Volume Title: Measuring and Accounting for Innovation in the Twenty-First Century

Volume Authors/Editors: Carol Corrado, Jonathan Haskel, Javier Miranda, and Daniel Sichel, editors

Volume Publisher: University of Chicago Press

Volume ISBNs: 978-0-226-72817-9 (cloth); 978-0-226-72820-9 (electronic)

Volume URL: https://www.nber.org/books-and-chapters/measuring-and-accounting-innovation-twenty-first-century

Conference Date: March 10–11, 2017

Publication Date: April 2021

Chapter Title: Expanded GDP for Welfare Measurement in the 21st Century

Chapter Author(s): Charles Hulten, Leonard I. Nakamura

Chapter URL: https://www.nber.org/books-and-chapters/measuring-and-accounting-innovation-twenty-first-century/expanded-gdp-welfare-measurement-21st-century

Chapter pages in book: p. 19 – 59

1

Expanded GDP for Welfare Measurement in the 21st Century

Charles Hulten and Leonard I. Nakamura

1.1 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" (10–11). However, its reach goes beyond the internet per se to include major advances in health care and higher education and structural changes in finance, 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, the location of production, and the way goods and services are distributed. One result has been an increase in the well-being of consumers.

The question addressed in this chapter is whether the procedures currently used to measure GDP adequately capture this increase. There are

Charles Hulten is professor emeritus of economics at the University of Maryland and a research associate of the National Bureau of Economic Research.

Leonard I. Nakamura is an emeritus economist at the Federal Reserve Bank of Philadelphia.

The views expressed in this chapter 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 the Bureau of Labor Statistics and the Bureau of Economic Analysis for help and advice. Jeanna Kenney provided excellent research assistance. For acknowledgments, sources of research support, and disclosure of the authors' material financial relationships, if any, please see https://www.nber.org/books-and-chapters/measuring-and-accounting-innovation-21st-century/expanded-gdp-welfare-measurement-21st-century.

good reasons to think that they do not.[1] 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 direct 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 because output is often measured in terms of inputs rather than outcomes, and as Griliches (1992, 1994) has noted, it is not even clear what actually constitutes output. The digital revolution has 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 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 about not just 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 (1966), 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

1. For example, Coyle (2014) remarked that GDP was "a measure of the economy best suited to an earlier era."

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 the digital revolution enable consumers to make more efficient use of their incomes. We term this form of innovation *output saving*, since a given level of welfare can be achieved with fewer resources, but it could equally be called *utility augmenting*, since it allows consumers to get more "bang for their buck." 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" shifts in the production function. The latter describes an increase in the productivity of inputs, while 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 general equilibrium dollar metric for measuring the benefits from innovations that go directly to the consumer. We add this dollar metric to conventional GDP to obtain an expanded concept of GDP. Expanded GDP (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 EGDP.[2] The goal is to decompose the growth rate of EGDP into output saving, resource saving, resource using, and input accumulation. This is essentially the conventional growth accounting framework with output-saving innovation added and costless product quality 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 EGDP may have grown faster than real GDP over the last three decades. Our estimates suggest that it did.

1.2 The Theory of Aggregate Output

1.2.1 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

2. The technical derivations and assumption can be found in our previous paper (Hulten and Nakamura 2018) on which the current chapter builds.

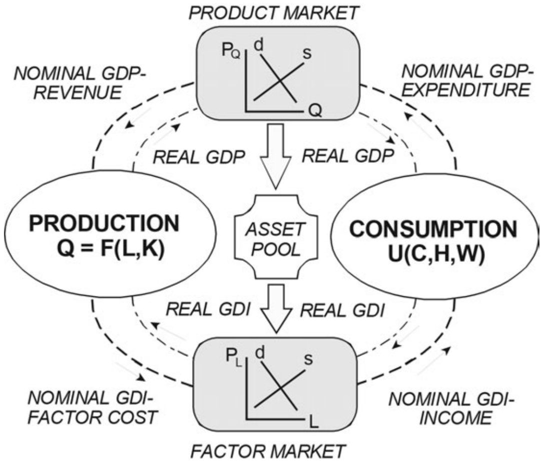

Fig. 1.1 Circular flow diagram for GDP and GDI

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.[3] The representation, shown in figure 1.1, 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 $F(\cdot)$; 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 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 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

3. See Patinkin (1973) for a discussion of the structure and history of the circular flow model.

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 United States 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 quantity, which is represented in figure 1.1 by the inner clockwise flow that tracks the movement of output and input quantities between producers and consumers. Prices are represented implicitly in figure 1.1 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 the innovation may appear very differently when it enters a sector of the economy than when it impacts final demand (e.g., Hulten 1978). 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 its conflation with goods 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.

1.2.2 Capital Formation

GDP and GDI are snapshots of the size of the aggregate economic flows in a time period. The bulk of US GDP goes to 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 1.1 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 1.1 labeled "asset pool."[4]

The pool of the productive capital contains different types of tangible capital (equipment and structures) as well as intangible capital. Intangible 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 improve the organization and management of firms on the other. Until quite recently, expenditures for intangible capital except computer software (only added in 1999) 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 the Bureau of Economic Analysis (BEA) of R&D and expenditures for artistic originals. This move added 3 percent to 4 percent to US 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).

1.3 GDP and Consumer Welfare

1.3.1 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 symmetrically 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 1.2, 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 the consumer through the product market, in which the volume and price of each good are determined by the interaction of supply and demand. Once the price and quantity of each good are determined, aggregate GDP follows immediately. Under the standard optimization assumptions, the resulting GDP represents the maxi-

4. This figure is based on figure 2 of Corrado and Hulten (2015).

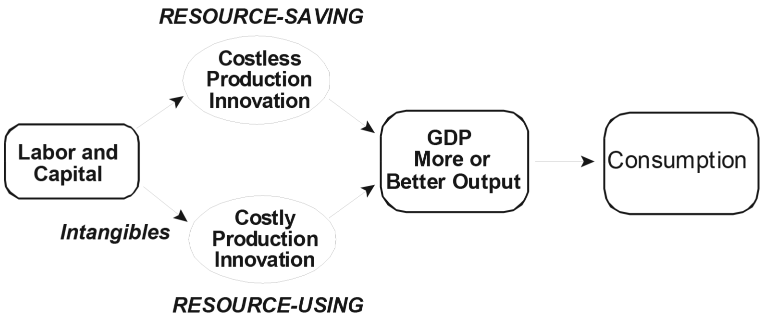

Fig. 1.2 Resource-saving and resource-using innovation

mum attainable utility. An increase in real output Q is assumed to increase utility, and a proportional increase in Q may 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).

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 total factor productivity (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 macroinnovation. 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 systematic investment in 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 innovation that increases the quality of existing goods or introduces new goods that is implicit in figure 1.2. 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 "better" is typically expressed as more output for purposes of measurement, to the extent that an adjustment is actually made.

