a similar item with which to replace it in the sample. This matched-model approach is the “cornerstone” of the CPI program. Groshen et al. report that, for the period 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.” (pp. 190-191). These percentages are on a monthly, not annualized, basis. They go on to say that:

“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” (p. 191).

The nature of the quality adjustments made to the prices of the missing two-fifths is one of the salient questions about the CPI’s ability to account adequately for product innovation.

According to the CPI Chapter 17 in the BLS Handbook, the adjustment involves an imputation procedure:

“Imputation is a procedure for handling missing information. The CPI uses imputation for a number of cases, including refusals, inability to collect data for some other reason (the item may be out of season), and the inability to make a satisfactory estimate of the quality change. Substitute items that can be neither directly compared nor quality adjusted are called noncomparable. For noncomparable substitutions, an estimate of constant-quality price change is made by imputation. There are two imputation methods: Cell-relative imputation and class-mean imputation” (p. 20).

It is these last two imputations that are the source of much controversy. When a new good like the cell-phone or the ATM arrives in the market place, it is assigned a price that reflects the average price change of the goods in the product class to which it is assigned (or the average price of a subset of goods in the class). In other words, the technological innovations embodied in wholly new goods are incorporated with a procedure based on the price of goods that do not embody the innovation.

This problem extends to the rotation of items into and out of the sampling frame. The BLS Handbook states, page 12 of the CPI Chapter 17, that

“To enable the CPI to reflect changes in the marketplace, new item and outlet samples are selected each year, on a rotating basis, for approximately 25 percent of the item strata in each PSU [primary sampling unit].”

This rapid substitution is a very welcome feature of the price program because it allows new goods to enter the CPI sample, including those that embody innovative new technology. Overlap procedures are used in incorporating the rotated sample into the index.

The price hedonic method is another way that quality and sample composition issues are handled in the CPI.<sup>11</sup> Groshen et al. report that “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. Housing-related expenditures account for most of this share” (page 192).<sup>12</sup> These statistics suggest that very few item categories are subject to the hedonic method, despite recommendation of the Stigler Commission (1960) review of price measurement that specifically referred to the Griliches study of hedonics in new cars. For its treatment of the price of cars, the BLS uses a measures of the resource cost of new car features rather than hedonic measures of the value of car features, in both the PPI and the CPI programs. In this method, the costs of new options added to the standard light vehicle are removed from new car prices in estimating inflation. Recent decades have been a period of remarkable technological innovation in autos, often at relatively low cost per automobile, using sensors, computer power, and software to improve driving. These improvements include safety warning signals, enhanced cruise control, self-parking, and backup vision. The BLS uses the cost method primarily for autos.

#### D. Critique of the CPI and the Problem of Quality Measurement

The totality of the CPI program is enormous, given the huge number of items in the universe of all consumer goods and services. It is all the more impressive because the process must be repeated month after month, without fail. And, this is far from the only BLS program, since they are also responsible for many other data collection programs. Moreover, it accomplishes its main mission: to provide a timely cost-of-living adjustment that is accepted by those affected by the outcome. This political economy aspect is perhaps its most important feature, given the large

---

<sup>11</sup> The basic idea of price hedonics is to regress the observed transaction price of a sample of goods on a set of characteristics to estimate the shadow price of each characteristic. The price of a bundle with more, or different, characteristics can then be estimated and, by extension, the price of a bundle that possesses more characteristics. Computers are a prime example. Here, the unit price of a new model of computer that embodies a faster processor speed, better graphics, and more memory often remains more-or-less the same (controlling for inflation) as the preceding inferior model.

<sup>12</sup> The hedonic regression for housing-related expenditures estimates the rate of deterioration of rental units over time so the reported inflation rates are higher than the rate of rental price increase to account for the worsening quality of the rental unit over time.

transaction costs involved in bargaining and renegotiation that would need to occur in the absence of an acceptable price index (indeed, this was the genesis of the CPI). To accomplish its mission, the BLS must contend with the dynamic nature of the economy, which continually throws up new challenges for the CPI.

The use of its price statistics by the deflation of nominal output to obtain the real inflation-corrected data for the national accounts is an important by-product, but not the main BLS mission. One consequence is that the BLS does not report the amount of the quality correction – its implicit estimate of  $\beta$  -- that gets embodied in its price estimates that are used for output deflation by the BEA. More attention has been given to the size of the implied bias in the price deflators (and the bias in  $\beta$ ) than on the size of  $\beta$  itself. The subject has generated numerous studies, articles, and conference volumes (including some in the CRIW series (Hulten, 2015)). However, these studies produce mixed results about the size of the CPI bias. The estimates by Groshen et al. present a recent assessment of the overall bias based on past studies (including Lebow and Rudd (2003) and Greenstein and McDevitt (2011)). They put the downward bias in the annual growth of real GDP at -0.26% in 2015 due to consumer goods, and a -0.15% bias due to private investment (real GDP growth was around 2.0% in that year). The former is particularly relevant for this paper, since the “PC services (including Internet)” component of the -0.26% downward bias was only -0.04% (the contribution of medical bias was -0.12%). The “raw” annual bias in PC/Internet services was an annual -6.50% (based on Greenstein and McDevitt), but the GDP share of this category was so small that the share-weighted growth bias barely moves the GDP needle.

Other studies have also found larger biases than Groshen et al. That by Bils (2009) concludes “price inflation for durables has been overstated by nearly 2 percentage points per year”, and found that the BLS procedures for the CPI for autos and trucks understated quality improvements by 2.6 percentage points a year over that period. Indeed, when a large part of the value of a new car is due to electronics and software, new car features have very little additional resource cost, and thus are unlikely to appear as a price reduction. Thus, the gradual advent of a driverless car, with concomitant increase in leisure for the driver and reduction in accidents, is not likely to appear in measures of output.

