

100 years of Economic Measurement in the Division of Research & Statistics: Beyond the streetlight

Carol Corrado, formerly of the Division of Research and Statistics
Arthur Kennickell, formerly of the Division of Research and Statistics

October 24, 2023
(Preliminary Draft, not for quotation)

This paper has been prepared for the R&S Centennial Conference, November 8, 2023. The authors thank Dave Byrne, whose contributions are close to that of a co-author, and Ari Gelfand, who provided research assistance.

Introduction

In commemoration of the 100th anniversary of the merger of the Division of Analysis and Research with the Office of the Statistician in 1923 to create the Division of Research and Statistics (R&S) of the Federal Reserve Board, the paper provides an overview and some insights into the subsequent role of measurement in the division.

In large part, R&S operates as a large factory for producing information relevant to its mission of supporting the Board through analysis of the economy and the effects of policy decisions. A small, but important, part of that information production involves the creation of primary data via survey data collection as in the case of the Survey of Consumer Finances. A much larger part involves uses of a wide variety of secondary data sources, sometimes blended with other sources. Some such secondary uses feed into the creation of higher-level published statistical series, such as the Financial Accounts of the United States or the Industrial Production (IP) and Capacity Utilization system, or to the creation of routinely used indicators such as household debt burdens. In many other cases, the informational output is the result of an exercise in modeling or estimation for internal use, such as potential output; others are shared with the public via FEDS Notes, such as a recently developed index of common expectations.¹ These estimates serve to clarify the current state of some aspects of the economy or to test the appropriateness of a particular view of economic structure. In a complex data ecosystem such as this, it is difficult at best to define neatly what constitutes “measurement,” the subject of this paper. It would not be possible in the scope of this paper to address every relevant angle. Instead, below we sketch a framework of information use and discuss some of the key related elements of measurement.

From the perspective of the present, three mutually supporting factors are most striking about the overall history of measurement in R&S. First, the varieties and sources of data have expanded enormously over time. At one time, simply having reliable national income aggregates and money supply data represented progress. Now, the growing range of relevant data—commercial data, sensor data, administrative data, and data of other sorts from multiple

¹ See Ahn and Fulton (2020).

government agencies—can seem almost boundless. Second, computing and related technologies to cope with such data have grown roughly in parallel. When two of the authors of this paper began working at the Board, anyone using more than one megabyte of computer memory on the mainframe, then the only source of computing other than a calculator or pencil and paper, was required to present a blue card signed by a division officer to a window in the computer center. Even with what was then considered an advanced computer, congestion, delays, and system crashes were often an issue. Such technology seems almost medieval, compared with current resources. These factors supported the growing complexity of the contemporary economy that required an expanded, and more detailed and sophisticated, analysis by Board staff. We will variously return to each of these factors in the discussion of specific analytical areas.

An important facet of measurement in R&S has been, and continues to be, the investigation of tensions within and across data sources for interpreting developments in the economy, or alternatively, tensions between a theoretical prior and incoming data or empirical test results. A remark in a paper by former R&S Division Director David Wilcox expresses this perspective.

When researchers test and reject an implication of a theoretical model, they usually assume that the model is in error and that subsequent investigation should be directed toward the development of alternative models that might better account for the observed characteristics of the data. They usually spend little effort investigating the characteristics of the data themselves or the suitability of the data for use in the application at hand. This paper reverses these priorities and investigates the source data and estimation methods used to construct the retail-sales and aggregate consumption data in the United States, searching especially for imperfections that might have implications for the outcome of empirical work. (Wilcox 1992, page 922)

This mindset (echoed in the paper's subtitle) plays an important role in the work that has supported and developed the division's statistical programs reviewed in the paper.

Some important measurement systems present alternative constructs that conceptually do not differ but rather make tensions in underlying source data and surveys transparent, e.g., the discrepancy between gross domestic product (GDP) versus gross domestic income (GDI) is

almost a “feature” of national accounts. Employment from the Bureau of Labor Statistics’ household survey is adjusted to its establishment employment survey’s concept by the BLS to illustrate alignment of its alternative sources for collecting data on employment (BLS does not feature this analysis, though a recent website design has made it easier for users to find).

Internally, IP is lined up with goods GDP; the IP flow-of-goods system with NIPA inventories; and, at one time at least, saving in the national accounts with its conceptually consistent counterpart in the financial accounts. Sorting through specific discrepancies is routine in the conduct of division work, but many of the puzzles or tensions that arise in incoming data are not handed to us as part of a statistical release or routine construct. They must be detected and sorted through before they can be identified as a problem in measurement rather than a problem in our economic priors.

The influence of this “beyond the streetlight” mindset goes beyond the work environment of the Board. Distinguished alums of R&S have taken lessons from studying tensions and puzzles in data to other government agencies, academia, and the broad world of policy—exports that could be said to rival the influence of the body of research undertaken by R&S per se. Furthermore, more than a dozen R&S alums serve or recently have served on advisory committees for U.S. statistical agencies, international agencies, and central banks in their capacity as experts in economic measurement. The division nurtures its influence on economic measurement research and practice by being a sponsor of the NBER’s Conference on Research on Income and Wealth (CRIW). The CRIW holds annual research conferences and a workshop at the NBER Summer Institute on economic measurement that bring together practitioners and leading researchers on topics of mutual interest.

This paper does not attempt to capture every aspect of economic measurement in R&S over the past 100 years. To do so would necessitate a far longer paper, which would be tedious reading for all but the most devoted data geeks. Instead, we review four areas of work that seem to us to be ones that are likely to have greatest salience. The first section briefly reviews the IP and capacity system, and the second section covers the history of the measurement research that worked to develop a clearer picture of productivity growth. The third reviews the

Financial Accounts of the United States (formerly Flow of Funds Accounts) from the pioneering research that led to their founding, their uses, and the development of the associated distributional accounts. The fourth section addresses the development of primary data sources, largely financial microdata, in the division, including an in-depth review of the emergence of the modern Survey of Consumer Finances. A final section concludes. That section makes note of a recent initiative to utilize a broad spectrum of sometimes non-traditional data sources to construct indicators of business conditions, the motivation for the IP index more than 100 years ago, which is where we begin.

1. The IP and Capacity System: A very brief review

The collection of financial market and banking statistics by a central bank seems an obvious duty. Less obvious is that many central banks issue nonfinancial economic statistics generated to support the implementation of monetary policy, e.g., the Bank of Japan issues the country's price indexes for producers ("corporate goods prices"), exports, and imports to support the Bank's aim of maintaining price stability.

The Federal Reserve's earliest nonfinancial statistics grew out of its need to monitor business conditions. The Board's monthly IP index presented a timely summary of monthly industrial activity 20 years before the Department of Commerce began issuing GNP and national income. Similarly, the Board's capacity utilization measures grew out of a need to monitor short-run pressures on aggregate supply that might aggravate inflation during the booming 1960s (de Leeuw 1966).

The 100th year mark of the Federal Reserve's publication of the monthly industrial production index in its Bulletin fell in 2018. A brief history of its early development through 1986 was posted on the Board's website [at this URL](#) in commemoration; it's an easy read and worthwhile checking out for the charts alone. Of course, there have been many additional developments in the subsequent 40 years that deserve additional remarks.

A recent methods development is that the estimation of monthly IP now makes use of quarterly survey data on plant utilization collected by the Census Bureau. This quarterly survey, which also yields plant operating hours, superseded the Census Bureau's annual Survey of Plant Capacity, which collected information on industry-level factory utilization in the fourth quarter of each year since the 1970s. The annual survey had been used to develop capacity estimates consistent with IP since 1989.² (Both surveys have been partially supported by the division since the early 1990s.)

A quantum improvement in both monthly and annual processing also took place since the late 1980s. Charlie Gilbert was responsible for most of these achievements, as spelled out to attendees at his recent retirement party after 40 years at the Board. (Too bad we do not have a recording of those tributes!). Of course, many R&S staff members also contributed to the remarkable automation of the present IP and capacity utilization system, including one of the division's statisticians, Bill Cleveland, and economists Eric Bartelsman and Norm Morin, to name a few.

While we do not provide an in-depth review of the emergence of the modern IP and capacity utilization system in this paper per se, our point of departure for the next section flows from a final remark in its online brief history that underscores an important aspect of the system's current relevance. The review concludes "in recent years...upgrades to the index have ... aimed at [keeping] pace with the rapid evolution of products in high-technology industries such as computers, communications equipment, and semiconductors," the subject to which we now turn.