1.3.2 GDP Expanded to Allow for Direct Consumption Benefits

Most thinking about GDP has focused on figures 1.1 and 1.2. Indeed, figure 1.2 illustrates the point at which the conventional measurement framework leaves off. However, an increase in the consumption efficiency and

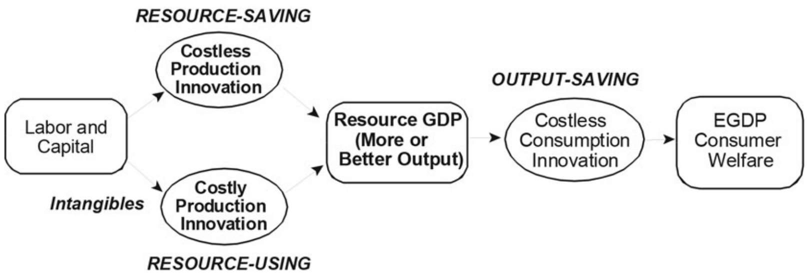

Fig. 1.3 Innovation including output-saving consumption technology

the increase in well-being it enables do 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 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.[5] 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. This form of innovation is "utility-augmenting," since it enables an increase in consumer welfare for the same amount of resources, or equivalently, it is "output-saving," since the prior level of welfare can be achieved with fewer resources. As a concrete example, 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 1.3 adds a consumption technology to the schema set out in figure 1.2. 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/utility-augmenting innovation operates as a link between resource output and commodity util-

5. The importance of the interaction between producer and consumer is also emphasized by Peter Hill (1999).

ity 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 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 of the product. In other words, 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).

The expansion of conventional output from figure 1.2 to figure 1.3 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 1.8, where we discuss state contingency in health and education.

1.3.3 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 1.3 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 two 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 those of GDP.[6] Figure 1.4 shows how this might work.

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

6. 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.

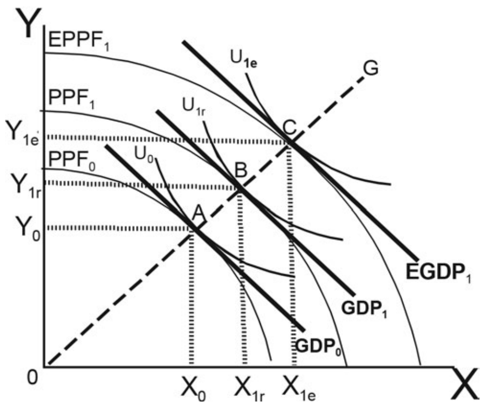

Fig. 1.4 GDP, EGDP, and the compensating variation

available in that year given the prevailing technologies for producing the goods (the first three stages of figure 1.3). 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_0^X and P_0^Y, and the line $P_0^X X_0 + P_0^Y Y_0$, 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 on the expansion path $0G$ at a higher indifference curve U_{1r} with an amount of real $GDP_{1r} = P_0^X X_{1r} + P_0^Y Y_{1r}$. The subscript r is used here to denote that the quantities of X and Y are measured in resource units. The dollar value of the real growth occurring between the two periods equals $GDP_{1r} - GDP_{0r}$, and the rate of growth is $(GDP_{1r} - GDP_{0r},)/GDP_{0r}$. The allocation of this rate among the growth in the inputs and technology can be estimated using the Solow (1957) residual method. $GDP_{1r} - GDP_{0r}$ 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 1.3. When the utility-augmenting 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 1.4 even though output in resource units (X_{1r}, Y_{1r}) remains unchanged, as do real GDP_1 and prices (P_0^X, P_0^Y). At these prices, the tangency between U_{1e} occurs at the point C. This tangency implicitly defines a new frontier labeled $EPPF_1$ to emphasize that it is the effective-output 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 subscript 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 results in the same proportion θ for both goods, as in figure 1.4, the result is $X_{1e} = (1 + \theta)X_{1r}$, and $Y_{1e} = (1 + \theta)Y_{1r}$. This is the phenomenon we have called utility-augmenting (or output-saving) technical change: an increase in utility for the same amount of resource-based output (occurring in this example at the rate θ).[7]

A little algebra establishes that the shift in utility from U_{1r} at B to U_{1e} at C is related to θ in the following way: $U_{1e} = (1 + \theta)U_{1r}$, under the simplifying assumptions of figure 1.4 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 1.4. 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 GDP*. $EGDP_1 = P_0^X X_{1e} + P_0^Y Y_{1e}$. It then follows that $EGDP_1 = (1 + \theta)(P_0^X X_{1r} + P_0^Y Y_{1r}) = (1 + \theta)GDP_1$. 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 $EGDP_1$ and GDP_1. In other words, $V = EGDP_1 - GDP_1 = (1 + \theta)GDP_1 - GDP_1$, from which it follows that $V = \theta GDP_1$ and that $\theta = V/GDP_1$. This result is significant for the issues at hand because it shows that the unobservable rate of output-saving technical change, θ, is potentially observable through the use of consumer surplus techniques.[8]

It is also important to emphasize that the definition of V used in arriving at EGDP is a general equilibrium concept involving both X and Y and that V must be estimated accordingly. The implication of this point is not readily apparent in figure 1.4 because it is drawn with indifference curves that shift in a parallel way and because the θ is the same for both X and Y. In this

7. Figure 1.4 is a simplified formulation from Hulten and Nakamura (2018). It is meant to illustrate the underlying role of a utility-enhancing shift in the consumption technology in a general equilibrium context. We have adopted a utility function that embodies simplifying assumptions. The indifference curves of $U(X, Y)$ are homothetic (radial blowups of a base curve), so the shifts have a neutral effect on the consumption Y/X ratio when relative prices do not change.

8. V in these equations is defined as the *distance* between the indifference curves in two time periods, and θ refers to the *rate* at which the consumption technology shifts over the interval. The interval may refer to one year (the simple case analyzed in this section) or the cumulative effects of many years. In general, V should not be used as a direct measure of θ and therefore should not itself be added to annual GDP to arrive at EGDP unless adjusted for the time horizon involved to get at θ.

situation, the expansion path of the economy, $0G$, is a straight line, and the price ratio P^X/P^Y is constant. When there are separate rates for each good, θ_X and θ_Y, the price ratio P^X/P^Y can change, as can the expansion path. In this case, the EV and CV differ, since they reflect different ratios. This is a familiar problem, but it implies that a partial equilibrium estimate of either θ_X or θ_Y separately holding the price of the other good constant, V_X or V_Y, does not capture the full impact of the change in the θ. Moreover, the sum of the resulting partial equilibrium V_X or V_Y is not equal to the general equilibrium V except under very strong restrictions on the utility function (Varian 1992). This, too, should be kept in mind when evaluating studies that add a partial equilibrium estimate of the willingness to pay for various technology goods to annual GDP.

1.3.4 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. The first is product-*dis*embodied innovation, μ, 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). The second is product-embodied innovation in consumption goods, which itself comes in two forms: improvements in the design of existing goods (quality change) and the advent of innovative new goods that embody characteristics not seen before or not available in past years.