Bils and Klenow (2001) use the Consumer Expenditure Survey to estimate “quality Engel curves” for 66 durable goods using the idea that richer households pay more for each good. They estimate that quality growth averages 3.7% percent per year for their sample of goods, with 2.2% showing up as pure price inflation, and conclude that BLS procedures do not fully account for the impact of quality upgrading. Based on their review of the available evidence, Shapiro and Wilcox (1996) place the midpoint (median) of their subjective probability distribution for the overall bias in the CPI at just under 1.0 percentage point per year with an eighty percent confidence interval stretching from 0.6 percentage point per year to 1.5 percentage points per year. Byrne, Fernald, and Reinsdorf (2016) provide estimates of the annual biases in investment price deflators, which range from 0.9% for software to 12% for computers and peripherals (the Greenstein and McDevitt estimate is in the middle of this range).

The focus of these studies is on the overall accuracy of price indexes and their major components. Other studies have a more narrow focus. Byrne and Corrado (2017a) provide estimates of the implied wired telecommunications services deflator based upon measures of the improving quality (and rapid deflation) of telecommunications equipment developed in Byrne and Corrado (2015) and methods described in Byrne and Corrado (2017b). They do so for *nonresidential* wireless services rather than for personal consumption expenditures, and they find a rate of deflation seven percentage points below the official measures from 2004 to 2014. Four other studies of the value of broadband are evaluated by Syverson. These provide estimates of consumer surplus that Syverson extrapolates to 2015 that range from a low of \$17 billion to a high of \$132 billion, including the Nevo et al. (2016) study of Internet access. It might also be noted that the hedonic regression for Internet broadband services used in the BLS PPI program includes a regression coefficient on download speed that suggests the 40 percent increase in speed experienced historically and would translate to a 12 percent further annual decrease in price.<sup>13</sup> This, in turn, would result in a decrease in the growth rate of the total PCE deflator that, if applied to both Internet access and cellular phone service, would increase real output by \$32 billion annually.

The rapid uptake of digital goods poses a special problem for price indexes where there are no prior goods on which to base quality comparisons. As noted, the BLS uses an imputation

---

<sup>13</sup> See <https://www.bls.gov/ppi/broadbandhedonicmodel.htm>.

procedure in such cases, which assigns the new good to a class of items at a base price equal to the means of the class (or a subset of the class). This is a problematic procedure to apply to the goods that are virtually the hallmark of the digital revolution. The magnitude of the annual upward new goods bias for the advent of the mobile cellular telephone was found by Hausman (1997) to be 1.9 % in the relevant component of the Consumer Price Index, telecommunications services. In a 1997 study of the introduction of a new brand of breakfast cereal, he found a 20 percent upward bias in the relevant subcategory -- a useful reminder that the new goods problem is not restricted to digital goods.

## VII. INFORMATION, THE INTERNET, AND THE CONSUMPTION TECHNOLOGY

### A. The Nature and Value of Information

Measuring the amount of the information that floods our senses every day is problematic and, in any event, it is not the volume of information in bits or bytes that matters for economic measurement. What matters is the perceived value to the recipient, and this depends on the way the information is organized, its relevance (often situational), its credibility or perceived accuracy, and its timeliness. Too much unstructured or irrelevant information can have a negative effect -- the noise-to-signal problem. The valuation of information is thus difficult, and it is compounded by the fact that most information flows without data-specific prices.

The information revolution has increased both “signal” and “noise”. For the purpose of this paper, we confine our attention to the disembodied output-saving innovations in the information that provide value to consumers, where value is determined by the amount they would be willing to pay if necessary, but which is in fact provided free of direct charge. We have formulated this as the parameter  $\mu$ . The magnitude of this parameter, as measured by the willingness-to-pay metric  $V$  of Section III, is of great consequence for the question of whether the growth rate of conventional real GDP provides a satisfactory measure of the dynamic changes in the economy over the course of the digital revolution. Addressing this question is the



overarching goal of this paper and, to this end, the rest of this section will marshal the available evidence on the size of  $V$  and  $\mu$ .<sup>14</sup>

### B. Current Treatment of Information in the Statistical System

BEA data from the U.S. national accounts by industry show that the GDP originating in the category “Information-communications-technology-producing industries” amounted to \$1.1 trillion in 2016, or about 6% of GDP. The scope of this category is rather broad, including the manufacturing of computer and electronic equipment which, when removed, causes this fraction to fall to 4.5%. A still narrower grouping with a focus on information services includes only “Data processing, Internet publishing, and other information services” (1.5%) and “Computer systems design and related services” (0.6%). Together, these two industries account for \$400 billion.

When the focus shifts to the consumer expenditures component of GDP (PCE), BEA data for the categories “Telecommunication services” and “Internet access” shows that consumers spent \$230 billion on the categories “Telecommunication services” and “Internet access” in 2016, or 1.8% of PCE and 1.2% of GDP. When expenditures for “Information processing equipment” and “Telephone and related communications equipment” are added to the list, the total increases to around \$380 billion, or 3.0% of PCE and 2.0% of GDP. By way of comparison, Groshen et al. report a GDP share for their category “PC services (including Internet)” of 0.6% for 2015 (the ratios we report are virtually the same for 2015 as in 2016). The larger point is that, in any case, the GDP associated with the digital economy are small using national accounting data.