2. IT Output, Productivity and Investment

Research contributions by R&S staff to the measurement of information technology (IT) domestic output, productivity, and investment are reviewed in this section. The review takes the paper's subtitle "beyond the streetlight" seriously and starts in the 1990s, when the importance of

² For readers interested in the estimation methods for capacity utilization, see Raddock (XXXX) and Corrado (XXXX); for its economic interpretation see Corrado and Matthey (1998).

the IT sector soared and interest in understanding the role of quality-change in inflation measures the division intensified.³ The subject of quality change was previously in the spotlight in the division when Charles Partee as R&S Division Director established a “Committee on Prices and Price Measurement” composed of academic consultants in 1965. A now classic collection of research papers edited by Zvi Griliches (Griliches 1971) was a product of this committee. We know of little else in the division on price indexes and quality change prior to this effort.⁴

Productivity in the 1990s and “The Man Who Knew”

The 1990s productivity revival in the United States began in the first half of the decade but discernment of the change took hold slowly among professional economists. Productivity analysts did not reach broad consensus that aggregate productivity growth picked up in the mid-1990s until the turn of the decade, when an acceleration in information technology (IT) capital deepening and faster trends in newly available productivity data became apparent (e.g., Oliner and Sichel 2000). By contrast, the FOMC’s call that labor productivity growth was accelerating—made by Chairman Greenspan in arguments to the committee at its September 1996 meeting—was remarkably early.

This event in monetary policy history and its beneficial aftermath for the economy has been chronicled for the public in the popular press and in books, including Bob Woodward’s *The Maestro* (2000) and Sebastian Mallaby’s biography of Alan Greenspan, *The Man Who Knew* (2016). A lesser-known aftermath is that inside the Board, the event prompted R&S to expend resources on better measuring and analyzing the IT sector of the U.S. economy.

Consider first the puzzle that surfaced in U.S. macro data in the summer of 1996: Wage rates were rising but the increase in labor costs was not boosting inflation. This conundrum could be explained by an acceleration in labor productivity (e.g., due to efficiency gains from IT-induced

³ BEA introduced quality-adjusted prices for computers in the national accounts in 1986 (Cole et al. 1986), after which IT equipment emerged as one of the fastest growing segments of aggregate demand.

⁴ Save for this goodie: As noted by Griliches in his classic paper on quality change in automobiles, which was reprinted in the Fed volume, the earliest construction of a hedonic price index was by Andrew Court while employed as an economist by the Automobile Manufacturers Association in Detroit in the 1930s (Court 1939). Files discovered while cleaning out an office in IO circa 1990 revealed that Court corresponded with staff in the old Business Condition section on his hedonic approach to measuring and analyzing price changes for automobiles.

innovations that were much discussed in the press at the time)—but a boost in productivity was not apparent in the available official data. This suggested to monetary policy commentators (including some committee members) that interest rates needed to move higher to staunch a looming pickup in inflation, a view also built into the staff forecast.

To Chairman Greenspan this line of analysis did not accord with the fact that corporate profits remained strong. In an environment of rising wages and stable prices, profits usually weaken, so he commissioned R&S staff to explore this observation. The objective was to take a “deep dive” and develop, to the extent possible, disaggregate estimates of labor productivity and unit costs for manufacturing and non-manufacturing industries as well as for corporations versus other legal forms of business. To sort through this request, it was necessary to approximate a three-way (industry, legal form of organization, time) decomposition of the available data on productivity and unit costs in the economy—a tough challenge, to be sure.

The ensuing work, portions of which were issued as a staff study and later published (Corrado and Slifman 1999), generated two primary conclusions. One conclusion was that official GDP growth was likely understated—the statistical discrepancy had swung sharply starting in 1993—and productivity calculated using income-side data (GDI) was increasing faster than the headline product-side (GDP) figure suggested; the 1997 *Economic Report of the President* later estimated that this gap implied that labor productivity from 1994 to 1996 grew 1.3 percentage points per year faster than officially reported. The other conclusion was that, using income-side data, the inferred profits of businesses in services industries (corporate and noncorporate) were well maintained while their labor productivity was estimated to be falling. How could these businesses remain profitable while their efficiency measured as real output per hour was outright dropping? One answer to this question was that the price measures used to obtain real output for these businesses were severely biased.

Using these observations, Chairman Greenspan convinced the committee that productivity was growing faster than reported in official data and that hiking the federal funds target rate was not warranted. The possibility that productivity may be mis-measured due to understated services prices or structural change not captured in national accounts likely was familiar ground

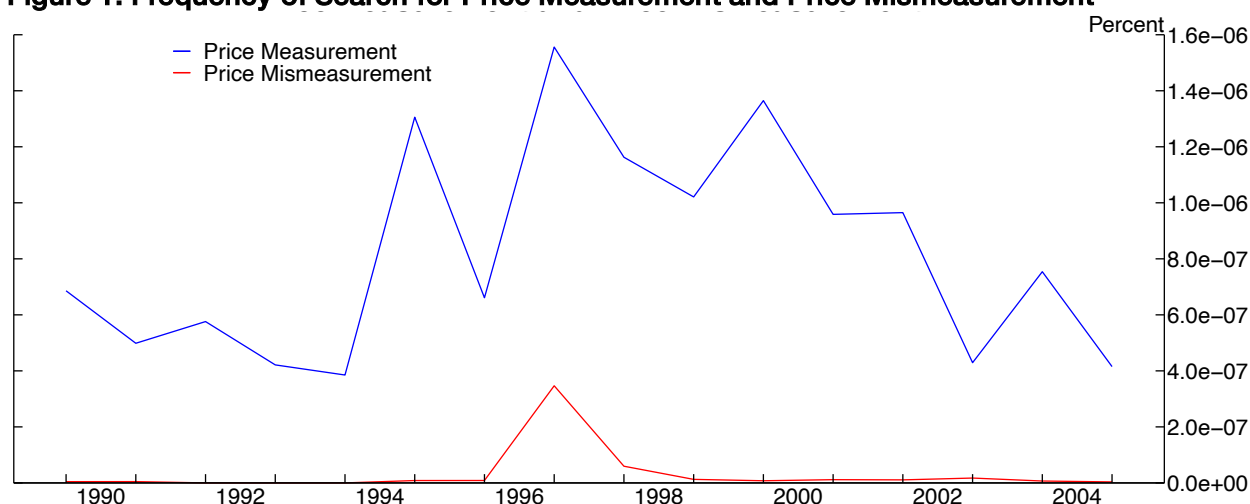
for committee members in view of the public discourse on monetary policy and technology at the time. This discourse covered issues ranging from technical issues of price measurement and discrepancies in national accounts to the “irrational exuberance” of the stock market and the possibility emerging from research using company-level data that an “IT revolution” was under way.

Measurement in the News

Chairman Greenspan famously included a critique of the CPI as a biased measure of price change in congressional testimony in January 1995—a critique drawn from an influential staff analysis (Lebow, Roberts, and Stockton 1994; see also Shapiro and Wilcox 1996; Lebow and Rudd 2003). On December 4, 1996, a group of economists appointed by the Senate Finance Committee reported on their study of the CPI, which estimated that the index overstated inflation by 1.1 percentage points annually (Boskin et al. 1996). This was less than a month after the release of the minutes of the September 1996 FOMC meeting, in which productivity figures based on income-side data and biases in services output prices were mentioned as factors in the committee’s decision. Economist debates over statistical constructs and methods usually do not make the headlines, but in this case the Boskin report led to an unprecedented interest in price measurement, e.g., occurrences of “price mismeasurement” and “price measurement” in the Google Books Ngram text corpus jumped dramatically in 1997, with the latter having first shot up after the Chairman Greenspan’s January 1995 congressional testimony (figure 1).

To be clear, the arguments presented to the FOMC in September 1996 did not claim that the biases in official measures of inflation had increased. They were rather points of evidence suggesting that the behavior of aggregate data on profits and price and wage inflation was not a puzzle, and that in fact labor productivity was increasing. While identifying technology shocks as sources of business cycle fluctuations in formal modeling of macrodata is now more common, it is fair to say that identifying shocks to aggregate productivity in real time, i.e., in macroeconomic policy forecasting, remains a formidable challenge.

Figure 1. Frequency of Search for Price Measurement and Price Mismeasurement



Source: Google Books Ngram Viewer; case insensitive results from the American corpus.

That said, there seems to be some longevity to the “income-side” lessons learned from the 1990s productivity resurgence. Staff research later found that GDI better represents the business cycle than does GDP (Nalewaik 2010; see also Beaulieu and Bartelsman 2004), and with covariance between them, logically, a combination of the two is a more stable indicator of actual real output. BEA began to publish an average of GDP and GDI as gross domestic output (GDO) in 2015. BLS’s headline figure for quarterly labor productivity, which is for labor productivity in the nonfarm business sector, remains GDP-based, leading commentators to continue to examine signals from profits data and the statistical discrepancy.⁵

The notion that productivity may be—at long last—accelerating due to efficiency gains from IT-induced innovations also was in the news at the time of the September 1996 FOMC meeting. The notion had appeared in press economic commentaries as early as 1993 when influential works based on company-level data by Erik Brynjolfsson and Loren (Brynjolfsson and Hitt 1996) began circulating in academic circles. Discussions intensified more broadly when Business Week (Mandel 1994) proclaimed that “the productivity surge of the last two years ... may reflect the

⁵ For example, a Treasury analysis (Harris and Mehrotra 2022), linked the statistical discrepancy (at the time of their writing) to a strong likelihood that labor productivity growth since 2020 was stronger in the aftermath of the pandemic than portrayed by official reports.

efforts of U.S. companies to finally take full advantage of the huge sums they've spent purchasing information technology." The lack of direct evidence for this claim in the available macrodata added to a sense that the data may be missing a key aspect of innovative activity in economies.