Quality change and new goods share the common feature that they are goods that embody desirable new features. However, they differ in the way the features affect utility. In the first, new varieties of *existing* goods enter the market with superior characteristics, and it is common to treat the superior variety as though it were equivalent to having more of the inferior variety it replaces. In terms of figure 1.4, this treats the good X_{1e} as a multiple $(1 + \beta)X_0$, holding μ and λ constant and letting β denote the rate of quality change (also, Y_{1e} is a multiple $(1 + \beta)Y_0$). In this formulation, "better" is assumed to be equivalent to more. This approach incorporates product quality innovation at a rate β into the analysis of figure 1.4 symmetrically with μ. Both are calibrated using the equivalent increase in the bundle (X_0, Y_0). The sum of the two equals the rate of output-saving innovation—that is, $\theta = \mu + \beta$.

The compensating variation V developed in figure 1.4 provides a metric for a generic θ but could in principle be applied to μ and β separately. However, because the latter is embodied in products that are transacted in markets, there is another avenue of approach to the problem of estimating β based on prices. It exploits the fact that, because the change in utility is assumed to be costless, the amount of money spent to purchase the quantities (X_{1r}, Y_{1r}) is the same as the amount associated with (X_{1e}, Y_{1e})—that is, that $P_0^X X_{1r} + P_0^Y Y_{1r} = P_0^{Xe} X_{1e} + Y_{1e}$. The P^{Xe} and P^{Ye} are the shadow prices of the effective outputs X_{1e} and Y_{1e} and are denominated in equivalent units (we assume here that there is no pure price inflation, so the accounting can be done in base-year prices). Since the same expenditure $P_0^X X_{1r}$ allows the consumer to acquire X_{1e}, P_0^{Xe}, $P_0^X X_{1r} = P_0^{Xe} X_{1e}$. It then follows that P_0^X / P_0^{Xe} is equal to the X_{1e}/X_{1r}, which in turn equals (1 + β). Thus as utility increases by the factor β, the cost of acquiring this utility falls. This formulation reduces the problem of estimating β to the problem of estimating the relevant price ratio. We will revisit this approach in the sections that discuss the associated empirical procedures and problems.

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 by consumers and that the quantity demanded would increase to the point at which the gap between the new marginal utility and acquisition price would be extinguished. However, the opportunity for this arbitrage does not exist in all cases. When a superior pharmaceutical drug arrives in the marketplace, the individual consumer does not respond by buying more of the drug until the marginal utility equals the old one but instead purchases the new standard regimen. Nor do people necessarily usually purchase more personal computers as their efficiency increases and the efficiency price falls; there may even be a shift to less-expensive tablets. 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 utility-enhancing innovation.

1.3.5 Quality Change Embodied in New Goods

The treatment of quality change in its β form relies on the assumption that "better" can be measured in terms of more of an inferior good. This is a tidy solution that locates β in the theoretical framework of figure 1.4 and is useful for empirical work. But "better as more" embodies the paradox that a good that is sufficiently superior that it needs separate treatment is also essentially a multiple of the replaced good. However, it may be more accurate to regard the superior variety as a new good that offers capabilities that the previous version did not. Again, a pharmaceutical drug with a high degree of efficacy does not achieve the same outcomes as multiple doses of an earlier treatment with a low degree of efficacy.

Unfortunately, treating a significant change in the β quality as a new good leads to a host of other problems. From a theoretical standpoint, a new good Z cannot be located on the XY axis of figure 1.4. It appears on a new Z axis and becomes incorporated in GDP as $P^X X + P^Y Y + P^Z Z$. Because of the sudden appearance of the Z, there is no prior price or quantity with which to estimate the gain in consumer utility from its arrival. The Hicks-Rothbart solution is to regard quantity of Z as zero prior to its introduction because its theoretical price was too high and there was zero consumer demand. The solution posits the existence of a "reservation" price that is just low enough to attract consumers to the market for Z. The difference between the reservation price and the actual price prevailing when the good is introduced is then used as a measure of the increase in utility resulting from the arrival of Z. The empirical problem is then to estimate this reservation price.

It should also be noted that the implementation of the reservation price approach requires econometric modeling. This, in turn, requires assumptions and procedures that lie outside of the normal sphere of data measurement. It is also time consuming and must be repeated for each new good, so it is not economical for use in statistical programs that produce annual data series that must be internally consistent over time. This problem applies to the Bureau of Labor Statistics (BLS) price program, and they thus use an imputation procedure that, as we shall see, has the general effect of linking the new good to the subcategory to which it is assigned at, or near, the mean value of the other goods in the subcategory. This way of incorporating new goods into the price indexes used to compute real GDP is conceptually the same as the way it treats quality change in existing goods, except that it refers to quality change in a class of goods that may or may not be closely related. This approximation procedure may thus miss much of the value of the innovation embodied in truly new goods like the internet.

1.4 The Estimation of Innovation and EGDP

1.4.1 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 ninefold 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 percent 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 well-being of individuals. The world of

1820 lacked effective medical treatments for most serious afflictions. The discovery of the germ theory of infection by Joseph 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 (today, it is hard to imagine surgery without it). Antibiotics in the 20th century allowed the treatment of routine infections 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 well-being. The medical revolution proceeds apace with important breakthroughs in surgery (noninvasive, 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 20th 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 19th 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 well-being of families, as did residential air conditioning. Many advances have come since the mid-20th century. As recently as 1950, a quarter of America's homes had no flush toilet, according the US Census Bureau housing data. In 1990, only 1 percent of US 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 one in five households had no telephone available. Wood was used as a major heating fuel in 1940 (23 percent) but virtually disappeared by 1970 (only 1.3 percent). 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 fraction of adults 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 1 in 10 in 2005 to two-thirds in 2015, and other Pew surveys found that the market penetration of smartphones more than doubled from 2011 to 2016, from 35 percent to 77 percent.[9] The rapid uptake was matched by 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 1 gigabit are available in a few locations, and 100 Mb and higher speeds are

9. US Census Bureau (2014); Perrin (2015); Pew Research Center (2017); Anderson (2015).

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 10,000 times, or a 40 percent annual rate.

1.4.2 Sorting Out the λ, μ, and β Effects

The overview of the preceding section suggests that a high degree of innovation activity accompanied a sustained growth rate of real GDP. The question raised in this chapter is whether the gains in individual well-being are fully valued by the corresponding gains in income per capita and, if not, how much additional welfare was generated by a shift in what we have called the consumption technology. In more precise parametric terms, innovation enters the picture via the λ, μ, and β. The remaining sections of this chapter review a more detailed look at the link between the growth in real GDP per capita and the growth in consumer well-being and EGDP, with a view toward assessing their potential magnitude and the implied biases vis-à-vis current statistical practice.