If this were the final word on the subject, then the aggregate consequence of the digital revolution may be smaller than many of its enthusiasts claim. However, this is far from the last word. Many of the information goods consumed are transferred without a direct charge, and there is thus no monetary value to include in GDP. The cost to providers of producing the good

---

<sup>14</sup> Any attempt to assess the role of information in promoting consumer utility should recognize its public good nature. It is both non-rival (one person’s use of the Internet does not crowd out anyone else’s use), and it is difficult and cumbersome to create markets that price individual “units” consumed. Determining the optimal amount of a public good and determining its value are classic problems in public finance. Many information goods can be classified as partial public (or “club”) goods for which access fees are charged (e.g., the use of the gasoline tax to finance road systems). Some are pure public goods, as with information broadcasted over networks.

is often defrayed using indirect or ancillary revenues. Google and Facebook illustrate this problem. They are firms that have as their primary functions serving consumers with search and social networking, respectively, and, each firm's economic model is to provide its primary function at no direct cost to the consumer, supporting this economic activity with advertising. The two companies, together, reported annual revenues in 2016 of over \$115 billion, largely from advertising, and had a total market value of roughly \$1 trillion as of mid 2017. This business model implies that the flow of payments does not relate to the price or quantity of the information goods provided to consumers. The monetary flows involved appear in GDP via the price and quantities of the goods that are advertised.

Some part of the total value of information is covered by system access fees charged for network use. These payments tend to be blanket fees that are unrelated, or only loosely related, to the quantity or value of the information or social interaction on which value is based. Moreover, it is also true that some of the information offered at a zero marginal cost over the Internet or other media is simply free, provided *pro bono publico* by Internet application developers (von Hippel's "free innovation", 2016), or crowd-sourced and without a measured resource cost.

The value of the information services actually recorded in GDP is in the range of \$100 billion to \$400 billion, depending on how broad a definition is used.<sup>15</sup> The question is how much this range understates the true value to consumers, as revealed by the price they would be willing to pay for the "free" information goods. This is the question to which we now turn.

### C. The Measurement Literature on the Internet's Contribution to Welfare

There are a small, but growing, number of studies that address the measurement issues implied by Schmidt and Rosenberg's remark that "the Internet has made information free, copious, and ubiquitous". They cover both the Internet and the explosion in timely information it enables, but

---

<sup>15</sup> It should be emphasized that the Internet is scarcely the only channel through which information reaches the population. Education is an even more important channel, whether learning takes place in schools or at home or among peers. Books and other media are important, as is life experience. Much of this escapes GDP, and a full account would be a challenging task. Our goal in this paper is limited to an analysis of how costless increases in digital sources of information can provide consumer benefits beyond those records in GDP, and thereby present a different assessment of economic progress.

also the devices needed to enable the digital revolution. The former are associated with disembodied output-saving technical change,  $\mu$ , and will be the focus of the studies reviewed below.

There are several ways to measure the value of the Internet's information and entertainment flows, one of which is to use econometric techniques to estimate the expenditure function or the compensating and equivalent variations associated with the utility function ( $V$ ), or the system of demand equations associated with these functions. This can, in principle, get at the non-GDP contribution to consumer welfare in a framework that also includes the GDP contribution, to the extent that goods are priced. This is the approach followed by Redding and Weinstein (2016).<sup>16</sup>

Another line of attack on the problem is to introduce time cost into the analysis of value. A search engine can be seen as creating consumer value by reducing the time cost involved in acquiring information, and Varian (2009) adopts this approach using a finding from Chen et al. (2013), who had students at the University of Michigan obtain answers to questions using either a search engine or the library of the University of Michigan. The students who used the search engine were more successful, getting answers to questions posed in an average of seven minutes compared to 22 minutes using the library. Varian calculated the implied value to individual consumer value of roughly \$500 per year. Goolsbee and Klenow (2006) use a value-of-time approach, but focus on the Internet as a whole using a parametric consumption function analysis. They estimate that the value of the time spent on the Internet translates into a consumer surplus of \$2500 to \$3800 per year. Syverson (2016) also conducts an exercise in which he updates the Goolsbee and Klenow estimate of the value of the Internet and obtains a measure of the aggregate increase in the value of broadband of \$842 billion post-2004 time period. Other creative approaches to the consumer surplus problem use questionnaires, surveys, and microdata provide another approach to the consumer surplus problem. The literature includes Brynjolfsson, Hu, and Smith (2003), Aguiar and Waldfogel (forthcoming), Quan and Williams (2015), and Einav et al. (2017).

Another way to deal with zero prices is with direct measures of willingness-to-pay. An unusual opportunity to estimate willingness to pay with a free good is discussed in Noll et al.

---

<sup>16</sup> The Redding-Weinstein methodology assumes that time-varying demand shifts cancel on average. This assumption may not be valid when net gains in consumer technology, such as those generated by the Internet, occur.

(1973). The slow diffusion of broadcast TV meant that some rural households had to pay for the broadcasts that were free elsewhere. Using demand analysis, they were able to estimate that households would be willing to pay some three percent of income for free TV. However, such natural experiments are rare in the literature. One alternative is simply to ask people about their willingness to pay for a search engine. Varian (2009) used Google consumer surveys to ask this question and found that, on average, consumers were willing to pay \$36 a year for search, a much smaller number than his back-of-the-envelope welfare calculation. However, more recent work by Brynjolfsson et al. (2017) suggests that the minimum payments consumers would accept (WTA) for loss of access to search engines may be as large as \$5000 a year. This estimate suggests a value of about \$1 trillion missing from GDP from search alone.

A small industry has arisen in evaluating Facebook willingness-to-accept. Brynjolfsson et al. (2019) report a willingness to accept Facebook of about \$506 per user, with 202 million users, or \$100 billion in aggregate. They also estimate that this amount adds between 0.05 to 0.11 percentage points to the growth of real GDP (in other words, the increment to  $\mu$  is between 5 and 11 basis points). In another study, an auction experiment conducted by Corrigan et al. (2018) put the value of doing without Facebook for an entire year at between \$1000 to \$2000 per adult person in the U.S., with an implied value of as much as \$250 to \$500 billion a year. The largest scale experiment, Alcott et al. (2019), finds a similar value.