Orientation of Price Measurement Research in the Division

As previously indicated, many R&S studies focused on advanced manufacturing products as part of an Industrial Output (IO) section initiative to improve the accuracy of the monthly IP index. The initiative began in 1998 as one piece of a wider effort led by Division Director Mike Prell to improve and update the division's statistics. The division's work on price measurement also engaged productivity researchers Steve Oliner and Dan Sichel, who were already parsing the contribution of IT capital to aggregate productivity (e.g., Oliner and Sichel 1994). In the economics profession at large, it is common for productivity researchers to pursue in-depth research on price measurement.

The Division's work proceeded with the aim of appropriately incorporating quality change in goods price measures—the “house-to-house combat” of price measurement as dubbed by Shapiro and Wilcox (1996). In view of the implications of price mismeasurement for monetary policy inflation targeting, and the fact that missed quality change due to technological change is thought to be the largest source of bias in official measures of price change, this focus was natural—whether the missed quality change was in investment prices or consumer prices.

In the event, most of the division's price measurement research has been directed at IT business output, which consists mainly of investment and intermediate goods that feed indirectly into the PCE inflation target of monetary policy. That said, there have been noteworthy efforts directed at accurately measuring consumer price change. The Shapiro and Wilcox study quoted above produced estimates of quality adjusted prices for cataract surgery, underscoring the need for better prices for health care services. (Division alums, Ana Aizcorbe and Louise Sheiner subsequently made contributions in the area.)

And following the body of work on IT goods prices described below, Dave Byrne and Carol Corrado estimated price indexes for consumer paid-for and “free” digital services (Byrne and Corrado 2021a, 2021b).⁶ Paid-for digital services are squarely within the scope of the Fed’s PCE inflation target, and their results implied an increasing bias in core PCE inflation since 2007. The methods they used are included in forthcoming international recommendations for measuring the digital economy in national accounts.

IT Sector Measurement: Data on “Ps” and “Qs”

The work conducted under the IO initiative began with the fast-growing domestic semiconductor industry, which accounted for more than 3 percent of manufacturing IP by 1996. Along with computers and communications equipment, the pace of technological change in semiconductors reflected the booming “tech” sector of the U.S. economy of the time.

The context of the planned work on IT in terms of data availability is important to understand. The possibility of acquiring datasets from private sources that contained very detailed, high frequency observations on both “Ps” and “Qs” was a novel opportunity. Prior to the second half of the 1990s, datasets of this type were not widely available for price research.⁷ R&S staff constructed price indexes for memory chips (DRAM), computer microprocessors (MPUs), and PCs using very detailed model-level unit price and revenue data purchased from private companies (Aizcorbe, Corrado, and Doms 2003; see also Aizcorbe 2006). The product-level price indexes for DRAM, MPUs and PCs were calculated using a matched-model formulation and, after a period of review, shepherded into monthly IP by Charlie Gilbert, whose routines to calculate chain-weighted indexes were made-to-order for crunching the complex, unbalanced panels typical of model-level datasets.

⁶ Paid-for services include fixed broadband internet services, cellular services, and video services, where the latter contained a component for streaming services from their inception in the over-the-top TV services market.

⁷ A line of related price research based on scanner data was emerging, however; it was the subject of a CRIW conference held in early 2000 (Feenstra and Shapiro, 2003). One objective of working with scanner data is to produce observations on homogeneous products in the form of unit prices and revenues, which is the type of data initially obtained by IO.

A matched-model price index will capture quality change as a composition effect, but based on past research it was thought that hedonic indexes were needed to capture quality change for high-tech products such as MPUs or PCs. The adequacy of the quality adjustment in matched-model indexes depends on whether the market for quality in new models or new versions of a product is in equilibrium. If equilibrium prevails, the introduction of a new model with a better price-quality trade off pushes down transaction prices of existing models to equalize quality adjusted prices. At the time, competition among so-called “IBM clones” in the PC market was fierce, AMD had become a serious competitor to Intel with the K6 processor, and the DRAM was a commodity product characterized by razor slim margins.⁸

The work on communications equipment began by studying a new, leading-edge product of the industry, local area network (LAN) equipment for which data were scarce. Mark Doms and Chris Foreman estimated hedonic indexes based on item list prices compiled from Cisco catalogs discovered in ARC’s offices (Doms and Foreman 2005; internal memo circa 1999). Doms subsequently developed quality-adjusted price indexes for the many types of equipment that powered fiber optic networks. The work was pathbreaking in the breadth of the products of the industry, new and old, that it covered (Doms 2005). But the ink was barely dry on this work when it became apparent that data sources for the growing wireless/cellular equipment segment needed to be developed. New data sources covering both wireless and wireline network products in substantial detail were eventually identified, making possible the development of quality-adjusted matched model price indexes for much of the industry (Byrne and Corrado 2015a, 2015b, initial working paper 2007).

⁸ As a practical matter, the appropriate method for estimating quality-adjusted prices for high-tech products is intrinsically related to data availability—and availability typically is scarce at industry inception, e.g., there may only be list prices of items produced by a leading company. Data availability tends to be more plentiful, and sometimes in the ideal form of model-level unit prices and revenues, as multiple products and brands are in competition. Whether the available data on a product class are sufficient to capture quality change via matched-model price indexes also depends on the frequency of new product introductions relative to the frequency of available data. For a discussion of these and related issues see Aizcorbe (2014).

The rapid adoption of smartphones after the introduction of the iPhone in late 2007 and the birth of Android—Google’s answer to the iPhone—in 2008 led to numerous iterations of smartphone hardware, which by then had become the platform of choice for new apps. A hedonic approach for measuring quality-adjusted cell phone prices was eventually developed (Aizcorbe, Byrne, and Sichel 2017) using quarterly data from IDC (purchased by BEA). The earlier work on MPUs also needed to adapt to industry changes and data availability, and price indexes driven by performance measures were introduced following a shift in Intel’s pricing strategy in the mid-2000s (Byrne, Oliner and Sichel 2018).

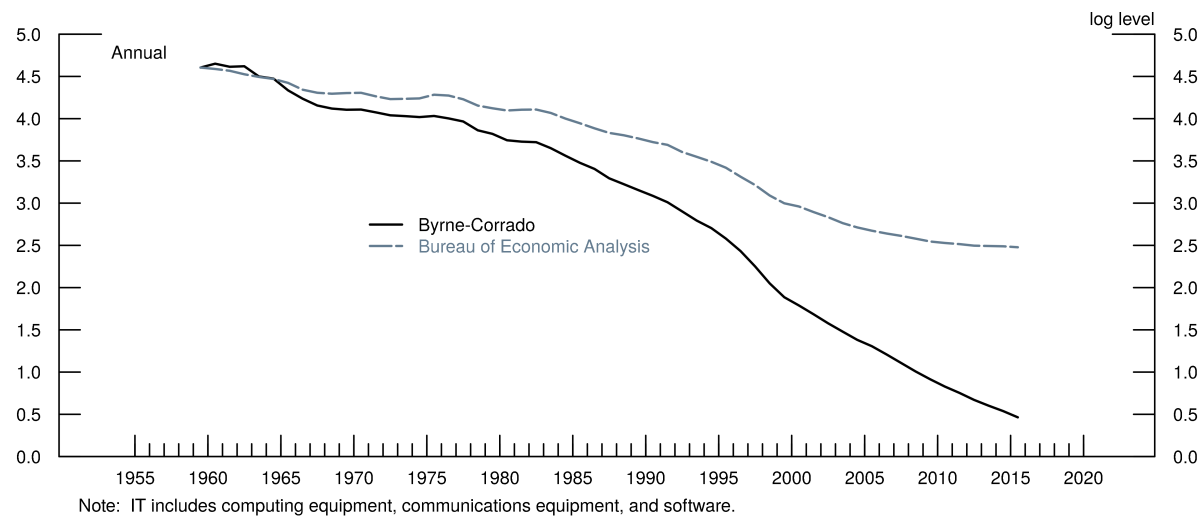
All told, the quality-adjusted IT price indexes developed by R&S staff have material impacts on real consumption and investment and their prices. Research price indexes compiled to reflect this cumulative effort are displayed in figures 2 and 3 below.

The details of these constructs are not reviewed here, but it is important to note that they are formulated in the same fashion as in the national accounts and draw upon research price indexes developed by academic researchers, e.g., Gordon (1990) and Berndt and Rappaport (2003), work at BEA (Grimm 1996, 1997), as well as Board staff. The implications of these indexes for understanding trends in technology, productivity, and “free goods” and consumer welfare are discussed extensively in the references cited in the figures; the IT equipment indexes were used in the Byrne, Fernald, and Reinsdorf (2016) to analyze the role of mismeasurement in the recent productivity slowdown. We will not review these studies (as much as we would love to wave the flag) but rather emphasize that “getting prices right” offers much to say about a range of economic issues as well as helping to assess underlying economic conditions (via IP and IT investment demand) and biases in the inflation rate target of monetary policy.

