MIT Open Access Articles

Measuring Welfare with Massive Online Choice Experiments: A Brief Introduction

NEW MEASURES OF THE ECONOMY[‡]

Measuring Welfare with Massive Online Choice Experiments: A Brief Introduction[†]

By Erik Brynjolfsson, Felix Eggers, and Avinash Gannamaneni*

One of the fundamental objectives in economics is to assess the well-being of people. Economists, policymakers, and journalists routinely use changes in GDP (Gross Domestic Product), and metrics derived from it, such as productivity (GDP divided by hours worked or other measures of input), as a proxy for changes in well-being. For example, a report from the Federal Reserve Bank of Boston referenced "…a measure for standard of living: average real gross domestic product (GDP) per capita."[1] However, GDP was never meant to be a measure of welfare. It is a measure of production. Even Simon Kuznets, the creator of GDP, in a report delivered to US Congress in 1934, warned against using it for other purposes: "The welfare of a nation can scarcely be inferred from a measurement of national income as defined [by the GDP]."

In some cases, GDP and welfare are correlated, but in many other situations this need not be the case, and even the signs of the changes in GDP and welfare can go in opposite directions. Alternative measures going beyond GDP and incorporating estimates of subjective well-being have been proposed, but a majority of economists do not consider them reliable enough for macroeconomic analysis and policymaking (den Haan et al. 2017).

In theory, consumer surplus is a better measure of consumer welfare. However, in practice, it is challenging to measure consumer surplus in a scalable manner since such measurement requires full estimation of demand curves. In Brynjolfsson, Eggers, and Gannamaneni (2017)—henceforth, BEG—we propose a new way of directly measuring consumer welfare using massive online choice experiments while staying within the neoclassical framework. Our approach takes advantage of the fact that in recent years it has become much easier to collect data online on a large scale. These advances have been essential for creating alternative measures of the economy including ours. In this article, we provide a brief introduction to our approach.

I. GDP versus Consumer Welfare

To illustrate the relationship between GDP and welfare, we consider three cases of goods that we term *classic goods*, *digital goods*, and *transition goods*. Classic goods, such as automobiles, food, haircuts, etc., are sold with a positive price. For these goods, GDP and welfare are usually positively correlated. Most of the goods of the twentieth century would fall in this category. Holding price constant, as we double the production and consumption of such goods, both GDP and welfare roughly double as well.

[‡]*Discussants:* Hal Varian, Google and University of California-Berkeley; Betsey Stevenson, University of Michigan; Charles Hulten, University of Maryland-College Park; Dennis Fixler, US Bureau of Economic Analysis.

*Brynjolfsson: MIT Sloan School of Management, 100 Main Street, Cambridge, MA 02142, and NBER (email: erikb@mit.edu); Eggers: University of Groningen, Faculty of Economics and Business, Nettelbosje 2, 9747 AE Groningen, The Netherlands (email: f.eggers@rug.nl); Gannamaneni: MIT Sloan School of Management, 100 Main Street, Cambridge, MA 02142 (email: avinashg@mit.edu). We thank the MIT Initiative on the Digital Economy for financial support, Carol Corrado for discussing the paper at the AEA annual meeting and participants at the IMF Statistical Forum, the NBER CRIW Workshop, and the Society for Economic Measurement for valuable comments. We are, of course, responsible for any remaining errors.

[†]Go to https://doi.org/10.1257/pandp.20181035 to visit the article page for additional materials and author disclosure statement(s).

[1] https://www.bostonfed.org/-/media/Documents/ledger/ledger2003/measure.pdf (accessed January 2018).

In theory, it is possible to perfectly infer welfare from national accounts for classic goods, provided that we make perfect quality adjustments in every period. However, in practice, it is infeasible for government agencies with limited budgets to keep track of thousands of quality changes happening every day, and so welfare calculations for classic goods will be imperfect, even if they are at least, generally, directionally accurate.

For other types of goods (digital and transition goods), GDP and welfare might even move in opposite directions. Many of the goods we consume in the twenty-first century are purely digital goods such as online encyclopedias, search engines, social media, digital maps, etc. Most of these digital goods are available at zero price to consumers and hence excluded from GDP.[2] As we increase the production and consumption of such goods, GDP does not change, however, welfare does increase.