Finally, Nakamura et al. (2018) argue that even if we measure the cost of “free” information and entertainment in terms of their cost of production, the gains from marketing-supported information and entertainment are substantial. Taken from the cost-side alone, total nominal value in 2015 was \$103 billion from Internet contributions to personal consumption expenditures. This cost estimate does not include the volunteer time invested by consumers in creating Internet content, nor does it attempt to estimate any consumer surplus, just business-paid input costs in producing Internet content. The authors argue that including their conservative methodology would lower the PCE deflator by roughly 0.1 percent.

In sum, the results of different approaches vary from as little as \$100 billion to considerably more than \$1 trillion. This range of values suggests that there is ample potential for welfare gains to the consumer beyond those that are not included in the value of personal consumption expenditures and GDP.

## VIII. HEALTH AND EDUCATION: THE ROLE OF STATE CONTINGENCY

Consumer welfare gains are often contingent on a person's current state of being, or on changes in that state. The benefit of a health care intervention or expenditure, for example, depends on the state of health, and it is often shocks to that state that trigger the demand for the intervention. Moreover, the success of the intervention is often contingent on the severity of the shock (the same is true of some legal and financial problems). Other interventions are intended to improve the ambient state of being. The benefits of obtaining an education, for example, involve a move from one level of knowledge to another. Similarly, some health interventions are intended to improve the ambient state of health, through healthier life styles and preventative medicine. Moreover, education and health interventions may interact in ways that strengthen each other.

Improvements in health and education affect consumer welfare both through increases in income and in the satisfaction derived from that income, i.e., through shifts in both the production and consumption technologies. Healthier and better educated workers tend to be more productive and thus command higher wages, and they also are able to make better choices and live more comfortable lives, freer of pains and cares. On average, there have been dramatic increases in both dimensions. Average life expectancy increased from 48 to 78 years over the course of the last century, with an attendant increase in quality-of-life years. On the education side, Census estimates show that only one-third of the adult population had a high-school degree in 1950 and only six percent had a four-year college degree or more. By 2013, almost 90% had at least a high-school degree and one third had an undergraduate college degree or higher.

The consumption technology as formulated in this paper refers to the average state-of-health or knowledge, whereas much of the actual contingency involves states experienced by individuals. A health care innovation, like minimally invasive surgery, will generally affect a subset of the population, and perhaps only a small subset. However, the gains to those affected may be quite large but appear small when averaged into the total population. Moreover, some innovations may allow a subset of those afflicted that were previously untreatable to be helped. The innovation may improve the welfare of that subset, but if the success rate of the treatment is lower for this group than for the population as a whole, and if success rates are used as an indicator of innovation, the metric may send a false signal.

An extension of EGD<sub>P</sub> model to allow for individual heterogeneity in contingent states is not easy, since it involves the utility of individuals and a way of aggregating their utilities. The standard way is to appeal to an explicit social welfare function (as opposed to the one implied by the use of averages). This step involves the introduction of value judgments into the measurement of GDP and EGD<sub>P</sub>. This is a major step, and since the basic thrust of this paper is to explore the EGD<sub>P</sub> concept *per se*, it is a step we will defer to subsequent research.

## B. INNOVATION IN HEALTH CARE

The 2017 review of the bias in price statistics by Groshen et al. identified health care as a major source of the accuracy problem. The health care has been a hard-to-measure industry for a long time because of the problems associated with the disconnect between expenditures and outcome that form the basis for the “output movement” described by Cutler and Berndt (2001). It has been the beneficiary of rapid innovation, much of which has improved outcomes for given levels of expenditure, which constitute our output-saving technical change. The case of minimally-invasive surgery has been noted already, but there are many other examples.

Recent studies have found large potential biases in health care. For example, Dauda, Dunn, and Hall (2018) find that annual medical price inflation declined by 4.8 percent relative to aggregate inflation rates over the period 2001 to 2014. With health care expenditures accounting for 17 to 20 percent of personal consumption during this period, this would add close to 1 percent to the growth rate of the total. They also report that, for heart attack, congestive heart failure, and pneumonia, 30-day-risk-adjusted mortality rates fell significantly over this 13 year period (-39 percent, -25 percent, and -40 percent, respectively), while 30-day-risk-adjusted expenditure rose much less rapidly (-1 percent, +20 percent, and +11 percent, respectively). In other words, outcomes have improved over the period with much less increase in spending, the very phenomenon our framework seeks to address.

Output-saving innovation is also present in the studies by Chernew et al. (2016), who report that disability-adjusted life years increased 1.8 years at age 65 between 1992 and 2008, of which they attribute 1.1 years to improved health treatment, particularly of heart disease and vision problems. Along the same lines, the Murphy-Topel (2006) calculation of the value of the increases in life expectancy finds a very large number, \$1.2 million per person, for the representative person in 2000 in the U.S. This results in a very large number, but it should be

noted that valuing human capital is a perilous enterprise, as is assigning changes in the value to factors other than medical treatment (Fogel, 2012). However, taken together, these health care studies highlight the importance of outcomes (longevity, mortality rates), as opposed to expenditures.

Another example of output-saving technical change comes from the recent study by Rothwell et al. (2016), who found taking aspirin for 12 weeks following a stroke or mini-stroke lowers the probability of a recurrent stroke or heart attack during that period from 4.3 percent to 1.9 percent. The cost of avoiding one stroke or heart attack is thus \$40, assuming an aspirin cost of \$.01 per tablet, orders of magnitude smaller than the consumer benefit, however measured.