The parameters λ, μ, and β and intangible capital are part of the larger framework underlying the figures. We have studied this framework in the two-sector (X, Y) case, but the problem at hand involves the impact of innovation on the growth rates of aggregate real GDP per capita and individual welfare, so it is appropriate to reformulate the problem in a one-sector form. The various components of interest come together to form the basic framework linking the growth in welfare per capita, $u - \ell$, to the growth in output per worker, $(q^r - \ell)$, and the parameters of output-saving innovation. This yields the basic economy-wide sources-of-welfare-growth equation of this chapter:

$$u - \ell = \mu + \beta + (q^r - \ell). \tag{1}$$

This equation indicates that the representative person's welfare depends on both the amount of income they have and how well they use it. (The variable $q^r - \ell$ here represents the growth rate of output per worker measured at resource cost, not effectiveness.) The term $q^r - \ell$ can be further decomposed to yield the conventional Solow sources-of-output-growth equation:

$$q^r - \ell = \lambda + \mathrm{v}_K(k - \ell) + v_N(n - \ell). \tag{2}$$

This second equation indicates that the growth rate of output per worker is composed of the following elements: the growth rate of tangible capital per worker $(k - \ell)$ and the growth rate of intangible capital per worker $(n - \ell)$, each weighted by their respective income shares, v_E and v_N. These income shares are proxies for the corresponding elasticities of output in the standard Solow sources-of-growth framework. The λ measures the resource-saving technical change, while $v_N(n - \ell)$ is a measure of resource-using intangible innovation.[10]

10. A more detailed description of the sources-of-growth model and the role of the income shares is given in the survey by Hulten (2001).

Two elaborations of (1) and (2) are necessary for the empirical literature described in the following sections. As previously noted, the statistics on real GDP in the United States embody a correction for quality change, implying that the observed growth rate is $q^e = q^r + \beta$ if the correction for β is complete and accurate. This correction implies, in turn, that equation (1) must also be modified to account for the fact that the use of q^e as the output growth means that β is suppressed into output and does not appear explicitly in (1), with the result that

$$u = \mu + (q^e - \ell) = \lambda^e + v_K(k - \ell) + v_N(n - \ell). \tag{3}$$

The q^e-based TFP residual conflates the true λ productivity, the shift in the production function, with the quality effect, with the result that $\lambda^e = \lambda + \beta$. In other words, the use of real GDP, as presented in official statistics, has the effect of concealing the true shift in the production function, unless the magnitude of β is known. However, the size of β is nowhere shown in the official statistics.

A second modification of this framework is needed because, as we shall see, the β that gets embedded in q^e and λ^e is estimated with a significant degree of bias, giving β' instead. The bias in β results in a corresponding bias in output growth, which becomes $q^{e\prime} = q^r + \beta'$. When this biased estimate is used in place of q^e, the growth equation becomes

$$u = \theta + [\beta - \beta'] + (q^{e\prime} - \ell) = \lambda^{e\prime} + v_K(k - \ell) + v_N(n\ell). \tag{4}$$

The $q^{e\prime}$-based TFP residual now conflates the productivity effect and the biased quality effect, with the result that $\lambda^{e\prime} = \lambda + \beta'$. As before with (3), neither the biased β' nor the degree of bias $[\beta - \beta']$ is recorded in official statistics. However, there are numerous occasional studies of the bias in price statistics that can be used to get an impression of its potential magnitude.

1.5 The Supply-Side Contribution to Overall Growth

1.5.1 The Sources of Output Growth

The sources-of-growth results for the US *private business economy*, based on equation (4), are shown in table 1.1 for the period 1948 to 2007. A version of this 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 percent. The sources of this growth are reported in the rows of table 1.1, 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 percent of the 2.4 percent output growth, of which 10 percent came from information and communications

Table 1.1 **Sources of growth in US private business sector (average of annual growth rates)**

	1948–2007	1948–73	1973–95	1995–2007
1. Output per hour [q^e—ℓ]	2.41	2.99	1.56	2.76
Percentage point contribution to output per hour of:				
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
Percentage of total contribution to output per hour of:				
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 to total factor productivity. The latter includes both β' and λ terms, since the hyperoutput 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 β' rather than a true β.

Source: Corrado and Hulten (2010, 2015).

technology (ICT) equipment per worker hour. Intangible capital contributed 17 percent, of which only 4 percent came from formal R&D. Changes in the composition of the workforce added 8 percent, while the TFP residual explained by the other sources made the largest contribution at 47 percent.

These estimates refer to the period as a whole. A look at the subperiods reveals some important within-period trends. It is significant for the taxonomy of innovation presented in section 1.3 that the long-term trend in TFP moved downward since the 1960s. TFP grew at an average annual rate of 1.8 percent over the period 1948 to 1965 and explained almost half of the growth rate of output per worker hour; the growth rate fell to 1.2 percent in the most recent period, 1995 to 2007, and its contribution to output growth fell from 60 percent to just over 40 percent. The declining trend in TFP is also evident in figure 1.5, 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 1.1, followed a generally

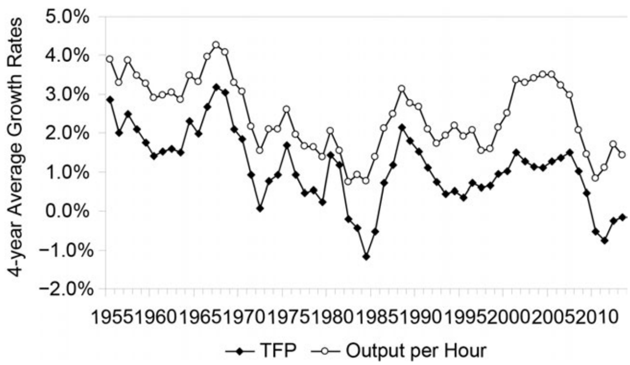

Fig. 1.5 Growth in output per hour and TFP, US NFB, 1955–2011

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 termed output-saving innovation) toward costly resource-using innovation, as represented 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 technological complexity rises, 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.

Resource-saving and resource-using technical changes are not the only factors in the innovation process. 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 1948–73 to 1995–2007, from 0.56 percent in the earlier period (19 percent of overall growth) to 1.31 percent (a 47 percent contribution). Thus the relative contribution of TFP has declined, but innovation and its correlates have not, although the composition has changed.

1.5.2 Critique of the Growth Accounting Results

Growth accounting produces estimates that are by far the most secure results in the empirical chain linking resources and technology to EGDP in figure 1.3. They are supported by national accounting data 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, which are partly due to spillover externalities of technical knowledge whose property rights are hard to protect, but also nontechnological factors such as the omitted variables like infrastructure capital, nonmarket 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 β and λ.

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 a 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 procyclicality of TFP seen in figure 1.5. It is for this reason that the time period covered in table 1.1 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 US 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 λ 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 (North American Industry Classification System [NAICS] industries 54 through 81) over the period 1987–2015. 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 percent); Ambulatory Health Care (−0.4 percent); Hospitals, Nursing, and Residential Care (−0.9 percent); Management of Companies and Enterprises (−0.4 percent); and Legal Services (−0.3 percent).

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.[11] Indeed, negative TFP 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" (7). He thus labeled the industries we are discussing as hard-to-measure industries. A consequence is that there 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^r right, much less the efficiency-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 and 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" (286).