Though there is a gap between the research price constructs shown in the above figures and BEA’s IT price indexes, BEA incorporated R&S staff research price indexes for wireline network equipment, wireless transmission equipment, and cell phones in the national accounts; the review by Byrne and Corrado (2017a) also contributed to an update of BEA’s method for calculating a price index for purchased computer software. Several research price indexes

developed by R&S staff are not included, however, and the history for others was truncated, resulting in the gap shown in the figures.⁹ Also noteworthy, staff-produced estimates of depreciation rates of mainframe computers (Oliner 1993) and PCs (Doms, Dunn, Oliner, and Sichel 2004) are incorporated into BEA's measures of national income (via capital consumption).

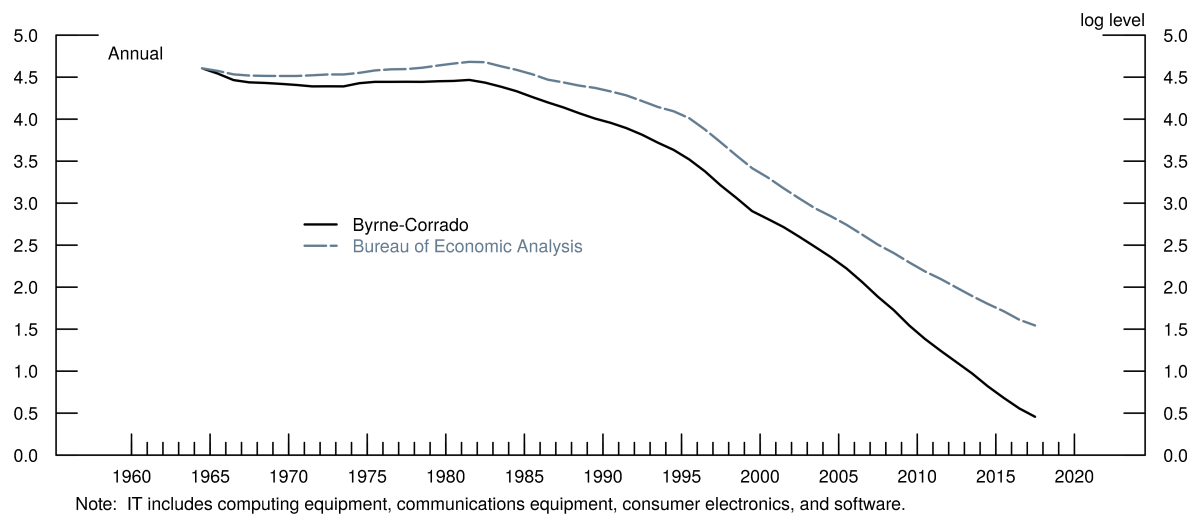
Figure 2. IT Investment Prices



Source: Price indexes reported in Byrne and Corrado (2017a, 2017b).

⁹ The exclusions and truncations include the Byrne (2015) price index for storage equipment; research on price indexes for PCs, including contributions by Berndt and Rappaport (2003) on PC prices that showed very fast drops in the 1990s; the PC prices developed using performance measures to capture quality change after the slowdown in microprocessor speeds in the early 2000s; the mainframe, server, and supercomputer price indexes Byrne and Corrado (2017a) developed for use beginning 1994, roughly the point at which BEA switched from their own computer price indexes to the BLS PPI; the history Byrne and Corrado (2015a, 2015b) developed for wireline communications equipment that included work by Gordon (1990), as well as their own work on satellite prices from industry inception in 1963 to 1995; and their history for wireless equipment, also from industry inception in 1986 to 2002, the point at which it was adopted by BEA.

Figure 3. IT Consumption Prices



Source: Price indexes reported Byrne and Corrado (2021).

The Cloud: “Ps”, “Qs” and “Ks”

The notion that IT services were becoming increasingly important for the macroeconomy was addressed by R&S staff early in the development of one of its key products, cloud computing services. Using posted prices scraped from the web, Byrne, Corrado, and Sichel (2021) estimated hedonic prices indexes for cloud services from their inception by Amazon in 2009.

The early grasp of the emerging cloud and of cloud technologies became relevant for analyzing and projecting IT investment. Cloud computing technologies enabled higher utilization of the installed computer server base, which helped to explain the weak trajectory of IT equipment investment in the aftermath of the global financial recession. In the event, the weakness also reflected an anomaly in BEA’s computer investment series that was highlighted by Byrne, Corrado, and Sichel (2017) and subsequently addressed in the national accounts.

For some time, IT investment had proxied as macroeconomic indicator of “tech” innovation, and the authors were well positioned to triangulate the apparent weakness in IT investment with the possibility that the national accounts may be missing the production/assembly of computers in the “server farms” of cloud services providers. By the time of their work on the impacts of

cloud technologies, owing to prior work on intangible investment that began more than two decades earlier, the authors had a simple, yet uncommon, routine in their back pocket for extracting and viewing time series of intermediate purchases from BEA's supply-use tables. This permitted them to spot what they reasoned was "missed" computer production by IT services industries in the national accounts.

Intangible Investment

R&S staff emerged as key players in developing the notion of the intangible economy in which a key portion of business investment is "missing" in national accounts. This refers to investments in intangible assets, e.g., the spending that creates a digital platform business model, efficient delivery systems, and global supply chains, that loom especially large in companies that drive global growth. The contribution of Carol Corrado and Dan Sichel, working with Chuck Hulten of the University of Maryland (Corrado, Sichel, and Hulten 2005, 2009; see also Corrado and Hulten 2010) was to develop a framework for understanding how an "expanded" investment construct can be used to analyze how intangible capital affects productivity and economic growth.⁴ Their approach is widely used, e.g., at the OECD and elsewhere, as well as incorporated in the latest edition of EUKLEMS productivity estimates, EUKLEMS & INTANProd. These uses and implications for productivity measurement are reviewed in a chapter in a forthcoming Brookings volume on productivity measurement edited by Louise Scheiner, with Marshall Reinsdorf (Corrado, 2023).

Before we say a few words on the relevance of the intangibles framework, consider the atmosphere the division at the time of its development. The growing IT sector and the gap between equity market and accounting valuations of firms that emerged during the 1990s led many economists and practitioners to believe that there was a "new economy." But what did that mean from a measurement point of view—was something important missing in our output and productivity data? To address this question, Corrado and Sichel, along with John Haltiwanger (also from the University of Maryland), organized an academic conference "Measuring Capital in the New Economy." The conference, which was sponsored by the NBER/CRIW, was held at the Board in April 2002.

Chairman Greenspan offered welcoming remarks to a packed Dining Room E that included many leading economic measurement and productivity scholars along with Federal Reserve and statistical agency staff. The new economy conference followed the dot-com crash that began in 2000 (and went to 2002) and thus lent itself to reflection on the longer-lasting measurement challenges that had emerged with new tools for accessing and sharing information via the World Wide Web. The initial research paper setting out how BEA could implement the capitalization of R&D in the national accounts was presented at the conference by Barbara Fraumeni (and 10 years later, that was done). All told five R&S staff were formal participants in the conference as authors or discussants. Besides Corrado and Sichel, Mark Doms presented the work described above, and Jason Cummins and Eric Bartelsman contributed; for more details, see the conference volume (Corrado, Haltiwanger, and Sichel 2005)

The original work on intangible capital focused on implications for productivity analysis, with intangible investment considered a broader indicator of the economy's investment in innovation than IT investment alone. Though this continues to be the dominant use of the framework, in recent years, analysis at the IMF, OECD and Bank of England has looked at the implications of the intangibles intensity of investment for the conduct of monetary policy. These largely cross-country studies are not reviewed here, except to note that findings range from a lowered interest sensitivity of aggregate demand to a flattening of the short-run aggregate supply curve, topics likely to garner further scrutiny in coming years.

3. Financial Accounts

The Federal Reserve first published its system of flow-of-funds accounts for the United States in 1955, reflecting a project it took over from the NBER in 1947. The system is the precursor to the quarterly *Financial Accounts of the United States* (FAs) statistics that are published today. The flow of funds system is based on the accounting scheme for “money flows” pioneered by Morris Copeland, who designed it to urge a better understanding of the circulation of funds between the financial and nonfinancial economy (Copeland 1947, 1952).

Previous R&S Division Director Ralph Young managed the transition of Copeland's project, soon to be joined by Dan Brill, who worked to develop the Federal Reserve's initial flow of funds system.¹⁰ Steve Taylor, who headed the Flow of Funds and Saving Section from 1961 until he retired in 1985, contributed significantly to the development of the Fed's system and was responsible for its early computerization and dissemination (e.g., Taylor 1958, 1963).

As stated in an IMF Handbook, the analytical power of the FOF system of accounts—like that of the income and product accounts—stems fundamentally from the interlocking character of the system—from the cross and down totals that balance for every period. Social accounting consistency requires that a flow change in any matrix cell be accompanied by corresponding changes in at least three other cells. (This is because asset demand functions are not independent in a fully specified multi-asset flow of funds model.) For example, if increased government capital formation is to be financed by government debt issues, some sector must absorb the issues. To do so that sector must have larger sources of funds or must reduce other acquisitions. By making use of this feature in various forms, it is often possible to trace the impact of each sector's financial behavior on the others and eventually on the nonfinancial economy, or vice versa.

