Well-Being and Public Policy

John Bronsteen
John.Bronsteen@chicagounbound.edu

Christopher Buccafusco
Christopher.Buccafusco@chicagounbound.edu

Jonathan Masur
dangelolawlib+jonathanmasur@gmail.com

Follow this and additional works at: https://chicagounbound.uchicago.edu/law_and_economics

Part of the Law Commons

Recommended Citation

This Working Paper is brought to you for free and open access by the Coase-Sandor Institute for Law and Economics at Chicago Unbound. It has been accepted for inclusion in Coase-Sandor Working Paper Series in Law and Economics by an authorized administrator of Chicago Unbound. For more information, please contact unbound@law.uchicago.edu.
WELL-BEING AND PUBLIC POLICY

JOHN BRONSTEEN
CHRISTOPHER J. BUCCAFUSCO
JONATHAN S. MASUR
WELL-BEING AND PUBLIC POLICY

John Bronsteen, Christopher Buccafusco and Jonathan S. Masur

THE LAW SCHOOL
THE UNIVERSITY OF CHICAGO

November 2014

Well-Being and Public Policy

John Bronsteen, Christopher Buccafusco & Jonathan S. Masur

Introduction

Governments rely on certain basic metrics and tools to analyze prospective laws and policies and to monitor how well their countries are doing. In the United States, cost-benefit analysis (CBA) is the primary tool for analyzing prospective policies, especially with respect to administrative regulations. Similarly, Gross Domestic Product (GDP) is perhaps the most prominent metric for monitoring a country’s progress. Both CBA and GDP are economic measures: CBA values outcomes by trying to approximate how they would be valued by the market, and GDP focuses on total economic output.

For decades, critics of such economic measures have argued that they ignore important aspects of value that are not fully reflected by output or by willingness to pay (WTP). This first manifested itself in the social indicator movement of the 1960s and 1970s (Stiglitz, Sen, and Fitoussi, 2009:208 & n.86; Veenhoven, 1996:2) that aimed to produce rival measures based on objective values such as health and education. The most famous such measure is the Human Development Index (HDI).

In recent years, one of the most important developments in social science has been the emergence of psychological research measuring subjective well-being (SWB) or “happiness.” Researchers have made great strides in replicating and validating their findings about happiness, and world leaders have called for this research to be used to supply SWB-based metrics and tools as alternatives to the existing economic ones (Cohen, 2011; Samuel, 2009; General Assembly Resolution 2011).

In response to these calls, early efforts have been made to use SWB research to create new social indicators. The efforts have been sensitive to a number of important obstacles. For example, subjective well-being cannot be validated via objective proof in the same way that economic output or longevity can be. In addition, quality of life may have components other than SWB, and SWB itself may have different components that cannot be easily combined.

But despite these concerns, scholars and policymakers understand the political power of a simple, unitary metric such as GDP or the HDI. And they also see the value of SWB measures in assessing the quality of human life more directly and perhaps more accurately than can be done via traditional economic measures, and more comprehensively than via other social indicators. As a result, they have shown intense interest in using SWB data to create alternatives to the existing tools and measurement devices.

In this chapter, we discuss some of the efforts that have been made in this regard.¹ We first briefly explain the way that SWB is measured and the way those measurements have been validated. We then explain our own contribution — well-being analysis (WBA) — which uses

¹ This chapter borrows directly in some instances from other recent work of ours, especially Bronsteen, Buccafusco & Masur 2013 and Bronsteen, Buccafusco & Masur 2015.
happiness data to analyze prospective policies more accurately than does CBA. Next, we cover the ways in which SWB data have been used to generate prices that can be used by traditional economic analysis. We then discuss attempts to revise cost-benefit analysis to deal with the limitations stemming from the fact that it uses wealth to assess the effects of policy on quality of life. Finally, we lay out the strides that have been made toward creating an SWB-based alternative to GDP.

**SWB Data**

Since the days of Jeremy Bentham, economics has focused on utility. Although Bentham and his followers believed that utility could plausibly be measured directly, most modern economists have sought proxies for utility (Colander 2007). For the most part, modern economics has assumed that the choices people make are the best available proxy for their well-being. Daniel Kahneman has called this “decision utility.” More recently, however, a new wave of social scientists has called for a return to Bentham and, accordingly, the measurement and analysis of “experienced utility” (Kahneman et al. 1997). These social scientists have attempted to develop more direct and valid proxies for individuals’ well-being that rely on self-reported happiness.