In defense of the BLS program, the BLS website that presents the non-

11. The Baumol disease explanation of the lower 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 output per worker 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 services rose from an average annual growth rate of 0.7 percent during the 1987–95 period to 2.6 percent in the years 1995–2001; for the goods-producing sector, the corresponding numbers were 1.8 percent and 2.3 percent, respectively. They also find that 80 percent 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.

manufacturing industry productivity estimates contains this 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."[12] It is hard to criticize the BLS for not fully solving the problems with service sector output measurement highlighted by Griliches.

1.5.3 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 term in (3) and table 1.1. 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 its intangible nature makes the extent of its presence hard to detect. Moreover, own-account production does not generate an explicit price and quantity from which its quantity and value can 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. As previously noted, the BEA moved in 2013 to capitalize R&D and artistic originals and to add them to GDP rather than treating them as within-firm intermediate goods that do not find their way into 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 Corrado, Hulten, and Sichel (2005, 2009). The estimates in table 1.1 are based on an updated version of the Corrado-Hulten-Sichel framework and thus differ from those presented in the BLS productivity tables, which include only a partial list.[13]

1.6 Estimates of Innovation on the Consumption Side of the Economy

1.6.1 An Overview of the Problems Involved

The previous section reviewed the empirical work on the two main variables of supply-side innovation: the Solow residual and the intangible capital effect. We turn now to the consumption side and the variables that shift the consumption technology, μ and β. This type of innovation is inherently

12. Multifactor Productivity and Related KLEMS Measures from the NIPA Industry Database, 1987 to 2016 (https://www.bls.gov/mfp/mprdload.htm).

13. One consequence of capitalized intangibles is that the relative importance of TFP as a source of growth falls from 50 percent to 39 percent when moving from the BLS TFP estimates to the fuller list (Corrado and Hulten 2014, table 3). Another consequence is that the resulting investment is added to GDP, which is thereby increased in size, but not so much in its rate of growth, which is only modest.

more difficult to measure because it involves a shift in utility, for which there are no regularly published estimates, whereas production-side innovation involves output, for which such estimates are available. Moreover, the latter is based on well-established concepts, while the factors that shift the consumption technology are new to this chapter. However, conventional statistical practice does include some of the effects of β in the adjustment of output for quality change, although the implied β is not shown explicitly and is associated with production, not consumption. Measuring the effects of μ is even more of a challenge, since it is not embodied in specific goods, though it does emanate from goods (as internet information does from computers and smartphones). This example points to another complication that arises because μ and β are linked in ways that make them hard to separate (medical care offers numerous other examples, such as the computer-based machinery that enables minimally invasive surgery).

This said, estimating μ and β can at least be approached via individual studies of their value as revealed by consumer preference. We will review some of these sources of information in the remaining sections of this chapter. We first focus on the measurement of quality change and the evidence of the potential size of β found in academic research and government programs. Much of the literature relating to β is actually about the bias with which β is estimated in official statistics, which is the rationale for the reformulation of our basic model to include the explicit bias term $[\beta - \beta']$ in (4). We postpone our discussion of the disembodied term μ in sections dealing with the internet, health care, and education.

1.6.2 Estimates of Product Quality Change

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; Groshen et al. 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" (124). This combat has, however, produced some notable victories, and the case of computers is a salient example. The BEA makes a quality adjustment to the output price of computers and peripheral equipment in personal consumption expenditures in order to reflect the advances in computing power enabled by Moore's law, with the result that the price fell at an average annual rate of −1 percent from 1960 to 1985, then by −21 percent per year from 1985 to 2000, followed by a −11 percent decline from 2000 to 2015. These declines imply a high rate of quality-induced price change P^e when compared to a baseline scenario of no change in the resource price P^r. And computers are not the only example of rapid quality change. The BEA's prepackaged computer software and accessories price deflator also includes an adjustment for quality change (Abel, Berndt, and White 2007), and it declined at an average annual rate of −17 percent over the period 1985–2000 and by −5.5 percent from 2000 to 2015.

Moore's law applies to goods directly affected by the silicon revolution, like computers, but its reach is far wider. Computer chips and software are embedded in many devices, from smartphones to vehicles and machine tools. Byrne and Corrado (2017b) provide estimates of the implied wired telecommunications services deflator based on measures of the improving quality (and rapid deflation) of telecommunications equipment developed in Byrne and Corrado (2015) and methods described in Byrne and Corrado (2017a). They do so for *nonresidential* wireless services rather than for personal consumption expenditures, and they find a rate of deflation 7 percentage points below the official measures from 2004 to 2014. This study points to the need to distinguish between the quality change in a good that accrues to consumers and that which affects the supply of goods passing through markets.

As for a broader range of goods, 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 percent showing up as pure price inflation, and conclude that BLS procedures do not fully account for the impact of quality upgrading.

Some mention must also be made of product innovation brought to marketplace in the form of new goods. Hausman (1996) examined the introduction of a new brand of breakfast cereal and found that the treatment of new goods in official statistics missed a significant amount of the innovation that had occurred. Hausman's 1999 study of the introduction of mobile cellular telephones reached the same conclusion.

1.6.3 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 the source of many of the price statistics used by the BEA to derive real GDP, but its main task is to prepare 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 (CPI) 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. Since an improvement in product quality provides more "bang for the buck" for each dollar spent and offsets the inflationary erosion, it 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 marketplace and find their way into the basket, and others 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 (US Bureau of Labor Statistics 2018).

The prescribed procedures are complicated and not easy to summarize. Fortunately, the survey by Groshen et al. (2017) 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. (2017) report that for the period from December 2013 to 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" (190–91). These percentages are on a monthly, not annualized, basis. They go on to say,

> 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. (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 (US Bureau of Labor Statistics 2018), 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. (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 marketplace, 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). Thus, as previously noted, 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, on 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]" (US Bureau of Labor Statistics 2018). This rapid substitution is a 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.[14] Groshen et al. (2017) report that "in the Consumer Price Index, about 33% 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" (192).[15] These statistics suggest that very few item categories are subject to the hedonic method, despite the recommendation of the Stigler Commission (1961) review of price measurement that specifically referred to the Griliches study

14. 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.

15. 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.

of hedonics in new cars. For its treatment of the price of cars, the BLS uses a measure 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.

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 the bureau is 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 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 and the changing quality of goods, but again, this is not its main mission. One consequence is that the BLS does not report the amount of the quality correction it makes—its implicit estimate of β. That is embodied in its price estimates that are used for output deflation by the BEA.

1.6.4 The Bias in Quality Measurement

More attention has been given to the size of the implied bias in the price deflators (and the bias in β) than to the size of β itself. The subject has generated numerous studies, articles, and conference volumes (including some in the Conference on Research in Income and Wealth Studies in Income and Wealth series). These studies tend to produce mixed results about the size of the CPI bias. The estimates by Groshen et al. (2017) present a recent assessment of the overall bias based on past studies (including Lebow and Rudd 2003; Greenstein and McDevitt 2011). They put the downward bias in the annual growth of real GDP at −0.26 percent in 2015 due to consumer goods and at −0.15 percent due to private investment (real GDP growth was around 2.0 percent in that year). The former is particularly relevant to this chapter, since the "PC services (including internet)" component of the −0.26 percent downward bias was only –0.04 percent (the contribution of medical bias was −0.12 percent). The "raw" annual bias in PC/internet services was an annual −6.50 percent (based on Greenstein and McDevitt 2011), but the GDP share of this category was so small that the share-weighted growth bias barely moved the GDP needle.