Flow-of-funds (FOF) projections conditioned on a GDP projection (and, in turn, the interest rate assumptions that conditioned it as is done at the Board) exploit the “vice versa” side of the interlocking feature, i.e., the implications of a GDP path for financial flows. At one time, the implications of the flow of funds analysis of the “corporate financing gap” went strongly the other way, i.e., as a “check” on the staff judgmental forecast. The corporate financing gap is the difference between the corporate capital spending and projected internal cash flow, which reveals the amount of funds that must be raised in capital markets by the corporate sector to

¹⁰ Ralph Young came to the Board as assistant director of R&S in 1946 and was appointed director in 1949 (Obituary in the Washington Post, April 8, 1980). Before that he directed the financial research program of the NBER, which included Copeland's money flows project, and was on the faculty of Wharton School and chair of the economics department (then in the Wharton School) at the University of Pennsylvania. Dan Brill was Copeland's chief assistant at the NBER on the money flows project (De Bonis and Gibliobianco 2012) and came to the Board with the project in 1947. He subsequently served as R&S Division Director.

meet its capital spending plans (given government saving/borrowing, household saving, rest of world payments, etc.).

This “check” played out in GDP meetings in the 1980s. After presentation of the GDP projection, Division Director Jim Kichline would ask Steve Taylor whether the projected “real side” placed unusual demands on capital markets. More than once the investment analyst and GDP coordinator were sent back to their drawing boards because of Taylor’s analysis of the implied financing gap. With increased participation of nonfinancial businesses in financial activities, the proliferation of financial intermediaries, and globalization of corporate business, this check became a less reliable tool for domestic macro analysis and fell out of use.

FAs and Growth in Nonbank Intermediaries

In simple terms, paraphrasing Teplin (2001, page 431), the FAs measure financial flows across sectors of the economy, tracking funds as they move from those sectors that serve as sources of capital, through intermediaries (such as banks, mutual funds, and pension funds), to sectors that use the capital to acquire physical and financial assets. With data extending back to the 1950s, i.e., nearly 75 years, the accounts provide a consistent set of time-series data on financial flows in the economy. They reveal central trends, such as the growth of debt and the development of new financial instruments for providing credit. They document the growth of nonbank financial institutions as intermediaries and show how they became woven into the fabric of the economy.

By the late 1990s, nonbanks held two-thirds of the total credit market assets issued by banks and nonbanks (Fischer 2015). The leverage that had built up in a small but systemically important portion of nonbanks (the so-called “shadow banks”) precipitated the global financial crisis (and the failure of Long-Term Capital Management before that)—the point being that if one wishes to study these major events, the institutional context for them is found in the FAs.

Uses of the FAs in Macroeconomic Analysis

The FAs include balance sheets for the household and other major nonfinancial sectors and did so from their earliest days. Balance sheets explicitly account for capital gains and losses to maintain stock-flow consistency from period to period. This consistency provides a dimension of analytic capability beyond the many matrices of transactions by instrument and institution that form the core of the accounts. The ability to calculate the market value of household net worth is a prime example of this additional capability; the capability to study home equity extraction and its influence on consumer spending and the demand for credit are related examples. The gap between market valuations of corporate capital in relation to its replacement value enables the study of investment demand using a Tobin's q model, etc.

From the perspective of domestic macroeconomic analysis, the FAs provide data designed to help understand the behavior of households and corporations. A (highly unscientific) perusal of R&S staff research in this area over the last 50 or so years suggests that the FAs have been used most widely for the analysis of household behavior. Central themes in this body of work include the modeling and forecasting of consumer spending using household net worth and studies of the relationship between household debt and consumer spending.

A body of early work centered on estimating wealth effects based on the Ando-Modigliani consumption-wealth channel of monetary policy transmission. The channel derives from the aggregate implications of the life-cycle model of consumer spending in which wealth is a determinant of consumption and changes in wealth are induced by monetary policy. The approach was incorporated in the Board's quarterly econometric model from its earliest days (e.g., de Leeuw and Gramlich 1969) and remained the paradigm for explaining consumer spending in the inaugural version of FRB/US model in 1996 (Brayton, Mauskopf, Reifschneider, Tinsley, and Williams 1997). These models exploit the regularity in aggregate data between saving rates and wealth-to-income ratios predicted by a simple life-cycle model (e.g., see Dynan and Maki 2001, footnote 1 and figure 1, pages 1 and 40, respectively [here](#)).

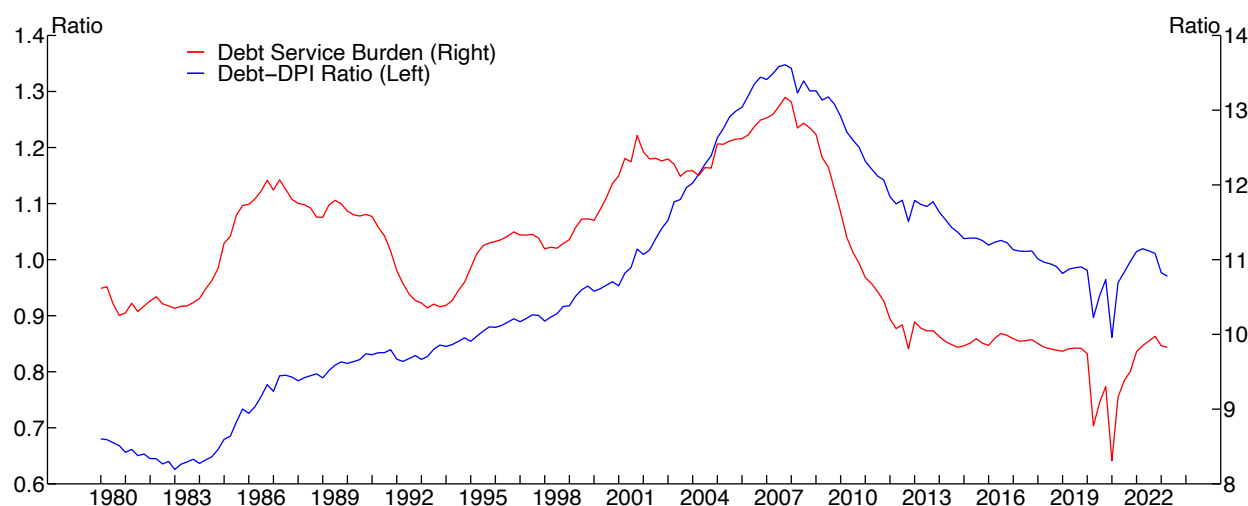
The question of whether and how stock market wealth "mattered" for consumption has been the subject of many empirical studies by R&S staff as reviewed in the paper on R&S macro research contributions in this conference. By the mid-1990s, many studies, inside the division

and beyond, were developing new findings on consumption behavior and stock market effects based on disaggregate and microdata, including the Survey of Consumer Finances whose development is discussed in the next section of this paper.

The currently favored paradigm for modeling consumer spending in FRB/US is that there are liquidity-constrained as well as nonliquidity-constrained households and different marginal effects of income according to income group. Limited borrowing opportunities of some households help to explain observed patterns of U.S. household wealth holdings as well as the fact that consumption tracks household income quite closely over the life cycle (e.g., Carroll and Summers 1991; see also Carroll 1997). These views seem consistent with the Federal Reserve's long-expressed sentiment in the Bulletin and FOMC minutes that elevated consumer debt may create financial distress should adverse employment and income conditions unfold. The utility of an income-level distinction when modeling consumption also helps explain why multiple investigations (in the division and elsewhere) that used aggregate time series failed to reveal a stable, direct link between higher levels of debt relative to income and changes in consumer spending.

Household sector debt outstanding marched steadily upward beginning in the early 1950s, though growth waxed and waned with the business cycle; figure 4 shows these data relative to income from 1980 on. The household debt-to-income ratio reached one by 2001, which was then unprecedented. The household debt service burden, measured as scheduled principal and interest payments on debt relative to income, also reached a new high that year. Subsequent years were of course marked by new highs and balance sheet turmoil associated with Great Recession.

Figure 4. Debt to DPI Ratio and Debt Service Burden



Note. Debt is household sector debt outstanding at end of period from the FAs. Disposable personal income (DPI) is after-tax income for the national accounts. Debt service burden is the scheduled principal and interest payment on household sector debt as a proportion of DPI as issued quarterly by the Federal Reserve.

The consumer credit data used to calculate the indicators shown in figure 4 are provided by the Board’s monthly Consumer Credit (G.19) release, a principle economic indicator.¹¹ The Federal Reserve’s monthly press release on consumer credit and quarterly statistics on household debt service ratios (mortgage and consumer) have probably garnered more ink in the press and Bulletin than any other type of financial data since their publication beginning in the 1950s. Even prior the ballooning of debt in the 2000s and associated household balance sheet stresses, a former Division director once remarked in the presence of one of the authors that the then manager of the consumer credit and debt service burden statistics, Charles Luckett, “was quoted in the press more than any figure in the Federal Reserve System.”