Over the past quarter century, researchers in the field of hedonic psychology have developed a variety of new tools for studying SWB. We will describe some of these tools, discuss issues about their validity and reliability, and briefly summarize a few of the key findings from the literature.

Hedonic psychologists believe that the best way to figure out how an experience makes a person feel is to ask her about it while she is experiencing it. The “gold standard” of such measures is the experience sampling method (ESM), which uses handheld computers or mobile phones to survey people about their experiences (Kahneman et al. 1997). Subjects are beeped randomly throughout the day and asked to record what they are doing and how they feel about it. The data that emerge from such studies provide a detailed picture of how people spend their time and how their experiences affect them. The data can also be combined with socio-economic and demographic data via regression analyses for even greater insight (e.g., do the unemployed spend more time in leisure activities than the employed and do they enjoy these activities as much?).

The oldest method of measuring SWB is the life satisfaction survey. These surveys ask individuals to respond to a question such as, “All things considered, how satisfied with your life are you these days?” (Pavot & Diener 1993). Respondents answer on a scale that ranges from “not very happy” to “very happy.” Life satisfaction surveys have been included in the U.S. General Social Survey since the 1970s; as a result, we now have substantial quantities of longitudinal data on thousands of individuals. The principal value in such surveys is the ability to correlate SWB data with a variety of other facts about people’s lives. Using multivariate regression analyses that control for different circumstances, researchers are able to estimate the strength of the correlations between SWB and factors such as income, divorce, unemployment, disability, and the death of family members (Lucas et al. 2003; Clark et al. 2008).
ESM and life satisfaction surveys have different advantages and disadvantages. ESM relies less on difficult cognitive processes like remembering and aggregating experiences, so it avoids certain kinds of errors. But while ESM can give a fine-grained picture of people’s lives, it tends to be very expensive to run for large samples.\(^2\) Life satisfaction surveys are relatively inexpensive to administer and can be easily included in a variety of larger survey instruments. Accordingly, they are valuable as sources of large-scale data about many subjects and of longitudinal data about changes in SWB over time.

The various SWB metrics have generally scored well in terms of their reliability and validity, often performing at least as well as more established economic measures of utility (Bronsteen, Buccafusco & Masur 2015). Meta-analyses of different well-being tools have found high levels of reliability for both life satisfaction and experience sampling methods. This is especially true of more advanced multi-item measures (Diener et al. 2009). In addition, SWB measures tend to correlate well with a variety of other relevant indicators, including other SWB data, third-party well-being estimates, the amount of time someone spends smiling, and neurological activity (Diener et al. 2009; Bronsteen, Buccafusco & Masur forthcoming). Finally, SWB metrics seem appropriately sensitive to external factors in peoples’ lives, including unemployment, health, social relationships, and income.

Despite these findings, there are still reasons to be cautious about some uses of SWB data. The major concern about hedonic data is the one that first dissuaded economists from attempting to measure utility directly—anxieties about interpersonal cardinality. How can we know, for example, that when two people both rate their SWB as “5 out of 10” they mean the same thing? While this is an important objection, it is one to which there are answers. First, if different people use the happiness scales differently, this will not necessarily lead to systematic errors if SWB data is used to formulate policy. Instead, it may simply produce noise. Different scale usage will only be a problem if it correlates with differences between two populations that are being compared.\(^3\) Moreover, the use of monetization as a proxy for utility suffers from the same interpersonal cardinality problem. Whereas a reported “5.0” might reflect a different level of well-being for one person than it does for another, there are definitely situations in which a particular sum of money will affect one person’s well-being differently from another’s. One thousand dollars might mean everything to a homeless person and nothing to Bill Gates. Because of the diminishing marginal value of money, two individuals with differing levels of personal wealth can obtain vastly different amounts of welfare from the same gain (or loss) of income (Frank 1997). And even two people with identical wealth could derive radically different amounts of utility from the same amount of additional money.

In assessing the value of these data, the question should not be whether they perfectly measure utility. SWB data are not and cannot be a perfect proxy for human welfare. The important question is how they compare with the available alternatives. As we explain in more

\(^2\) Some of these difficulties can be minimized with the day reconstruction method (DRM) pioneered by Daniel Kahneman and his colleagues (Kahneman et al. 2004). DRM uses daily diary entries about each day’s experiences to reconstruct an account of subjects’ emotional lives. DRM studies correlate strongly with ESM studies and can be run at lower cost.