### C. THE CASE OF EDUCATION

There have been major gains in educational attainment in the U.S., but also large expenditures and poor test results (see summary in Hulten and Ramey, 2019). Education premia have led to rising incomes for much of the population, and increased productivity has propelled output growth. The average quality of life has doubtless risen as well, but how much more tuition college students would be willing to pay over and above the amount they already pay for this enhanced quality of life is unclear.<sup>17</sup> In this section, we explore another aspect of education's impact of individual welfare: the importance of initial states and individual heterogeneity in assessing the welfare benefits of education.

Formal education is an output of the schooling industry, but student learning is the relevant outcome. Schooling is an important channel through which learning occurs, but family, peers, and personal experience all make important contributions to these outcomes. Student "inputs" of effort are also important and depend on idiosyncratic characteristics like motivation and general openness to change. As Hulten and Ramey (2019) observe, "[poor] K-12 results cannot be attributed to the quality of schooling alone. Many other non-school inputs also affect student outcomes. Moreover, research suggests that the cognitive and noncognitive skills developed by age three have fundamental effects on the ability to learn. Thus, K-12 schools have little control over key inputs into their production functions."

---

<sup>17</sup> Education plays an important role in the quality of life. It exposes people to ideas and possibilities that expand consumer horizons and enhance the enjoyment of life. Put in economic terms, it allows people to get more enjoyment out of each dollar they spend, as with the shift in the consumption technology.

Improvements in the outcomes of historically underserved student populations have a large payoff to society and, importantly, to those individuals who stand to benefit. When looked at from the standpoint of the individual, each person has the equivalent of the basic macro equation (1), which was formulated to describe the welfare of the representative individual. The welfare gain of actual individuals is also a matter of increased personal income and non-monetary improvements in the quality of their life. Improvements in the quality of their education increase both, and both must be included in the assessment of outcomes. Subsuming these gains in a measure based on average experience risks missing some of the most important gains in outcomes.

#### IX. Final Thoughts on the Path Ahead for EGD<sup>2</sup> Measurement

We have proposed Expanded GDP as a new measure of aggregate economic activity that builds on existing GDP by adding to it the consumer's willingness-to pay for output-saving innovation. Our review of the empirical literature and the available data suggest that this effect is non-negligible, perhaps amounting to as much as a trillion dollars or more. While it is true that the GDP share of the digital economy is relatively small, as some have noted, we have shown in our earlier paper that the effect on EGD<sup>2</sup> growth can be quite large despite this small share. In a previous study, we conducted a thought experiment in which the bias in price-deflators noted by Groshen et al. that when combined with the impact of output-saving technical change could easily be a full percentage point (100 basis points) higher. Given that the average annual top line growth of the Private Business sector shown in Table 1 of this paper is 2.76% for the period of 1995 to 2007, a 100 basis point increase is significant.

We emphasize that this hypothetical estimate is *not* intended as our best guess at the contribution of output-saving innovation to expanded economic growth, but it *is* intended to show that the output-saving effect, and the resulting EGD<sup>2</sup>, are potentially too large to be ignored. But, if these potential consumer welfare effects are too large to be ignored, should they not be recognized in GDP estimates? GDP is like an album that contains annual snapshots of the economy as it evolves over time. It provides a way of measuring the growth in the economy and recording the ways it has changed. To be useful, the content of the annual pictures must reflect the economy as it is at each point in time, and as it evolves, the content of the pictures must



reflect the way it has evolved. One aspect of the current evolution is that much of the value creation discussed in this paper occurs outside of formal markets, and must be added to the picture.

This said, part of the value of GDP lies in the continuity of the time record it provides. The question is not simply “what do we look like today,” to use the album analogy, but “what do we look today in comparison with how we looked in the past”. There is thus a tension between updating the accounts to reflect the current economy and maintaining comparability over time. One way to deal with this quandary is through the use of satellite accounts to bridge the gap. A satellite account preserves the main accounting structure of GDP, while at the same time providing a home for the more speculative estimates emerging from the study of the current technical revolution.

The EGD<sup>1</sup> model set out in this paper, and in our earlier work, provides a model of a satellite account that includes EGD<sup>1</sup>. It follows the Koopmans’ injunction that measurement practice should be anchored in the economic theory that it empirically informs. This goal is already set out in the CRIW volume *A New Architecture of the U.S. National Accounts*, and adding an EGD<sup>1</sup> satellite as shown in Figure 3 of this chapter would be a further step along this path. It is not a minor step, but one that would require new sources, surveys, and techniques to go beyond the one-off studies in the literature reviewed in this chapter. Fortunately, the very technical revolution that has increased the importance of the non-market economy has also created new techniques and methods for processing the huge amount of information involved in understanding customer preferences and the value created by (and for) the firms concerned. This new expertise could potentially be applied to the estimation of EGD<sup>1</sup>, and perhaps form the basis for an experimental innovation satellite account to the main GDP program.

Fortunately, a start in this direction has already been made at BEA with its innovation accounting and limited capitalization of intangible assets. This innovation account could be expanded in several important ways. One is to extend the current list of intangible capital included in GDP to encompass a broader range of intellectual property, enterprise-specific human capital, and organizational assets. Another important step is for BLS and BEA to work together to improve price statistics so that they more accurately reflect product-innovation. Taking on the challenge posed by new goods, like the Internet and mobile communication

devices, is of central importance in this regard. Another major step within the scope of existing statistical programs is for the BLS to report separately the extent of product innovation already embodied in its quality-corrected prices estimates. Finally, the research from the “outcome movement” in health care research should be accorded a high priority.<sup>18</sup>

The task of building a full innovation satellite account is daunting. The history of the national accounts is a history of overcoming one daunting challenge after another. The result of these efforts has been what Samuelson and Nordhaus have called “One of the Great Inventions of the 20th Century”.<sup>19</sup>

## REFERENCES

Abramovitz Moses, 1956, “Resource and Output Trends in the United States since 1870”, *American Economic Review*, 46(2), 5–23

Aguiar, Luis, and Joel Waldfogel, forthcoming, “Quality Predictability and the Welfare Benefits from New Products: Evidence from the Digitization of Recorded Music,” *Journal of Political Economy*.