Enhanced Financial Accounts Initiative

¹¹ Consumer credit blends information from weekly reporting commercial banks, the Survey of Finance Companies, and multiple other sources, e.g., credit unions and lenders providing student loans. For further information see the “About” tab on the G.19 web page. For a broad review of developments in consumer credit see the book, *Consumer Credit and the American Economy* by ex-pats Tom Durkin and Greg Elliehausen and collaborators (Durkin, Elliehausen, Staten, and Zywicki 2014).

The Enhanced Financial Accounts Initiative was launched in 2014, described as “an ambitious and long-term effort to enhance the Financial Accounts of the United States by providing additional detail and disaggregation, higher-frequency data, and additional documentation and analysis of financial data, in order to improve our picture of financial intermediation and activity in the United States.” The project grew out of a need for more detailed information to investigate the causes and consequences of the financial crisis and to create indicators of financial stability. Three initial enhancements were released in 2015, each designed to provide additional detail on assets and liabilities of commercial banks, off-balance sheet items of banks, and newly compiled information on syndicated loans issued by banks and other financial institutions.

One of the most influential contributions of the EFA initiative has been issuance of the Distributional Financial Accounts (DFAs) that provide quarterly estimates of the distribution of household wealth from 1989Q3 to present. The roots of its construction predate the financial crisis, with early work on merging the flow of funds with statistics from the survey of consumer finances reported in internal memos from the division to Chairman Greenspan in 1990s. Carroll (2016) argued the need for distributional national statistics to understand aggregate household spending behavior at a CRIW conference on expanded measures held in 2012. The conference also included related papers addressing integration of flow of funds and SCF statistics, and of the flow of funds with national accounts, which presaged the current FAs. For further details, see Jorgenson, Landefeld, and Schreyer (2016).

The explosion of interest in the wealth distribution in the wake of influential work by Piketty (2013) and Saez and Zucman (2016) highlighted the limitations of the Fed’s publicly available data on U.S. household wealth that existed at the time, i.e., the relative low frequency of the SCF, lack of microeconomic heterogeneity in the flow of funds, and the fact that household sector wealth in the flow of funds included nonprofit institutions.¹² Saez and Zucman’s wealth estimates were based on an indirect method, capitalizing income from assets, which is very sensitive to rates and other assumptions used. Without going in details, R&S staff interacted

with these authors at CRIW summer workshops and elsewhere, eventually contributing to an improvement and revision of their estimates that better aligns with the Federal Reserve's data.

At the same time R&S's own work to further develop and enhance its estimates of the upper wealth tail was publicly discussed and reviewed (e.g., Bricker et al 2016, 2018). The ensuing release of the DFAs in 2019 appeared especially timely and responsive to the public's heightened interest in distribution statistics. This impressive contribution reflected a large team effort within the division; the key references at the time of introduction, Batty et al. (2019a, 2019b), a research paper and a FEDS Note, have 13 major authors.

4. Micro Data

Aside from the construction of the FAs and consumer credit indicators used to analyze the interrelationships between households and businesses and financial flows described above, the context much of the data developed for financial analysis in the division has revolved around the collection of micro data, i.e., observations on individual consumers or companies used to construct sample statistics. In addition, over the past approximately 40 years, there has been growing recognition that macro data alone are often not sufficient for understanding the evolution of the economy or its response to policy interventions. The example of the move to disaggregate income groups in FRB/US to understand the effects of debt and borrowing limitations, as discussed above, is a good illustration.

Statistical Surveys in R&S

The collection or assembly of financial micro data in R&S has a long history. In earlier times, before the separation of R&S and Monetary Affairs in 1987, most of the information collected consisted of money and credit statistics. Such information had particularly great relevance in times when policy focused more directly on management of the money supply and the money demand function played a significant role in macroeconomic analysis. R&S even contributed concepts and data for the analysis of monetary aggregates and their demand. William Barnett (1980) devised Divisia indexes as an alternative to the standard monetary aggregates based on an approach that weights monetary components by their contribution to liquidity. Paul Spindt

and colleagues (Avery, Elliehausen, Kennickell, and Spindt 1986) conducted surveys of the use of currency and transaction accounts, with the goal of being able to specify money demand equations at the micro level.

The oldest surviving micro data project in R&S is the quinquennial Census of Finance Companies, which is used to define the population of finance companies for the Survey of Finances Companies, which collects balance sheet information on firms identified from the Census. These two surveys provide a basis for a monthly panel survey of finance companies. Tabulations, the primary use of these data, began as an official published series in 1955 (G.19), and continue to be used in conjunction with bank-focused data to analyze the financing needs of households and businesses, especially small businesses. The surveys also provide information that is incorporated in the FAs.

The finance company survey provides a good example of how information collection must respond to changes in the environment for measurement. Unlike the case with banks, there is no official register of finance companies. Business lists, such as Dunn and Bradstreet, can be helpful in identifying potential candidates, but given the proliferation of finance companies, especially those nested inside larger nonfinancial corporate entities, there can be substantial omissions or questionable classifications. Moreover, the survey began suffering from nonresponse, i.e., companies were becoming less willing to provide the necessary information.

The level of difficulty and approximation required to maintain the finance company data reached a point by 2010 where a fundamental revision of the survey was needed. R&S staff developed novel statistical methods to both estimate the universe of such entities and account for the failure of entities to respond. The new methods also offered the capacity to provide a measure of uncertainty of the estimates, via standard errors for the resulting statistics, a first for any official series at the Board (see Chen *et al.*, 2013).

In its earliest days micro data outside of tabulated results did not play a significant role in the work of R&S. Following World War II, however, the division supported survey work at the University of Michigan to collect information on consumers' expectations, purchase intentions,

use of credit and other financial information. At the time, the economic psychologist George Katona was developing means of assessing economic behavior, including purchase intentions, through measurement of attitudes and other attributes of consumers. Improvements in card sorting (literally, for sorting punched cards according to information encoded on the cards¹³) made it possible to compile more sophisticated cross-tabulations of the survey's observations. The Michigan Survey of Consumers, which descends from that work, provides valuable statistics used in the division's analysis to this day.

By the early 1960s, computer technology had advanced enough that it was possible to consider collecting and starting to use microdata in some form other than tabulations and simple cross-tabulations. In this regard, the R&S division produced a literal breakthrough in the study of household finances and survey methodology with its 1962 Survey of Financial Characteristics of Consumers under the guidance of Dorothy Projector and Gertrude Weiss (1966). The survey questionnaire was developed in collaboration with researchers at the University of Michigan, and the data were collected by the Census Bureau. This survey was the first ever to collect detailed and complete household financial balance sheet information from a representative sample of individual households. It accomplished this by using IRS individual income data merged with 1960 Census data in its sample design. The approach allowed the survey to systematically sample the specific income groups and to over-sample the higher income groups to ensure there would be enough observations to support analysis of household investment behavior. Perhaps no other survey of consumer finances, before or after, has had the level of resources to support the accuracy of its results. For example, for corporate stock holdings the survey respondents were asked to report the company names and the number of shares owned. Staff in R&S looked up each stock in a specific issue of the Wall Street Journal and calculated and recorded the total value of the households' stock holdings. Without the benefit of today's model-based imputation methods, staff carefully imputed missing data items by exploiting available information and the technology of the day.

¹³ For the younger reader, see https://en.wikipedia.org/wiki/Punched_card_input/output

As recounted to one of the authors of this paper by Earling Thorensen, the last of the survey staff remaining at the Fed in the early 1980s, one of Dorothy Projector's greatest aspirations was to be able to compute the Gini coefficient for the U.S. wealth distribution, something that had only become possible to calculate with the completion of the survey data. She was given permission by the Board to have exclusive access to the mainframe computer overnight for this purpose. Reportedly, because machine memory was so limited at the time, all night the machine repeatedly disgorged punched paper tape encoding intermediate stages of the calculation and subsequently re-read the tape. By morning, the ultimate result was only a single number, but a landmark.

There was sufficient support to collect a second set of balance sheet data from the same households for the next year in the Survey of Changes in Household Finances (Projector, 1968), also a unique accomplishment at that time. There is even evidence in the small set of remaining project files that there was a subsequent effort to repeat the survey, and even some indication that data for one wave were collected and transferred to machine-readable form via a type of scanning that had been developed around then. The details of why that work, and the larger survey project, were terminated are not known, but there are some hints. One obvious critical factor was the serious constraint on data processing via computer. The large amount of time needed for dealing with data manually or developing workarounds to cope with computer limitations also meant that there was less time for the use of the data in ways that were relevant to the mission of the Board. One story was that Dorothy Projector was asked to make some seemingly simple calculation that would have had direct value for current policy work. Might she really have said, as claimed, that she had more important things to do to keep the survey on track? Whether true or not, efforts in this area subsequently slept for two decades.

In between, a circa 1976 request from the Senate Banking Committee for information about the use of credit sparked a partial revival of micro data collection on household finances in the division. The 1977 Consumer Credit Survey was produced under the leadership of Thomas Durkin (Durkin and Elliehausen, 1978) and transmitted to Congress. As its name implies, the

survey was primarily focused on credit-related information, though some limited other wealth and income information was collected.