\(^3\) In addition, the U-Index proposed by Krueger et al. is designed to mitigate differences in scale usage (Krueger et al. 2009).
Well-Being Analysis

Virtually every law makes people’s lives better in some ways but worse in others. For example, an air-quality law could make people healthier, but it could also force them to pay more money for the products they buy. Every proposed law thus raises the question: Would its benefits outweigh its costs?

To answer that question, there needs to be a way of comparing seemingly incommensurable things like health and consumer consumption. The most widely adopted method, which has now become standard within the American regulatory state, is cost-benefit analysis (CBA). CBA involves “monetizing” costs and benefits—translating real-world effects into their equivalent dollar values—and then comparing them in monetary terms. Thus, conducting a CBA of an air quality law would involve estimating the monetary costs of the law to consumers; estimating the monetary value they would place on cleaner air; and then comparing the two quantities. Every economically significant regulation from executive-branch agencies must, by law, be evaluated via CBA or an analogue. This has been the case since 1981, when President Reagan mandated it by executive order. That order has been reaffirmed by every president since, including Presidents Clinton and Obama (Bronsteen, Buccafusco & Masur 2013).

Yet this monetization process is hardly straightforward. The benefits produced by laws and regulations are often nonmarket goods such as health and safety, and thus there are no direct means for determining what value individuals place on them. CBA must thus turn to indirect mechanisms: either researchers must conduct contingent valuation surveys in which they ask, hypothetically, how much an individual would pay for a particular benefit, or they must attempt to determine revealed valuations by studying individuals’ market decisions such as the wage premium that workers demand to take a more dangerous job.

If an agency is considering regulating some aspect of workplace safety, for example, it will have to compare the costs of that regulation (including implementation costs, increased prices, and unemployment) with the benefit of improved employee health. The data for these estimates are derived from revealed preference studies that demonstrate the implicit value that workers attach to their lives by comparing the salaries that workers are paid in workplaces with different levels of risk (Hammitt 2007). Alternatively, if an agency is calculating the benefits of environmental regulation, it will typically attempt to estimate the value of, for example, an increase in protected wetlands by asking people to indicate how much they are willing to pay (WTP) for the benefit. In these stated preference or contingent valuation studies, a representative population will be surveyed and asked hypothetical questions about their WTP for various benefits (Desvousges et al. 1993). Revealed and stated preference studies provide a method for converting non-market benefits into monetizable and comparable units.

---

4 Exec. Order No. 12,866, 3 C.F.R. 638.
Both of these methods are flawed in a variety of ways, and numerous studies have demonstrated their shortcomings. The validity and reliability of the data produced by revealed and stated preference studies are questionable. These studies place considerable cognitive demands on subjects that undermine the trustworthiness of the data they generate. In the workplace safety example, employees are assumed to understand the level of risk that different jobs pose, and they are assumed to make rational tradeoffs between increased wages and potential health risks. A wealth of psychological research suggests that people fare very poorly at these sorts of tasks. People’s minds are not designed to differentiate between exceedingly small risks, and when asked to do so rationally, they frequently fail (Sunstein 2002).

Most importantly, there is one insuperable flaw, common to both of CBA’s methods of monetization, which looms particularly large: both mechanisms require individuals to predict how much they will enjoy some good, or how much they will suffer from some harm, if they experience it in the future. In the vast majority of cases, the individuals in question will have never experienced the harm or benefit. Thus, CBA will be attempting to put a value on, for instance, avoiding a case of cancer by looking to how much people who have never had cancer are willing to pay to avoid it.