Other studies have also found larger biases than those reported in Groshen et al. (2017). Bils (2009) concludes that "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 a concomitant increase in leisure for the driver and reduction in accidents, is not likely to appear in measures of output. The aforementioned 2001 Bils and Klenow study of 66 durable goods also concluded that BLS procedures do not fully account for the impact of quality upgrading. Other studies are consistent with this conclusion. 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 80 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 percent for software to 12.0 percent for computers and peripherals (the Greenstein and McDevitt [2011] estimate is in the middle of this range).

The four studies of the value of broadband evaluated by Syverson (2017) provide estimates of consumer surplus that he extrapolates to 2015 that range from a low of $17 billion to a high of $132 billion, including the Nevo, Turner, and Williams (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.[16] 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.

1.7 Information, the Internet, and Consumption Technology

1.7.1 The Nature and Value of Information

Measuring the amount of 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,

16. See US Bureau of Labor and Statistics (2019).

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 chapter, 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 for that which is in fact provided free of direct charge. We have formulated this as the parameter μ. The magnitude of this parameter, as measured by the willingness-to-pay metric V of section 1.3, 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 chapter, and to this end, the rest of this section will marshal the available evidence on the size of V and μ.[17]

1.7.2 Current Treatment of Information in the Statistical System

BEA data from the US 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 percent 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 percent. A still narrower grouping with a focus on information services includes only "Data processing, Internet publishing, and other information services" (1.5 percent) and "Computer systems design and related services" (0.6 percent). 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" show that consumers spent \$230 billion on the categories "Telecommunication services" and "Internet access" in 2016, or 1.8 percent of PCE and 1.2 percent 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 percent of PCE and 2.0 percent of GDP. By way of comparison, Groshen et al. (2017) report a GDP share for the category "PC services (including

17. Any attempt to assess the role of information in promoting consumer utility should recognize its public good nature. It is nonrival (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.

Internet)" of 0.6 percent 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 is 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 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 results 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, in their annual reports 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 prices 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 2016; Sichel and von Hippel 2019), or crowdsourced 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.[18] 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.

1.7.3 The Measurement Literature on the Internet's Contribution to Welfare

A small but growing number of studies address the measurement issues implied by Schmidt and Rosenberg's (2014) remark that "the Internet has

18. 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 chapter is limited to an analysis of how costless increases in digital sources of information can provide consumer benefits beyond those recorded in GDP and thereby present a different assessment of economic progress.

made information free, copious, and ubiquitous." They cover both the internet and the explosion in timely information it enables but also the devices needed to enable the digital revolution. The former are associated with disembodied output-saving technical change, μ, 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 (2020).[19]

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, Jeon, and Kim (2014), 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 7 minutes compared to 22 minutes using the library. Varian calculated an implied 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 \$2,500 to \$3,800 per year. Syverson (2017) 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 in the post-2004 time period. Other creative approaches to the consumer surplus problem use questionnaires, surveys, and microdata. The literature includes Brynjolfsson, Hu, and Smith (2003); Aguiar and Waldfogel (2018); Quan and Williams (2016); and Dolfen et al. (2019).

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, Peck, and McGowan (1973). The slow diffusion of broadcast TV meant that some rural households had to pay for broadcasts that were free elsewhere. Using demand analysis, they were able to estimate that households would be willing to pay some 3 percent of income for free TV. However, such natural experiments are rare in the

19. 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.

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 engines, a much smaller number than his back-of-the-envelope welfare calculation. However, more recent work by Brynjolfsson et al. (2018) suggests that the minimum payments consumers would accept (willingness to accept or WTA) for loss of access to search engines may be as large as $5,000 a year. This estimate suggests a value of about $1 trillion missing from GDP from search alone.

A small industry has arisen in evaluating consumer willingness to accept the price of Facebook. Brynjolfsson, Eggers, and Collis (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 0.05 to 0.11 percentage points to the growth of real GDP (in other words, the increment to μ is between 5 and 11 basis points). In another study, an auction experiment conducted by Corrigan et al. (2018) puts the value of doing without Facebook for an entire year at $1,000 to $2,000 per adult person in the United States, with an implied value of as much as $250 billion to $500 billion a year. The largest-scale experiment, Allcott et al. (2019), finds a similar value.

Finally, Nakamura, Samuels, and Soloveichik (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. However, it is important to recall the caveats of section 1.33 of this chapter. The studies reviewed in this section are mostly focused on individual goods like Facebook, and the results are partial equilibrium estimates of their value and thus are incomplete efforts to get at our EGDP. While doing so is a valuable step in this direction, goods with the broad scope of Facebook and the internet are bound to affect relative prices for many other goods in the economy, and the *ceteris paribus* assumption of partial equilibrium analysis is increasingly problematic as the importance of a good increases. Moreover, the important study by Brynjolfsson, Eggers, and Collis (2019) illustrates another issue raised in passing in section 1.3.3: the aggregate willingness to accept Facebook is large in dollar terms, but when expressed as an annual rate rather than a cumulative total, the contribution to GDP is found to amount to only 0.05 to 0.11 percentage point.

1.8 Health and Education: Individual Heterogeneity and the Role of State Contingency

The consumption technology as formulated in this chapter refers to the average state of health or knowledge, whereas much of the actual gain from innovation is contingent on an individual'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 lifestyles and preventative medicine. Moreover, education and health interventions may interact in ways that strengthen each other.

A health care innovation, such as minimally invasive surgery, will generally affect a subset of the population, and perhaps only a small subset. The gains to those affected may be quite large, but they 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 the EGDP program 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 EGDP. This is a major step, and since the basic thrust of this chapter is to explore the EGDP concept per se, it is a step we will defer to subsequent research.

1.8.1 Innovation in Health Care

The review of the bias in price statistics by Groshen et al. (2017) identified health care as a major source of the accuracy problem. 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 forms 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 2001–14. With health care expenditures accounting for 17 percent 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 attacks, 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 of an 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 and Topel (2006) calculation of the value of the 20th century increases in life expectancy from 48 to 72 finds a very large number, $1.2 million per person, for the representative person in 2000 in the United States. However, 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). Still, taken together, these health care studies highlight the importance of outcomes (longevity, mortality rates) as opposed to expenditures.

Another example of utility-enhancing technical change comes from the recent study by Rothwell et al. (2016), who found that taking aspirin for 12 weeks following a stroke or ministroke 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.

1.8.2 The Case of Education

There have been major gains in educational attainment in the United States 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.[20] 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 and maturation are the relevant outcomes. 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 . . . 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" (8).