Allcott, Hunt, Luca Braghieri, Sarah Eichmeyer, and Matthew Gentzkow, 2019, “The Welfare Effects of Social Media,” January, working paper.

Anderson, Monica, 2015. “Technology Device Ownership: 2015”,  
<http://www.pewInternet.org/2015/10/29/technology-device-ownership-2015>.

Becker Gordon M, Morris H. DeGroot and Jacob Marschak, 1964, “Measuring Utility by a Single-response Sequential Method,” *Behavioral Science* 9: 226–232.

Bils, Mark, 2009, “Do Higher Prices for New Goods Reflect Quality Growth or Inflation?” *Quarterly Journal of Economics*, May, 637-675.

Bils, Mark, and Peter J. Klenow. 2001. "Quantifying Quality Growth." *American Economic Review*, 91(4): 1006-1030.

Brooks, Arthur, 2008. “How to Succeed in Life ... by Really Trying: Money Doesn’t Buy Happiness, but Success Does”, *The American*, May/June, 27-34, 110-111.

---

<sup>18</sup> It must also be said that BLS is continually working to improve the CPI and the PPI. For example, BLS is moving to what has been called a diagnosis or a disease-centric approach (Roehrig, 2017). Progress is also underway at BEA on the in other problem of measuring outcomes in the provision of health care services. Much progress has been made, but the path ahead is long and difficult.

<sup>19</sup> Cited by J. Steven Landefeld (2000).

Brynjolfsson, Erik, Avinash Collis, W. Erwin Diewert, Felix Eggers, and Kevin J. Fox, Felix Eggers, “The Digital Economy, GDP and Consumer Welfare: Theory and Evidence,” October 2018.

Brynjolfsson, Erik, Felix Eggers, and Avinash Gannamaneni, forthcoming, “Using Massive Online Choice Experiments to Measure Changes in Well-being,” *Proceedings of National Academy of Science*.

Brynjolfsson, Erik, Yu (Jerry) Hu, and Michael D. Smith, 2003. “Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers,” *Management Science* 49, 1580-1596.

Byrne, David M. and Carol A. Corrado, 2015, “Prices for Communications Equipment: Rewriting the Record,” *Finance and Economics Discussion Series 2015-069*. Washington: Board of Governors of the Federal Reserve System,

Byrne, David, and Carol Corrado, 2017. “ICT Prices and ICT Services: What do they tell us about Productivity and Technology?” *Finance and Economics Discussion Series 2017-015*. Washington: Board of Governors of the Federal Reserve System,  
<https://doi.org/10.17016/FEDS.2017.015>.

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*, March.

Chen, Yan, Grace Young Joo Jeon and Yong-Mi Kim, 2014, “A Day without a Search Engine: An Experimental Study of Online and Offline Search.” *Experimental Economics* 17 (4): December, 512-536.

Chernew, Michael, David M. Cutler, Kaushik Ghosh, Mary Landrum, 2016, “Understanding the Improvement in Disability Free Life Expectancy in the U.S. Elderly Population,” *NBER Working Paper 22306*, June.

Corrado, Carol A., and Charles R. Hulten, 2010. “How Do You Measure a ‘Technological Revolution’?” *American Economic Review*, 99-104.

Corrado, Carol A., and Charles R. Hulten, 2014. “Innovation Accounting”, in D.W. Jorgenson, J.S. Landefeld, and P. Schreyer, eds., *Measuring Economic Progress and Economic Sustainability*, Studies in Income and Wealth, vol. 72, The University of Chicago Press for the National Bureau of Economic Research, Chicago, 2014, 595-628.

Corrado, Carol A., and Charles R. Hulten, 2015, “Connecting the Supply and Demand Sides of the Economy: Productivity, Asset Valuation, and Tobin’s ‘q’,” in C.R. Hulten and M.B. Reinsdorf, eds., *Measuring Wealth and Financial Intermediation and Their Links to the Real Economy*, Studies in Income and Wealth, vol. 73, Chicago:University of Chicago Press,125-147.

Corrado, Carol A., Charles R. Hulten, and Daniel E. Sichel, 2005, “Measuring Capital and Technology: An Expanded Framework” in Corrado, Carol, John Haltiwanger and Daniel Sichel, eds., *Measuring Capital in the New Economy*, Chicago: University of Chicago Press, 11-46.

Corrado, Carol A., Charles R. Hulten, and Daniel E. Sichel, 2009, “Intangible Capital and U.S. Economic Growth”, *Review of Income and Wealth*, 55, 3, September, 661-685.

Corrigan, Jay R., Saleem Alhabash, Matthew Rousu, and Sean B. Cash, “How much is social media worth? Estimating the value of Facebook by paying users to stop using it,” *PLoS ONE* 13(12). December 19, 2018, <https://doi.org/10.1371/journal.pone.0207101>.

Coyle, Diane, *GDP: A Brief But Affectionate History*, Princeton: Princeton University Press, 2014.

Cutler, David M. and Ernst R. Berndt, 2001, “Introduction,” in David M. Cutler and Ernst R. Berndt, eds., *Medical Care Output and Productivity*, NBER, University of Chicago Press, 1-12.

Dauda, Seida, Abe Dunn and Anne Hall, 2018, “Are Medical Care Prices Still Declining? A Systematic Examination of Quality-Adjusted Price Index Alternatives for Medical Care,” working paper presented at NBER/CRIW Summer Institute.