Psychologists have amassed mountains of evidence demonstrating that individuals very frequently make significant mistakes when attempting to estimate how they will feel about some future experience (e.g., Wilson & Gilbert 2005). These mistakes, called “affective forecasting errors,” are extraordinarily difficult to avoid. In particular, multiple studies indicate that people struggle with attempting to gauge the magnitude and duration of the effects of disabilities (Gilbert & Wilson 2007; Wilson & Gilbert 2005). If people mis-estimate the impact of disabilities, then inferring perfect judgment from their behaviors about health risks will systematically bias estimates about the welfare effects of regulations. Similarly, in the environmental protection example, there is good reason to doubt that people will be able to accurately predict how much their lives will be improved by things such as increased wetlands protection (Desvousges et al. 1993). Importantly, these errors are not random; they exhibit systematic biases that do not just create noise but also produce more intractable distortions in valuation. Affective forecasting errors result in some goods being overvalued and others undervalued relative to their actual effects on people’s lives. Accordingly, there is every reason to believe that CBA—which relies on forward-looking estimations—is making significant errors in pricing costs and benefits (Sunstein 2007). Policymakers who rely upon CBA are likely arriving at flawed conclusions regarding which laws and regulations they should implement.

With CBA’s critical flaw in mind, we designed an alternative method for comparing the positive and negative consequences of a law (Bronsteen, Buccafusco & Masur 2013). This method, which we termed “well-being analysis” (WBA), would directly analyze the effect of various costs and benefits on people’s subjective well-being or quality of life. For example, a policymaker would assess an air quality law by comparing how much more people would enjoy their lives if they became healthier with how much less they would enjoy their lives if their consumption were reduced. This is the most natural and direct way to put seemingly incommensurable things on the same scale. And it yields the specific answer that is needed: whether a law will make people’s actual experience of life better or worse on the whole.
WBA relies on the same basic cost-benefit-weighing principle that undergirds CBA: all else equal, regulations whose benefits exceed their costs are valuable because they enhance overall welfare. The main difference between the two techniques involves the way in which costs and benefits are calculated and compared. Instead of monetizing the effects of regulation, WBA “hedonizes” them. That is, it measures how much a regulation raises or lowers people’s enjoyment of life. To do that, it relies on hedonic psychology data that measure how different factors affect people’s enjoyment of their lives. The positive and negative hedonic impacts can then be compared with one another. They are the relevant costs and benefits.

Instead of converting regulatory effects into monetary values, WBA converts them into well-being units (WBUs). WBUs are intended to be subjective, hedonic, cardinal, and interpersonally comparable units that indicate the degree of a person’s happiness for a given period of time. They are, in some respects, similar to the quality-adjusted life years (QALYs) that are increasingly popular in health economics.

WBA maps a person’s subjective well-being onto a scale that runs from -10 to 10, in which 10 indicates perfect happiness (subjectively defined), -10 indicates perfect misery, and 0 indicates neutrality or the absence of experience. This type of scale would allow individuals to register experiences that are worse than nonexperience (undergoing torture, for instance) and would simplify the comparison between experience and nonexperience. Each decile of the scale is equivalent and indicates a 10 percent change in the person’s subjective well-being. One well-being unit is equivalent to 1.0 on the scale for a period of one year. Thus, if a person lives to the age of 100 and has a subjective well-being of 7.0 out of 10.0 for each year, that person has experienced 700 WBUs (7.0 WBU/year × 100 years). If an event such as illness causes a person’s SWB to drop from 7.0 to 5.5 for a period of ten years, that person loses 15 WBUs (1.5 WBU/year × 10 years).

This type of scale has significant benefits for any type of decision analysis, particularly regulatory analysis, because it enables the direct comparison of the hedonic impacts of proposed policy changes. Imagine, for example, that the Occupational Safety and Health Administration (OSHA) is contemplating a simple regulation of workplace safety that will prevent 100 workers from each losing an arm while on the job. Implementing such a measure, however, will increase the costs of production and force factories to fire 300 workers in the affected industry.

CBA would attempt to calculate the value of the regulation by monetizing the costs and benefits it generates. With respect to the costs, CBA could estimate the lost wages of the 300 unemployed people (although not the lost well-being entailed by those lost wages). The benefits, however, are trickier. Establishing a market price for the loss of an arm is a fraught enterprise. Given these shortcomings, the value CBA applies to the loss of an arm will be beset by a number of systematic errors, especially individuals’ poor ability to predict how events like losing an arm will affect them. Accordingly, CBA may substantially and systematically misstate the benefits of the regulation.