Emergence of the Modern SCF

Around 1980, a confluence of factors heightened interest in collecting micro data on household finances. These factors included the enactment of the Depository Institutions Deregulation and Monetary Control Act of 1980 and the desire to analyze the rapid emergence of financial innovations in the form of new debt instruments for households. The desire for this type of analysis went beyond the Board and included government agencies such as the Department of Health and Human Services, the Department of Labor, the Federal Deposit Insurance Corporation, the Federal Trade Commission, the Office of the Comptroller of the Currency, and the Treasury. Owing to synergies both in needs and efficiencies, a comprehensive survey, the 1983 Survey of Consumer Finances (SCF) was developed in R&S to serve the needs of interested parties. The survey was under the leadership of Robert Avery with Gregory Elliehausen and in coordination with F. Thomas Juster at the Survey Research Center at the University of Michigan, which collected the data. One of the authors of this paper became involved later, in the technical processing of the collected microdata.

Like the 1962 survey, the 1983 SCF was anchored around the collection of household balance sheet information. But it went further in trying to capture additional details, including broader information on financial services, pension coverage, employment, much more contextual detail, and some information on attitudes and credit use which had a history in earlier work. One critical parallel between the two surveys was in the sample designs, though there were also important differences. Although it may not have seemed like it at the time, in the 1960s it was a less fraught matter to blend IRS and Census information to create a sample design that could oversample the wealthy in a meaningful way. For the 1983 survey, Fritz Scheuren, then director of the Statistics of Income Division (SOI) of the IRS, courageously made it possible for the survey to oversample wealthy households using section 6103n of the Internal Revenue Code. Even so, the permitted approach to oversampling high-income households had limitations. SOI selected a sample of wealthy households, independent of the selection of the

national random sample that constituted the bulk of the respondents. Unlike the case for the general sample, interviewers could not approach a person in the SOI sample unless they had explicitly agreed to participate in the survey, by returning a postcard stating that fact. The postcard was included in an envelope with a return address for the Office of the Comptroller of the Currency, a government agency not so universally known. Not surprisingly, only about 10 percent of the tax-based sample opened the envelope and agreed to participate, and of those not all agreed to participate fully enough for their data to be useful. Despite the level of precaution in insulating the SOI sample from any sense of being unwillingly approached, an editorial in the Washington Post about the threat to privacy very nearly put an end to the use of IRS data.

Although SOI provided a statistical weight for observations based on the tax data, they withheld information about the design of that sample or the nature of the weights, which would be needed for merging it with the separately selected broad random sample. It was only many years later that even partial information could be revealed. Consequently, a large amount of effort was needed to develop a credible way to merge the special and general samples for analytical purposes. In the end, there was necessarily an unknown level of approximation, which drove efforts to validate the data in other ways, such as through comparisons with data from what is now called the Financial Accounts of the U.S. (Avery *et al.*, 1988).

At first, policy interest at the Board in the survey focused most strongly on credit use by households. That interest supported a plan to undertake more regular collection of such data, preferably on a triennial basis. This led to planning for a 1986 survey, which was envisioned as a re-interview of the 1983 participants, to allow a calculation of saving and to support more sophisticated analysis for which panel data were needed. A critical difference between the situation then and in 1983 is that no outside funding was available, and the Board was unwilling to provide the level of support necessary to conduct the effort on the same scale. Every aspect of the survey was affected by the reduction in scope, but none as much as the necessary reduction in the survey content. Although there were important benefits from the survey, the limitations were obvious. As one R&S officer remarked at the time (paraphrasing) “One of the

most important lessons we have learned here is not to try to do this on the cheap.” Fortunately for the future, this reinterview survey proved itself especially notable for being able to provide information to gauge the distribution of effects on household finances of the “black Monday” market crash in 1987.

Yet some of the funding problems persisted with the 1989 survey. To bridge the difference between the cost and what the Board was willing to pay for the next triennial survey, one of the authors of this paper who was then the project director needed to visit other federal agencies to ask for money to support the data collection, with the nearly universal response being “You are from the Federal Reserve and you are asking ME for money?!” In the end, an important contribution came from a grant from the National Institute on Aging and a small, but crucially scale-tipping one from the Congressional Joint Committee on Taxation that allowed the project to proceed. Fortunately, the Board agreed to support subsequent surveys at full cost.

The 1989 survey marked an important methodological turn for the survey. For this reason, this survey is usually taken as the baseline when comparing SCF results across years. What would be more immediately obvious to a user would be the changes to the questionnaire. The structure and content were broadly reformed. Among other things, more questions were added to assess households’ view of their own preferences and their situation. There were a variety of significant changes on the statistical side, but the most important was the change in the sample design.

Because the desire for panel data remained strong, the 1989 survey incorporated a complicated sample design to allow both representation of the 1983 panel and a 1989 cross-section. For the panel observations, significant effort was devoted in the survey to understand financial in-flows and out-flows due to routine wealth rearrangement activities such as selling a home or because of changes in household structure since the 1983 baseline. Although this decomposition effort was unsuccessful, the panel overall was successful in terms of usability. Unfortunately, the complexity of managing data with both cross-sectional and panel representation collided with the limited human and technical resources. One key computer program needed to address problems of missing data required approximately six months of error-free run. Although

computer technology subsequently advanced to the point that such a long program could be run within a day, the human resources constraint remained. As a result, no further SCF panel data were collected until a special re-interview of the 2009 SCF participants in 2010 to help understanding the effects of the financial crisis on individual households.

The cross-sectional sample was the focus of much innovation. The requirement that the people selected from the tax data explicitly volunteer to participate was dropped and replaced by an option to decline. As expected, this change had a substantial positive effect on contact and participation. In addition, for the first time the survey was able to directly access SOI data, under strict security provisions, to use more sophisticated means to construct the special oversample and to use those data as well for post-survey adjustments. Because the SCF data had become an important input to modeling of tax policy in the Treasury and the Joint Committee on Taxation, this use of the information was allowable under section 6103(n) of the tax code (a provision for sharing data for purposes related to tax administration). This linkage, and its continuation in more sophisticated ways in the later surveys, is essential for the credibility of the SCF in capturing the upper tail of the wealth distribution and for the close correspondence of the survey estimates with those for households in the Financial Accounts of the United States. The survey was truly a pioneer in data linkage, which now plays an increasingly key role in empirical research in many fields, including economics.

Through subsequent waves of data collection, the SCF continued to be a source of sometimes path-breaking innovation in a variety of areas, including methodology, implementation, quality control, and privacy issues.¹⁴ This innovation was driven only to a moderate degree by abstract principles, but more strongly by the need to cope with the ever-evolving problems of conducting such a survey and making the resulting data fit for use. As Catherine Haggerty, the project manager of the SCF for many years with the collector of the data, NORC at the University of Chicago, once aptly remarked, “Of all the surveys we do, the SCF is the most

¹⁴ See Kennickell (2017) and other papers in the same issue of the *Statistical Journal of the IAOS*.

challenging.” No such project can operate for long on autopilot. A firm commitment to a model of evaluation and continuous improvement is important.

For example, gaining cooperation in surveys has grown more difficult over time in all surveys in the U.S., and the sensitive nature of the SCF subject matter may make it even more susceptible to such problems. The further cooperation deteriorates, the greater possibility there is for selectivity bias to become a problem. Although a variety of changes in the patterns or intensity of effort were developed to mitigate this problem, beyond some limits such approaches become more imperfect and may lead to escalating costs, a key practical concern.

Consequently, focus on potential selectivity biases and means of mitigating them have been a core research and development issue for the SCF. In the past, a large part of that effort was focused on exploiting the power of the SOI data for understanding the patterns of selectivity among the upper reaches of the wealth distribution and for mitigating them. Going forward, deterioration of cooperation more broadly, as is seen clearly in many surveys, may more strongly raise the need to address selectivity issues for other parts of the distribution as well, perhaps by expanding the use of SOI data or other sources that could help to align the observed population with the key economic and demographic dimensions of the actual population.

The SCF has a rich and broad use for policy and research within the Board. Most importantly, it allows for seeing how economic changes or policy play out among a wide variety of groups, not just for a “representative consumer.” It is, of course, also a natural place to try to understand inequality of wealth, across wealth groups, racial groups, and other groups.¹⁵ With time, the Board overcame its reluctance to highlight inequality issues and began including concentration ratios in articles published in the Bulletin about the SCF. Among important recent uses of the SCF within the Board is as an input to the creation the Enhanced Financial Accounts, which provides quarterly data of household finances distributed over wealth groups, as mentioned earlier. Because the survey has comparable data on wealth and income, as well as some information on households’ consumption, it has the potential to address the joint distribution

¹⁵ See Avery et al (1988) and Kennickell (2003).

of wealth, income and consumption, an area that has garnered increasing attention since at least the time of the Stiglitz-Sen-Fitoussi Report (Noll, 210). {Here is the complete citation: Noll, HH. The Stiglitz-Sen-Fitoussi-Report: Old Wine in New Skins? Views from a Social Indicators Perspective. Soc Indic Res 102, 111–116 (2011). <https://doi.org/10.1007/s11205-010-9738-9> } It has also supported a large body of academic research and influenced the creation of similar surveys elsewhere. For example, the Household Finance and Consumption survey of the European Central Bank is in many ways modeled on the SCF.