Einav, Liran, Peter J. Klenow, Benjamin Klopach, Jonathan D. Levin, Larry Levin and Wayne Best, 2017. “Assessing the Gains from E-Commerce,” working paper, <http://web.stanford.edu/~leinav/wp/ecommerce.pdf>

Fogel, Robert W., 2012, *Explaining Long-Term Trends in Health and Longevity*, Cambridge: Cambridge University Press.

Goolsbee, Austin and Peter J. Klenow, 2006, “Valuing Consumer Products by the Time Spent Using Them: An Application to the Internet,” *American Economic Review Papers and Proceedings*, 96(2): 108-113.

Gordon, Robert J., 2016. *The Rise and Fall of American Growth: The U.S. Standard of Living since the Civil War*, The Princeton Economic History of the Western World, Princeton, New Jersey: Princeton University Press, January.

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, 1992, “Introduction,” in Griliches, Zvi, ed., *Output Measurement in the Services Sectors*, Chicago: University of Chicago Press, 1-22..

Griliches, Zvi, 1994. “Productivity, R&D, and the Data Constraint.” *American Economic Review* 84, March, 1-23.

- Groshen, Erica L., Brian C. Moyer, Ana M. Aizcorbe, Ralph Bradley, and David Friedman, 2017, "How Government Statistics Adjust for Potential Biases from Quality Change and New Goods in an Age of Digital Technologies: A View from the Trenches", *Journal of Economic Perspectives*, 31, 2, Spring, 187–210
- Hausman, Jerry, 1997. "Valuation of New Goods Under Perfect and Imperfect Competition," ed. T. Bresnahan and R. Gordon, *The Economics of New Goods*, Chicago: University of Chicago Press, 209-237.
- Hausman, Jerry, 1999. "Cellular Telephone, New Products and the CPI," *Journal of Business and Economics Statistics*, vol 17(2), 188-194.
- Hill, Peter, 1999, "Tangibles, Intangibles and Services: A New Taxonomy for the Classification of Output," *Canadian Journal of Economics* 32, April, 426-446.
- Hulten, Charles R., 1978. "Growth Accounting with Intermediate Inputs," *Review of Economic Studies*, 45, 3, October, 511-518.
- Hulten, Charles R., 2001. "Total Factor Productivity: A Short Biography," in Charles R. Hulten, Edwin R. Dean, and Michael J. Harper, eds., *New Developments in Productivity Analysis*, Studies in Income and Wealth, vol. 63, The University of Chicago Press for the NBER, Chicago, 1-47.
- Hulten, Charles, 2015, "Measuring the Economy of the 21st Century", *NBER Reporter*, No 4, December, 1-7.
- Hulten, Charles, and Leonard Nakamura, 2017. "Accounting for Growth in the Age of the Internet: The Importance of Output-Saving Technical Change," NBER Working Paper 23315.
- Hulten, Charles R., and Valery A. Ramey, "Education, Skills, and Technical Change, Implications for Future U.S. GDP Growth: An Introduction and Overview", in Hulten and Ramey, eds., *Education, Skills, and Technical Change: Implications for Future U.S. GDP Growth*, Studies in Income and Wealth, The University of Chicago Press for the NBER, Chicago, 2019, 1-19.
- Koopmans, Tjalling C., "Measurement Without Theory," *The Review of Economics and Statistics*, 29, 3, August 1947, 161-172.
- Lancaster, Kelvin J., 1966a. "Change and Innovation in the Technology of Consumption", *American Economic Review*, 56, 1/2, March 14-23.
- Lancaster, Kelvin J., 1966b. "A New Approach to Consumer Theory", *Journal of Political Economy*, 74, 2, April, 132-157.
- Landefeld, J. Steven, "GDP: One of the Great Inventions of the 20th Century," *Survey of Current Business*, January 2000.

Maddison, A. (2007), *Contours of the World Economy, 1–2030 AD. Essays in Macro-Economic History*, Oxford University Press, Oxford.

Moulton, Brent R. and Karin E. Moses, 1997, “Addressing the Quality Change Issue in the Consumer Price Index,” *Brooking Papers in Economic Activity*, 1, 305-366.

Murphy, Kevin and Robert H. Topel, 2006. “The Value of Health and Longevity,” *Journal of Political Economy* 114, October, 871-904

Nakamura, Leonard, Jon Samuels and Rachel Soloveichik, 2018, “ ‘Free’ Internet Content: Web 1.0, Web 2.0, and the Sources of Economic Growth,” *Federal Reserve Bank of Philadelphia Working Paper 2018-17*, May.

Nevo, Aviv, Jon L. Turner, Jonathan W. Williams, 2016, “Usage Based-Pricing and Demand for Residential Broadband,” *Econometrica* 84 (2), March, 411-443.

Noll, Roger G., Morton J. Peck, and John J. McGowan, 1973, *Economic Aspects of Television Regulation*, The Brookings Institution: Washington DC.

Nordhaus, William D., 1996, “Do Real-Output and Real-Wage Measures Capture Reality? The History of Lighting Suggests Not.” In Timothy Bresnahan and Robert J. Gordon, eds., *The Economics of New Goods*, University of Chicago Press: Chicago, 27-70.

Nordhaus, William D., “Schumpeterian Profits and the Alchemist Fallacy Revised,” Yale Working Papers on Economic Applications and Policy, Discussion Paper No. 6, April 2, 2005.

Patinkin, Don (1973). “In search of the ‘Wheel of Wealth’: On the Origins of Frank Knight’s Circular Flow Diagram,” *American Economic Review*, 63: 5 (December), 1037-1046.

Perrin, Andrew, 2015. “Social Networking Usage: 2005-2015”, Pew Research Center. October 2015, <http://www.pewInternet.org/2015/10/08/2015/Social-Networking-Usage-2005-2015>

Pew Research Center, 2017. “Mobile Factsheet” <http://www.pewInternet.org/fact-sheet/mobile>

Quan, Thomas W. and Kevin Williams, 2016. “Product Variety, Across Market Demand Heterogeneity, and the Value of Online Retail,” Cowles Foundation Discussion Paper 2054, November.