WBA would approach the measure in the same general fashion but with different analytical data. Like CBA, WBA would attempt to quantify the cost of unemployment. But instead of looking solely to the workers’ lost wages, it would calculate the hedonic cost of being
unemployed by surveying individuals who have actually become unemployed and comparing their responses pre- and post-unemployment. Some data suggest that unemployment has a significant effect on well-being (Lucas et al. 2004). Thus, the welfare costs of unemployment may be much greater than CBA predicts. On the other side of the ledger, WBA is well positioned to hedonize the benefits of the regulation. Studies of people who have lost limbs provide credible information on the hedonic loss associated with losing an arm, and thus the benefits of avoiding such a loss (Frederick & Loewenstein 1999). Again, the results are likely to be different from those determined by CBA. Studies show that individuals who lose limbs often adapt substantially to their new condition, recovering most of their lost happiness within a few years (Frederick & Loewenstein 1999). This result is contrary to the predictions of healthy people, who typically assume that such disabilities will be devastating and discount the possibility that they will adapt to the loss (Ubel et al. 2005; Ubel et al. 2001). Accordingly, WBA holds the potential to deliver results that are far more accurate in welfare terms than CBA, despite the fact that WBA is a much “younger” procedure.

WBA thus holds two primary advantages over CBA. The first is that it will provide more accurate measures of benefits and costs because it relies on actual human experiences, rather than speculation about unknown futures. The second is that in relying upon subjective well-being rather than money, it employs a closer proxy for welfare than does CBA. Importantly, these two advantages are severable. Even if agencies should be attempting to maximize wealth rather than welfare, as some commentators believe (Weisbach 2013), WBA will still yield more accurate “prices” for the consequences of regulation than CBA.

Happiness Data as Prices

Although we believe that WBA offers the best method for estimating the effects of laws and policies, researchers have developed other approaches that improve upon traditional CBA. The biggest challenge for CBA involves estimating the value of non-market regulatory benefits, especially health and safety improvements. Often, CBA will assess the value of a regulation in terms of the number of lives it saves multiplied by an estimate of the value of a statistical life (VSL) that has been computed from revealed preferences data (Sunstein 2004). In addition to suffering from the standard flaws common to all revealed preferences data, these estimates have been criticized for ignoring the length of the lives that get saved. VSL treats equally saving the lives of hospice patients and saving the lives of teenagers. To remedy this, some scholars have proposed analyzing regulatory costs and benefits using the number of years of life that are saved and the value of a statistical life year (VSLY), which take into account the estimated life expectancy of those lives saved by regulation. But just as VSL ignores the length of lives saved, VSLY ignores their quality: five years spent in perfect health would be treated equivalently to 5 years spent in poor health.

Recently, some scholars have recommended instead using quality-adjusted life years (QALYs) in cost-benefit analysis. QALYs rely on survey responses from the public, patients, and medical professionals to estimate the impact of various health states on a person’s quality of life (Weinstein et al. 2009). QALYs were initially developed in the field of cost-effectiveness analysis to provide data on the efficient use of scarce resources in medical decisionmaking, but some commentators—including courts and agencies—see value in the use of QALYs in CBA.
As yet, however, QALY analysis faces a number of methodological hurdles before it can be successfully incorporated into CBA. The first difficulty with adopting QALY analysis as part of traditional CBA is determining how to monetize QALYs. When QALYs are used in cost-effectiveness analysis in healthcare decisionmaking, no effort is made to quantify the value of a QALY. But to incorporate QALYs into CBA, policymakers must place a dollar value on QALYs. As yet, however, no acceptable number has emerged (Hirth et al. 2000).

More problematic is the method that researchers use to elicit QALY values. The most typical methods for valuing QALYs are surveys, in which people are asked how much they would pay for a year of life at a certain level of health. Just as with contingent valuation and WTP studies, this forces people to guess at how they will feel in the future, raising many of the same problems with affective forecasting errors that undermine traditional CBA. This problem is compounded by the fact that QALY studies often require healthy individuals to make value judgments about health states that they have never experienced. To be valuable in welfare analysis, QALYs should reflect how people feel in various states of health. Instead, when healthy people are asked about states of poor health they will tend to provide answers about how they feel about those health states. These data do not provide a sound basis for regulatory policymaking.