Increasingly more types of goods are transitioning from classic to free digital goods. While previously these types of goods were counted in GDP measures, they are excluded from GDP once they transition to free digital goods. The encyclopedia industry offers an excellent illustration of transition goods. Previously, people bought and paid for physical copies of encyclopedias such as *Encyclopedia Britannica*, and these transactions contributed to GDP. However, over the past 15 years, Wikipedia has replaced *Encyclopedia Britannica* as the premier reference source. Wikipedia is a truly free good without any advertising revenue. Contributors donate their content and editing services, and voluntary donations from readers pay for the relatively small operating expenses like computer services. Because it has zero price, Wikipedia is excluded from GDP measures. As a result, the contribution of encyclopedias to GDP decreased because people shifted from paying for *Encyclopedia Britannica* to consuming Wikipedia for free. However, consumers are clearly better off because they now have access to a much larger quantity of encyclopedic

reference for free.[3] Therefore, GDP and welfare can move in opposite directions for such transition goods.

II. Measuring Welfare Directly Using Massive Online Choice Experiments

In BEG, we explore discrete choice experiments as a way of estimating demand curves of goods with or without a price and inferring consumer surplus generated by these goods. We are comfortable focusing on consumer surplus and not producer surplus because Nordhaus (2005) finds that from 1948–2001 consumers captured over 96 percent of the social returns from innovations.

In BEG, we conduct several variations of such experiments, with or without incentive compatibility, including single binary discrete choice (SBDC) experiments, best-worst scaling, and Becker-DeGroot-Marschak (BDM) lotteries. We show that each of these approaches generate plausible demand curves, and valuations are consistent across all of these approaches. We also provide results from incentive-compatible studies where subjects have to make choices involving real consequences, i.e., receiving a monetary reward for giving up the good. This makes it possible to compare the results of experiments which pose purely hypothetical choices to those which are incentive compatible.

SBDC experiments are the most scalable and in our implementation, they involve short single question surveys that we run on Google Surveys.[4] We ask respondents to make a simple hypothetical choice between keeping access to a particular good, or giving it up for a certain amount of time in exchange for a monetary reward. Several thousands of such choice experiments can be run every day on a massive scale. In BEG, we illustrate this approach by surveying over 100,000 consumers to estimate the consumer surplus generated by the most widely used digital goods. This approach can be expanded massively to estimate the consumer surplus generated by digital and non-digital goods alike.

---

[2] In some cases, they are advertiser-supported, and because advertising is an intermediate good, it is excluded from GDP. What's more, advertising revenues and consumer welfare need not be correlated (Spence and Owen 1977).

[3] The Current English Wikipedia has as much content as 2,580 volumes of Encyclopedia Britannica (https://en.wikipedia.org/wiki/Wikipedia:Size_in_volumes).

[4] https://www.google.com/analytics/surveys.

To briefly introduce our approach, here we provide an example of estimating consumer surplus generated by Wikipedia. We conducted SBDC experiments similar to the experiments from BEG for giving up access to Wikipedia for one year. More specifically, we asked the following question: *Would you prefer to keep access to Wikipedia or go without access to Wikipedia for 1 year and get paid $E?*

• Keep access to Wikipedia
• Give up Wikipedia and get paid $E

We chose the following values for E: $1, $10, $20, $30, $40, $50, $60, $70, $80, $90, $100, $1,000, $2,000, $5,000, $10,000, $20,000, $50,000. For each E, we asked around 500 respondents, adding up to a total of 8,514 respondents in the United States. Note that this approach yields a measure of willingness-to-accept rather than willingness-to-pay. Figure 1 displays the estimated relationship for Wikipedia. As we increase E, more and more people choose the cash option and give up Wikipedia, yielding a downward sloping curve. According to our estimates, the median person living in the United States demands $150 to give up Wikipedia for 1 year (95 percent confidence interval = [$124, $182]). This translates to around $50 billion in consumer surplus per year created by Wikipedia in the United States alone, a value that cannot be inferred from GDP data.

In BEG we apply SBDC experiments to the most widely used categories of digital goods and find that search engines, email, and digital maps are the most valued categories, followed by video, E-commerce, social media, messaging, and music services. We also explore this approach for non-digital goods by computing the consumer surplus generated by breakfast cereal. The results show that the consumer surplus for breakfast cereal is not very different from revenues for breakfast cereals. For such goods, it might not be problematic to infer changes in welfare from changes in GDP.