Redding, Stephen J., and David E. Weinstein, 2016, “A Unified Approach to Estimating Demand and Welfare,” Working Paper, June.

Rothwell, Peter M., Ale Algra, Zhengming Chen, Hans-Christoph Diener, Bo Norrving, Ziyah Mehta, 2016, “Effects of Aspirin on Risk and Severity of Early Recurrent Stroke after Transient

Ischaemic Attack and Ischaemic Stroke: Time-Course Analysis of Randomised Trials,” *Lancet* 388, May, 365-75. [http://dx.doi.org/10.1016/S0140-6736\(16\)30468-8](http://dx.doi.org/10.1016/S0140-6736(16)30468-8)

Schmidt, Eric, and Jonathan Rosenberg (with Alan Eagle), 2014. *How Google Works*, Grand Central Publishing, September.

Shapiro, Matthew D., and David W. Wilcox, “Mismeasurement in the Consumer Price Index: An Evaluation.” *NBER Macroeconomics Annual 1996*, Volume 11, Ben S. Bernanke and Julio J. Rotemberg, Editors, MIT Press, 1996, 93 -154).

Sichel, Daniel E., " The Productivity Slowdown: Is a Growing Unmeasurable Sector the Culprit?," *The Review of Economics and Statistics*, Volume 79 , No. 3, 1997, 367-370.

Sichel, Daniel E., and Eric von Hippel (2019), *Household Innovation, R&D, and New Measures of Intangible Capital*.” MIT Sloan School of Management Working Paper.

Smith, Aaron, and Monica Anderson, Social Media Use in 2018, Pew research Center, March 1, 2018 <http://www.pewinternet.org/2018/03/01/social-media-use-in-2018/>

Smith, Adam, *The Wealth of Nations*, volume 1, 1776, Irwin Paperback Classics in Economics, Richard D. Irwin, Homewood Illinois, 1963.

Solow, Robert M., 1956. “A Contribution to the Theory of Economic Growth,” *Quarterly Journal of Economics*, 70, 65-94.

Solow, Robert M., “Technical change and the aggregate production function,” *Review of Economics and Statistics*, 39, August 1957, 312–320.

Stigler, George, ed. *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, 1961. Stigler commission: Price Statistics Review Committee,

Stigler, George, and Gary S. Becker. 1977. “De Gustibus Non Est Disputandum,” *American Economic Review*, 67 March, pp. 76-90.

Syverson, Chad, 2017, “Challenges to Mismeasurement Explanations for the US Productivity Slowdown,” *Journal of Economic Perspectives*, Spring, 165-86.

Triplett, Jack E. and Barry P. Bosworth, 2004, *Productivity in the U.S. Services Sector: New Sources of Economic Growth*, Brookings Institution Press: Washington DC.

U.S. Bureau of Labor Statistics, *Handbook of Methods*, Consumer Price Indexes, Chapter 17 <https://www.bls.gov/opub/hom/pdf/homch17.pdf>

U.S. Bureau of Labor Statistics, *Handbook of Methods*, Producer Price Indexes, Chapter 14 <https://www.bls.gov/opub/hom/pdf/homch14.pdf>

U.S. Census Bureau, 2014, “Computer and Internet Access in the United States: 2012”, February, 2014, <https://www.census.gov/data/tables/2012/demo/computer-Internet/computer-use-2012.html>

Varian, Hal, (2009), “Economic Value of Google,” BEA Advisory Council Meeting, December <http://cdn.oreillystatic.com/en/assets/1/event/57/The%20Economic%20Impact%20of%20Google%20Presentation.pdf>

Varian, Hal. 2016, “A Microeconomist Looks at Productivity,” BEA Advisory Council Meeting, December <https://www.brookings.edu/wp-content/uploads/2016/08/varian.pdf>

von Hippel, Eric, *Free Innovation*, MIT Press, 2016.



**TABLE 1**  
**Sources of Growth in U.S. Private Business Sector**  
**(average of annual growth rates)**

	1948- 2007	1948- 1973	1973- 1995	1995- 2007
1. Output per hour $[q^e - \ell]$	2.41	2.99	1.56	2.76
<i>percentage point contribution to output per hour of:</i>				
2. Tangible capital $[s_K(k-\ell)]$	0.65	0.76	0.52	0.67
Memo: ICT equipment	0.23	0.11	0.28	0.37
3. Intangible capital $[s_N(n-\ell)]$	0.42	0.30	0.39	0.74
Memo: R&D (NSF/BEA)	0.10	0.08	0.07	0.17
4. Labor composition	0.20	0.15	0.26	0.2
5. TFP $[\beta^* + \lambda]$	1.14	1.78	0.39	1.16
<i>percent of total contribution to output per hour of:</i>				
2. Tangible capital	27%	25%	33%	24%
Memo: ICT equipment	10%	4%	18%	13%
3. Intangible capital	17%	10%	25%	27%
Memo: R&D (NSF/BEA)	4%	3%	4%	5%
4. Labor composition	8%	5%	17%	7%
5. TFP	47%	60%	25%	42%

ICT refers to Information and Communications Technology Equipment, BEA to the Bureau of Economic Analysis, NSF to the National Science Foundation, TFP is Total Factor Productivity. The latter includes both  $\beta^*$  and  $\lambda$  terms, since the hyper-output concept,  $Q^e$ , is used in these data rather than resource-based output,  $Q^r$ . The procedures used to estimate product quality innovation are, at best, incomplete, hence the  $\beta^*$  rather a true  $\beta$ . Source: Corrado and Hulten (2010, 2015).