Improvements in the outcomes of historically underserved student populations have a large payoff to society and, importantly, to those individuals who stand to benefit. Tracking the gains to the average student will tend to understate the gains to this population not only in terms of increased personal income but also in the nonmonetary improvements in the quality of their life. Subsuming these gains in a measure based on average experience thus risks missing some of the most important welfare benefits of improved educational outcomes.[21]

1.9 Final Thoughts on the Path Ahead for EGDP Measurement

In his 1994 American Economic Association (AEA) presidential address, Griliches observed, "It is not reasonable for us to expect the government to produce statistics in areas where concepts are mushy and where there is little professional agreement on what is to be measured and how" (14). This observation applies in full force to the current measurement problems associated with the technological revolution currently underway. These problems are as much a matter of inadequate theoretical development as of inadequate statistics. Addressing the former is the rationale for our current work. To this end, we have proposed the theoretical construct of expanded GDP as a new measure of aggregate economic activity that builds on existing GDP. Our review of the empirical literature and the available data suggests that this

20. 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.

21. Quality-adjusted labor is considered exogenous in our discussion, but education partially endogenizes it.

effect is nonnegligible, 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 EGDP growth can be quite large despite this small share (Hulten and Nakamura 2018). In our previous study, we conducted a thought experiment in which the bias in price deflators noted by Groshen et al. (2017) 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 1.1 of this chapter is 2.76 percent for the period 1995–2007, 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 consumption technology and its utility-enhancing effect are potentially too large to be ignored. We recognize that adding a consumption technology to the conventional GDP framework is by no means an easy task, and one not to be undertaken lightly. Part of the value of GDP lies in the continuity of the time record that allows for meaningful comparisons with past eras, and 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.

Fortunately, the BEA has already made a start in this direction with its innovation accounting and limited capitalization of intangible assets. This innovation accounting 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 the BLS and the BEA to work together to improve price statistics so that they more accurately reflect and classify product innovation. Taking on the challenge posed by new goods, like the internet and mobile communication devices, is of central importance in this regard. Yet 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 price estimates. Finally, the research from the "outcome movement" in health care research should be accorded a high priority.[22]

The task of building a full innovation satellite account is daunting. The history of the national accounts is a history of overcoming one daunting

22. It must also be said that the BLS is continually working to improve the CPI and the PPI. For example, it is moving to what has been called a diagnosis or a disease-centric approach (Roehrig 2017). The BEA has also made much progress on the problem of measuring outcomes in the provision of health care services, but the path ahead is long and difficult.

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."[23]

References

Abel, Jaison R., Ernst R. Berndt, and Alan G. White. 2007. "Price Indexes for Microsoft's Personal Computer Software Products." In *Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches*, edited by Ernst R. Berndt and Charles R. Hulten, 269–89. NBER Studies in Income and Wealth 67. Chicago: University of Chicago Press.

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. 2018. "Quality Predictability and the Welfare Benefits from New Products: Evidence from the Digitization of Recorded Music." *Journal of Political Economy* 126 (2): 492–524.

Allcott, Hunt, Luca Braghieri, Sarah Eichmeyer, and Matthew Gentzkow. 2019. "The Welfare Effects of Social Media." NBER Working Paper No. 25514. Cambridge, MA: National Bureau of Economic Research.

Anderson, Monica. 2015. "Technology Device Ownership: 2015." Pew Research Center, October 29. http://www.pewInternet.org/2015/10/29/technology-device-ownership-2015.

Bils, Mark. 2009. "Do Higher Prices for New Goods Reflect Quality Growth or Inflation?" *Quarterly Journal of Economics* 124 (2): 637–75.

Bils, Mark, and Peter J. Klenow. 2001. "Quantifying Quality Growth." *American Economic Review* 91 (4): 1006–30.

Brynjolfsson, Erik, Avinash Collis, W. Erwin Diewert, Felix Eggers, and Kevin J. Fox. 2020. "Measuring the Impact of Free Goods on Real Household Consumption." *AEA Papers and Proceedings* 110 (May): 25–30.

Brynjolfsson, Erik, Avinash Collis, and Felix Eggers. 2019. "Using Massive Online Choice Experiments to Measure Changes in Well-Being." *Proceedings of National Academy of Science* 116 (15): 7250–55.

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 (11): 1580–96.

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

Byrne, David, and Carol Corrado. 2017a. "ICT Asset Prices: Marshaling Evidence into New Measures." Finance and Economics Discussion Series 2017-016. Washington, DC: Board of Governors of the Federal Reserve System. https://www.federalreserve.gov/econres/feds/files/2017016r1pap.pdf.

Byrne, David, and Carol Corrado. 2017b. "ICT Prices and ICT Services: What Do They Tell Us about Productivity and Technology?" Finance and Economics Discussion Series 2017-015. Washington, DC: Board of Governors of the Federal Reserve System. https://doi.org/10.17016/FEDS.2017.015.

23. Cited by Landefeld (2000).

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*. Washington, DC: Brookings Institution Press.
Chen, Yan, Grace YoungJoo Jeon, and Yong-Mi Kim. 2014. “A Day without a Search Engine: An Experimental Study of Online and Offline Searches.” *Experimental Economics* 17 (4): 512–36.
Chernew, Michael, David M. Cutler, Kaushik Ghosh, and Mary Beth Landrum. 2016. “Understanding the Improvement in Disability Free Life Expectancy in the U.S. Elderly Population.” NBER Working Paper No. 22306. Cambridge, MA: National Bureau of Economic Research.
Corrado, Carol A., and Charles R. Hulten. 2010. “How Do You Measure a ‘Technological Revolution’?” *American Economic Review* 100 (2): 99–104.
Corrado, Carol A., and Charles R. Hulten. 2014. “Innovation Accounting.” In *Measuring Economic Progress and Economic Sustainability*, edited by D. W. Jorgenson, J. S. Landefeld, and P. Schreyer, 595–628. NBER Studies in Income and Wealth 72. Chicago: University of Chicago Press.
Corrado, Carol A., and Charles R. Hulten. 2015. “Financial Intermediation in the National Accounts: Asset Valuation, Intermediation, and Tobin’s q.” In *Measuring Wealth and Financial Intermediation and Their Links to the Real Economy*, edited by C. R. Hulten and M. B. Reinsdorf, 125–47. NBER Studies in Income and Wealth 73. Chicago: University of Chicago Press.
Corrado, Carol A., Charles R. Hulten, and Daniel E. Sichel. 2005. “Measuring Capital and Technology: An Expanded Framework.” In *Measuring Capital in the New Economy*, edited by Carol Corrado, John Haltiwanger, and Daniel Sichel, 11–46. NBER Studies in Income and Wealth 65. Chicago: University of Chicago Press.
Corrado, Carol A., Charles R. Hulten, and Daniel E. Sichel. 2009. “Intangible Capital and US Economic Growth.” *Review of Income and Wealth* 55 (3): 661–85.
Corrigan, Jay R., Saleem Alhabash, Matthew Rousu, and Sean B. Cash. 2018. “How Much Is Social Media Worth? Estimating the Value of Facebook by Paying Users to Stop Using It.” *PLoS ONE* 13 (12). December 19. https://doi.org/10.1371/journal.pone.0207101.
Coyle, Diane. 2014. *GDP: A Brief but Affectionate History*. Princeton, NJ: Princeton University Press.
Cutler, David M., and Ernst R. Berndt. 2001. “Introduction.” In *Medical Care Output and Productivity*, edited by David M. Cutler and Ernst R. Berndt, 1–12. NBER Studies in Income and Wealth 62. Chicago: University of Chicago Press.
Dauda, Seidu, 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.
Dolfen, Paul, Liran Einav, Peter J. Klenow, Benjamin Klopack, Jonathan D. Levin, Larry Levin, and Wayne Best. 2019. “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, Austan, 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): 108–13.
Gordon, Robert J. 2016. *The Rise and Fall of American Growth: The U.S. Standard of Living Since the Civil War*. Princeton Economic History of the Western World. Princeton, NJ: Princeton University Press.