The SCF faces three informational limitations that historically have constrained its use. First, the survey is triennial, and as such it serves most directly only as a periodic benchmark of the structure of household finances and related matters. The second limitation is the length of the delay in the data release implied by the complexity of collecting and processing the data. The delay makes it a practical impossibility to think it could be used for strictly contemporaneous analysis even for the year of the survey. Modeling can be used to address these two issues, potentially by introducing a variety of external data, to project the relationships in the data to other periods. The use of SCF data in the distributional accounts, as mentioned earlier, is one such example. Some earlier discussions had explored the alternative possibility of using the SCF sample and data between data collection periods to design and execute rapid small-scale surveys focused on current topics of interest. Such data could be conditioned on the substantial body of data on households in the baseline survey.

Finally, there is typically no linkage of households across waves of the survey. Such information can be important in model building for addressing issues of causality or quantifying the degree of fluidity in the household economy that repeated cross-sectional observations obscure. As noted above, in the post-1989 period, the only follow-up panel of SCF data was collected in 2010, based on the 2009 SCF sample. The absence of SCF panel data has long been a complaint among academics. However, collecting such data is more complicated than collecting cross-sectional data and the post-processing of the data is substantially more difficult, owing to the cross-constraints implied by multiple observations on the same unit. Moving in that direction would require more staff and money, for a project that is already quite expensive.

Other Surveys

Another important development of micro data within R&S was the Survey of Small Business Finance (SSBF), a survey conducted four times between 1987 and 2003, with John Wolken as the project director (Mach and Woken, 2006). As is well known, small businesses (officially, those with fewer than 500 employees) are an important part of the dynamics of the job market and a source of innovation. But often lacking the history and financial weight of larger firms, financing for such businesses can face quite different obstacles. This survey was designed to cover both the important aspects of the businesses' balance sheets, their financial relationships, and other characteristics necessary for understanding the context in which the businesses operated and obtained credit and other financial services. This project, which collected a unique set of information, was regrettably canceled after the 2003 wave of the survey. Its influence lingers. A small amount of the information on small businesses can be obtained from the SCF, and since 2014, a Federal Reserve System effort has supported the annual Small Business Credit Survey, which continues some of the efforts of the SSBF.

The number of special-purpose efforts to collect survey data, including inclusion of questions on omnibus surveys, is so large and varied that it would be impractical to try to cover it all. Many efforts have been aimed at understanding the current economic conditions of households. Some have addressed specific questions in response to a request, sometimes from Congress. One unusual and amusing such request was for an estimate of the number of dollar coins in 2010 of distinct types (Eisenhower, Susan B. Anthony dollar, Sacagawea dollar, and the Presidential series, and the Native American series) that were held either in Federal Reserve vaults or private coin terminals on behalf of the Federal Reserve. Nearly one billion such coins were held across all 103 locations. Because it would have been impossible to enumerate all the coins, a sample survey seemed appropriate, if a suitable framework for efficient random sampling could be devised. Developing a stratified random sample of locations was straightforward. But the only common factors across locations were that the total number of coins in each location was known, all coins were held in bags of an approximately fixed content and an approximately fixed number those bags were held in individual containers. Otherwise, the physical arrangements across locations were different. Because the coins did not circulate

very often, there would be a tendency for bags of coins to cluster by date of issue. To ensure an appropriate and comparable dispersion of the sample across all locations, it was necessary to develop an abstract topology of containers and bags that could apply equally across locations, regardless of the positioning of the containers. For the curious reader, in the end it was estimated that 60 percent of the coins were from the Presidential series, with a standard error of 1.3.

Blending and Merging Data

R&S contains a substantial number of economists who work on financial institutions and markets, for the purposes of assessing market risks and financial stability, competitiveness of markets, and other aspects of the behavior of financial institutions. Although the Board, through the Monetary Affairs Division and the Division of Supervision and Regulation, collects data relevant to these tasks, such as the information in the National Information Center on the corporate structure of banks, in many cases the information is available at too low a frequency or too low a level of detail to be useful without additional data. A variety of sources of market data and data from vendors such as Bloomberg provide higher-frequency data, and R&S staff are adept at blending and managing these sources.

Using and blending multiple sources of data in a straightforward way in the financial context or more broadly requires clear and compatible definitions of variables across sources. In some cases, industry standards such as ones developed through the International Organization for Standardization (ISO) provide this service. Where such standardization is not present, substantial additional work is often required to develop appropriate approximations. A particularly compelling example in the context of financial analysis is the identity of a legal entity. Identity can be context dependent or multidimensional. For example, an entity might be seen as only an aspect of a larger corporate structure or simply as an object itself. A web of corporate relationships often can be exceedingly difficult to determine. In addition, many data sources use proprietary codes to identify such entities and charge a fee to use those codes. Mapping across data sources can be difficult or imperfect, often involving a degree of “hand” work. A classic example of the effect of these problems was the difficulty in identifying and

connecting entities in the corporate organization of Lehman Brothers in the beginning of the 2008 financial crisis (Bottega and Powell, 2011). In the aftermath of that crisis, the 2010 G20 meeting of finance ministers and central bank governors recognized the importance of improved financial data and called for the development of a global system for entity identification. Many national regulatory authorities and central banks came together, with support from the Financial Stability Board, to develop a suitable technical framework and a governance framework.¹⁶ R&S, together with other U.S. agencies, played a key role in the technical development and implementation of this frameworks. The resulting operational body, the Global LEI Foundation, went live in June of 2014 and continues to develop.

5. Concluding Remarks

Several themes emerge from our review of statistical programs in R&S and chief among them is the importance keeping these programs up to date. Surveys such as the SCF must continue to adjust to changes in the financial landscape, availability of relevant new data sources, new demands for data, and statistical issues such as maintaining meaningful response rates. Price measures will remain relevant as indicators of technology only as long as they remain up to date in terms of products covered and appropriate methods are used. Financial accounts must keep up with the proliferation of financial institutions and instruments and emerging analytical needs.

Another key theme is that financial microdata, new nonfinancial constructs, and miscellaneous economic indicators collected or assembled elsewhere increasingly became key resources in R&S. The division's march in this direction looks even to have accelerated in recent years with the Expanded Measurement Agenda launched to explore the utility of nontraditional sources in the wake of the COVID-19 crisis (Cajner et al 2022), In addition, a Federal Statistical Research Data Center (FSRDC) recently was established at the Board, easing access to a variety of government data sources that are available only in a secure setting.¹⁷ The passage of the

¹⁶ See www.lei.org/about/index.htm and www.gleif.org/en/about/history for details.

¹⁷ FSRDCs have made remote access to restricted-use government microdata possible for researchers working in locations ranging from Boston to Pittsburgh, to Chicago and elsewhere.

Foundations for Evidence-Based Policymaking Act (Public Law 115-435, often referred to as the “Evidence Act”) in 2018 explicitly calls for greater ability for agencies to share and link data for research aimed at policymaking, making it highly likely the division’s FSRDC facility will grow in importance. As we have seen, there is a long experience of linking or otherwise blending a broad swath of financial and nonfinancial data in R&S.

The division’s capability in sample design and technical processing of data is widely admired and has had significant positive spillover effects on the work of the Board as a whole. Support from R&S provides sample designs for ancillary surveys of financial institutions in Monetary Affairs. By the same token, other divisions, e.g., Supervision and Regulation as well as Monetary Affairs, generate a lot of data R&S uses for understanding financial conditions. Of course, understanding the resilience and potential risks financial institutions may pose for the broader economy is priority for many at the Board, but we believe it is fair to say that R&S provides the major tools and datasets that sustain and advance this effort.

The body of R&S research on real IT output, productivity, and investment continues to provide insights for policymakers, researchers, and practitioners seeking to assess and improve economic performance. Working to construct quality-adjusted price indexes for high-tech products and industries builds near geek-level command of technologies and, less obviously, of industry production geography. This knowledge has proved relevant to analysis of the Chips Act and related industrial initiatives of the current administration. The intangibles framework helps to provide perspective how the current wave of AI technology may affect the economy, and while not discussed in this paper, how the pace and composition of intangible investment influences total factor productivity through spillovers.

Finally, besides the analysis of financial conditions per se, the data from the FAs, consumer credit, and the SCF have provided the basis for many studies of consumer behavior and assessments of data on consumers, areas where R&S staff have been especially prolific. Topics investigated include how the level, dynamics, and distribution of wealth and debt, including the cost of debt and capital gains and expectations of them, affect consumer behavior—and the adequacy of data sources for understanding these matters. Grasping these issues are central to

understanding how financial conditions affect consumer spending and to formulating theories and testing them based on disaggregated or microdata.

All told, we believe it is fair to say that the Board would be in a state of near know-nothingness about current economic conditions, financial stability, the performance of the economy (especially the role of technology in driving it), and many fundamental factors affecting consumer spending and business capital formation without the economic measurement programs in R&S.