Using a best-worst scaling approach, we also obtain an ordinal ranking of median valuations of numerous digital and non-digital goods, including toilets at home, meeting friends in person, and airline travel. We find that access to toilets at home is the most highly valued of the goods we consider, but internet access comes a close second. The median value for digital goods such as search engines and email is greater than the median value for than meeting friends
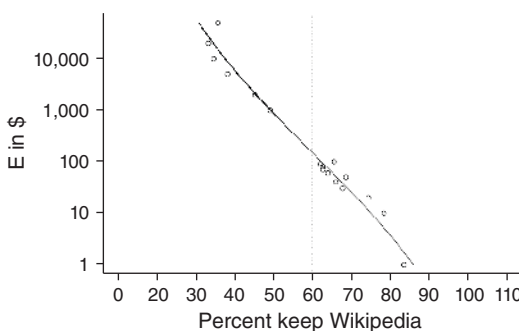

FIGURE 1. WILLINGNESS TO ACCEPT (WTA) FOR LOSING ACCESS TO WIKIPEDIA FOR ONE YEAR

in person. The median consumer also values Facebook more than airline travel. We also conduct an incentive compatible study of Facebook in 2017 and find that the median consumer demands around $38 to give up Facebook for 1 month.

In Brynjolfsson et al. (2018), we integrate our welfare measures obtained through such massive online choice experiments back into the GDP framework. We derive adjustment terms to national accounts to account for new goods, new free goods, and continuing free goods so that welfare can be inferred for GDP, and we calibrate these terms for several free digital goods. These adjustment terms are based on consumers' willingness to accept values to give up access to these digital goods and are obtained through our online choice experiments approach.

III. Discussion

Our choice experiments approach suffers from limitations related to contingent valuation studies as outlined in Hausman (2012). In BEG, we tackle these limitations including the key issue of hypothetical bias. We conduct a large set of hypothetical and incentive compatible studies to study the bias present in hypothetical choice experiments. We find that, in contrast to much of the literature on contingent valuation, in our setting hypothetical bias typically leads to *underestimating* the consumer surplus valuations. In Brynjolfsson, Eggers, and Gannamaneni (2018) we dig deeper into theories of why hypothetical bias exists. In order to identify drivers of the hypothetical bias we systematically vary factors

that we identified from the literature and qualitative pretest, including loss aversion, System 1/System 2 judgments, strategic responses, and option value.

While GDP is measured quite precisely, our approach provides less reliable estimates of welfare, but precision can be improved with increased sample sizes. Moreover, we are of the opinion that it is better to be imprecisely correct than precisely wrong while measuring well-being.

Our approach can be scaled easily to hundreds of thousands of goods by running several thousand choice experiments every day. This approach can be implemented more frequently than the CPI and can be used to track changes in well-being over time. Moreover, goods can be easily added or removed from the basket.

That system of national accounts centered on GDP was one of the greatest inventions of the twentieth century. In the twenty-first century, the proliferation of digital data and an infrastructure of easily, cheaply and quickly surveying millions of people provide an opportunity to develop new measures of welfare which can be used to supplement GDP. In our research agenda, we propose one such method of directly measuring consumer welfare by conducting massive online choice experiments.

REFERENCES

Brynjolfsson, Erik, Erwin Diewert, Kevin Fox, Felix Eggers and Avinash Gannamaneni. 2018a. "The Digital Economy, GDP and Consumer Welfare: Theory and Evidence." Unpublished.

Brynjolfsson, Erik, Felix Eggers and Avinash Gannamaneni. 2017. "Using Massive Online Choice Experiments to Measure Changes in Well-being." Unpublished.

Brynjolfsson, Erik, Felix Eggers and Avinash Gannamaneni. 2018b. "Hypothetical Bias in Choice Experiments." Unpublished.

den Haan, Wouter, Martin Ellison, Ethan Ilzetzki, Michael McMahon, and Ricardo Reis. 2017. "Happiness and Wellbeing as Objectives of Macroeconomic Policy: Views of Economists." *VOX EU*. March 30. http://voxeu.org/article/views-happiness-and-wellbeing-objectives-macroeconomic-policy.

Hausman, Jerry. 2012. "Contingent Valuation: From Dubious to Hopeless." *Journal of Economic Perspectives* 26 (4): 43–56.

Kuznets, Simon. 1934. "National Income." 73rd Congress, 2nd Session, Senate Document 124.

Nordhaus, William D. 2005. "Schumpeterian Profits and the Alchemist Fallacy." Yale Economic Applications and Policy Discussion Paper 6.

Spence, Michael, and Bruce Owen. 1977. "Television Programming, Monopolistic Competition, and Welfare." *Quarterly Journal of Economics* 91 (1): 103–26.