Because of these problems, social scientists have increasingly turned to data from hedonic psychology to estimate the monetary effects of non-market goods. Unlike the studies typically used to inform CBA, happiness measures do not impose significant cognitive demands on subjects. Instead of having to estimate the probability of a risk, the effect of that risk, and the amount of money that would compensate for it, people only have to answer simple questions about how they are feeling and how satisfied they are with their lives (Clark & Oswald 2002). Researchers can use hedonic data and income data from cross-sectional and longitudinal panel surveys in regression analyses that control for other demographic variables to estimate the “shadow price” of a non-market good (Powdthavee & van den Berg 2011). For example, if researchers know the degree to which increased income improves well-being and how much an individual’s well-being is typically reduced by disability, they can estimate how much additional income a disabled person would need to make her as happy as a non-disabled person.

Using this method, researchers can establish “exchange rates” between income and a wide variety of non-market goods (Dolan, Fujiwara & Metcalfe 2011). Multiple studies have estimated shadow prices for health states using happiness data (Ferrer-i-Carbonell & Van Praag 2002; Groot & Van den Brink 2004; Oswald & Powdthavee 2008; Powdthavee & van den Berg 2011). For example, Powdthavee and van den Berg (2011) use different well-being measures to estimate the amount of additional income necessary to compensate an individual for a variety of health conditions. They estimate that, for a person with average household income, skin conditions reduce well-being by approximately £5000 per year, while experiencing migraine headaches would require an additional £3.2 million per year to make sufferers as happy as healthy people. In another paper, Oswald and Powdthavee (2008) estimate the compensating income associated with the death of a family member. They range from £39,000 per year for the death of a sibling to £286,000 per year for the death of a partner. Numbers such as these could inform both CBA and torts damages awards (Sunstein 2008).
Over the past decade, researchers have conducted pricing studies of a wide variety of non-market goods (and bads) based on happiness data (van Praag & Baarsma 2005). These include marriage and divorce (Blanchflower & Oswald 2004; Clark & Oswald 2002), crime (Powdthavee 2005), terrorism (Frey et al. 2009), natural disasters (Luechinger & Raschly 2009), and the environment (Welsch & Kühling 2009). Valuations of environmental goods are particularly helpful in the context of CBA. The environment is a major subject of regulation, yet it is especially difficult to value using revealed and stated preference studies. People may not appreciate the effect of the environment on their welfare, and they may be resistant to placing explicit monetary values on things like clean air. Accordingly, hedonically derived estimates about the compensating values of water, air, and noise pollution are especially promising.

As with other uses of happiness data for public policy, attempts to calculate implicit prices from changes in income and subjective well-being face difficulties. Questions arise over whether the happiness data can be treated cardinally, which would produce the most helpful and easily interpretable results. In general, however, issues of cardinality are no greater for happiness data than they are for money (Bronsteen, Buccafusco & Masur 2013). Another issue with studies that attempt to determine implicit prices of goods based on their hedonic impact is the potential endogeneity of income and happiness. Income and happiness may exhibit collinearity, and increased income may produce indirect as well as direct effects on well-being. More recent research, however, has attempted to measure and control for this endogeneity (Dolan, Fujiwara & Metcalfe 2011).

WBA and versions of CBA that use SWB data as prices are not mutually exclusive, and governments might alternatively prefer one or the other in different circumstances. Unlike WBA, in which all of the inputs are converted into hedonically measured WBUs, CBA with SWB-based prices retain monetary figures for all goods that are generally priced in markets, including wages, consumption, and the direct costs of implementing regulations. The relative merits of the two different metrics will thus depend upon whether CBA’s use of monetization is a sufficiently accurate proxy for welfare in some cases and whether it can be done more inexpensively than WBA.