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 to *Output Measurement in the Service Sectors*, edited by Zvi Griliches, 1–22. NBER Studies in Income and Wealth 56. Chicago: University of Chicago Press.
Griliches, Zvi. 1994. "Productivity, R&D, and the Data Constraint." *American Economic Review* 84 (1): 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): 187–210.
Hausman, Jerry. 1996. "Valuation of New Goods under Perfect and Imperfect Competition." In *The Economics of New Goods*, edited by T. Bresnahan and R. Gordon, 209–37. NBER Studies in Income and Wealth 58. Chicago: University of Chicago Press.
Hausman, Jerry. 1999. "Cellular Telephone, New Products, and the CPI." *Journal of Business and Economics Statistics* 17 (2): 188–94.
Hill, Peter. 1999. "Tangibles, Intangibles and Services: A New Taxonomy for the Classification of Output." *Canadian Journal of Economics* 32 (2): 426–46.
Hulten, Charles R. 1978. "Growth Accounting with Intermediate Inputs." *Review of Economic Studies* 45 (3): 511–18.
Hulten, Charles R. 2001. "Total Factor Productivity: A Short Biography." In *New Developments in Productivity Analysis*, edited by Charles R. Hulten, Edwin R. Dean, and Michael J. Harper, 1–47. NBER Studies in Income and Wealth 63. Chicago: University of Chicago Press.
Hulten, Charles. 2015. "Measuring the Economy of the 21st Century." *NBER Reporter* 4:1–7.
Hulten, Charles, and Leonard Nakamura. 2018. "Accounting for Growth in the Age of the Internet: The Importance of Output-Saving Technical Change." NBER Working Paper No. 23315. Cambridge, MA: National Bureau of Economic Research.
Hulten, Charles R., and Valerie A. Ramey. 2019. "Education, Skills, and Technical Change, Implications for Future U.S. GDP Growth: An Introduction and Overview." In *Education, Skills, and Technical Change: Implications for Future U.S. GDP Growth*, edited by Charles R. Hulten and Valerie A. Ramey, 1–19. NBER Studies in Income and Wealth 77. Chicago: University of Chicago Press.
Lancaster, Kelvin J. 1966. "A New Approach to Consumer Theory." *Journal of Political Economy* 74 (2): 132–57.
Landefeld, J. Steven. 2000. "GDP: One of the Great Inventions of the 20th Century." *Survey of Current Business*, January, 6–14.
Lebow, David, and Jeremy B. Rudd. 2003. "Measurement Error in the Consumer Price Index: Where Do We Stand?" *Journal of Economic Literature* 41 (1): 159–201.
Maddison, A. 2007. *Contours of the World Economy, 1–2030 AD: Essays in Macroeconomic History*. Oxford: Oxford University Press.
Moulton, Brent R., and Karin E. Moses. 1997. "Addressing the Quality Change Issue in the Consumer Price Index." *Brookings Papers on Economic Activity* 1:305–66.
Murphy, Kevin M., and Robert H. Topel. 2006. "The Value of Health and Longevity." *Journal of Political Economy* 114 (5): 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.

Nevo, Aviv, John L. Turner, and Jonathan W. Williams. 2016. "Usage Based-Pricing and Demand for Residential Broadband." *Econometrica* 84 (2): 411–43.

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

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

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): 1037–46.

Perrin, Andrew. 2015. "Social Media Usage: 2005–2015." Pew Research Center, October 8. http://www.pewInternet.org/2015/10/08/2015/Social-Networking-Usage-2005-2015.

Pew Research Center. 2017. "Mobile Fact Sheet." 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. 2020. "Measuring Aggregate Price Indexes with Taste Shocks: Theory and Evidence for CES Preferences." *Quarterly Journal of Economics* 135 (1): 503–60.

Roehrig, Charles. 2017. "A Comparison of Bureau of Economic Analysis and Bureau of Labor Statistics Disease-Price Indexes." Bureau of Economic Analysis Working Paper 2017-3. https://www.bea.gov/system/files/papers/WP2017-3.pdf.

Rothwell, Peter M., Ale Algra, Zhengming Chen, Hans-Christoph Diener, Bo Norrving, and 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 (10042): 365–75. http://dx.doi.org/10.1016/S0140-6736(16)30468-8.

Schmidt, Eric, and Jonathan Rosenberg (with Alan Eagle). 2014. *How Google Works*. New York: Grand Central.

Shapiro, Matthew D., and David W. Wilcox. 1996. "Mismeasurement in the Consumer Price Index: An Evaluation." In *NBER Macroeconomics Annual 1996*, Vol. 11, edited by Ben S. Bernanke and Julio J. Rotemberg, 93–154. Cambridge, MA: MIT Press.

Sichel, Daniel E. 1997. "The Productivity Slowdown: Is a Growing Unmeasurable Sector the Culprit?" *Review of Economics and Statistics* 79 (3): 367–70.

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.

Solow, Robert M. 1957. "Technical Change and the Aggregate Production Function." *Review of Economics and Statistics* 39 (3): 312–20.

Stigler, George, ed. 1961. *The Price Statistics of the Federal Government*. Report to the Office of Statistical Standards, Bureau of the Budget. Prepared by the Price Statistics Review Committee of the NBER. Stigler Commission. New York: National Bureau of Economic Research.

Syverson, Chad. 2017. "Challenges to Mismeasurement Explanations for the US Productivity Slowdown." *Journal of Economic Perspectives* 31 (2): 165–86.

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

US Bureau of Labor Statistics. 2018. *Handbook of Methods*. Consumer Price Indexes, chapter 17. https://www.bls.gov/opub/hom/pdf/homch17.pdf.

US Bureau of Labor Statistics. 2019. "PPI Introduces Hedonic Quality Adjustment for Internet Access Indexes." Modified March 6, 2019. https://www.bls.gov/ppi/broadbandhedonicmodel.htm.

US Census Bureau. 2014. "Computer and Internet Access in the United States: 2012." February. https://www.census.gov/data/tables/2012/demo/computer-internet/computer-use-2012.html.

Varian, Hal. 1992. *Microeconomic Analysis*. 3rd ed. New York: W. W. Norton.

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.

von Hippel, Eric. 2016. *Free Innovation*. Cambridge, MA: MIT Press.