Distributionally Weighted CBA

Another type of welfarist decision procedure that bears a family resemblance to WBA is distributionally weighted CBA. Like WBA, distributionally weighted CBA attempts to cure problems that arise from traditional CBA’s use of wealth as a proxy for welfare. As originally conceived by Robin Boadway (1974, 1975, 1976), distributionally weighted CBA springs from two separate insights about wealth distribution and welfare, one economic and one philosophical. The first is the well-understood idea of diminishing marginal utility of money. The more wealth an individual possesses, the less each marginal dollar means to that individual. As wealth increases, changes in wealth (increases or decreases) will have smaller effects on the individual’s welfare. That is, a millionaire should not care if she gains or loses $100, but for someone earning $10/day that sum of money could have a significant impact on welfare. Thus, equivalent amounts of money do not necessarily provide the same amount of welfare to different people.
The second insight is that in a fair or just society, all changes to individual welfare might not be equal in moral terms and should not be treated equivalently (Samuelson 1947:221; Harsanyi 1977). Imagine a society with just two individuals: Rich, who has a great deal of welfare, and Poor, who has very little welfare. A classical utilitarian would measure the welfare of that society—that is, compute the social welfare function (SWF)—as the sum of the individual welfare of Rich and Poor. A gain in welfare for Rich would increase the welfare of the society as much as an equivalent gain in welfare for Poor—no more and no less. However, this strictly utilitarian formulation ignores considerations of equity and distribution of welfare. If Rich gains welfare, that increases the welfare gap between rich and poor and makes the society even more inequitable. On the other hand, if Poor gains welfare, the society’s inequality diminishes. Depending on one’s normative view of social welfare, this might make the society better off. Boadway’s original insights have spawned a significant literature on the technical complexities of distributional weighting (e.g., Adler 2013; Cowell & Gardiner 1999; Creedy 2006; Dasgupta & Pearce 1972; Dasgupta, Sen & Marglin 1972; Dreze & Stern 1987), as well as a number of important critiques (Weisbach 2013).

To remedy the disconnect between wealth and welfare, distributionally weighted CBA typically incorporates two separate weighting terms: one to adjust for the different welfare values of money to richer and poorer individuals and a second to adjust for the different contributions to the social welfare function of changes in the welfare of different individuals. The inputs to the cost-benefit calculus, which are derived using traditional methods (WTP and contingent valuation studies), are then multiplied by these weighting terms. Ideally, economists would measure the first weighting term empirically, by determining the rate at which the marginal utility of money declines for the typical individual. The second term is a matter of normative philosophy (Adler 2013). A society that valued equality would use a weighting factor that was smaller for better-off individuals and larger for less well-off individuals, to reflect the judgment that changes in the welfare of individuals who are already doing very well is less important than the welfare of individuals who are faring poorly. A society that valued inequality would use a weighting factor that placed greater weight on the welfare of individuals who were already doing well. A strictly utilitarian society would use no weighting factor whatsoever. These two terms are often combined into a single term for any given individual, and this single term is typically referred to as the distributional weight. However, they are best understood separately, and we will consider them separately here.

What is the relationship between distributionally weighted CBA and WBA? Both tools arise from concerns about using wealth as a proxy for welfare, but their solutions to the problem are different. As we explained above, WBA incorporates two elements that are absent from traditional CBA. First, it uses hedonic data, rather than WTP studies or contingent-valuation surveys, to value goods. And second, it measures a close proxy for actual welfare, rather than wealth. Distributionally weighted CBA does not incorporate either of those elements. Proponents of distributionally weighted CBA would continue to rely upon the same WTP studies that traditional CBA uses to translate goods into monetary terms and would then add weights to correct for wealth effects.

Similarly, there is no analog in WBA to the fact that distributionally weighted CBA might place different values on the welfare of particular individuals when calculating overall
social welfare. Recall that a strictly utilitarian decision procedure would value the welfare of all individuals equally, while a more egalitarian society would value increases in the welfare of less well-off people more than increases in the welfare of better-off people. WBA was explicitly constructed as a utilitarian decision procedure in order to mimic CBA (Bronsteen, Buccafusco & Masur 2013). Yet this is not a necessary feature of WBA. That is, a decisionmaker could employ WBA in conjunction with any type of social welfare function, whether it is weighted to favor egalitarian outcomes or not. WBA as we originally conceived of it is not designed to be distributionally weighted in this manner. But if a decisionmaker were able to resolve the difficult normative questions involved in deciding upon a weighting factor, WBA could be easily adjusted to fit the desired normative outcome.

Indeed, even the original version of WBA overlaps considerably with distributionally weighted CBA in that each values changes in wealth differentially depending on the wealth of the affected individual. In this way, distributionally weighted CBA at least aims to capture actual changes in welfare more closely than does traditional CBA — the same goal that WBA seeks to achieve. Rather than simply measuring changes in total wealth, distributionally weighted CBA makes more of an attempt to measure changes in actual welfare. But where distributionally weighted CBA attempts to correct for flawed welfare proxies, WBA simply abandons them in favor of better ones.

Distributionally weighted CBA has also been criticized on the ground that it will promote welfare less efficiently than could be done through alternative means. Critics note that the tax system is generally believed to be a more efficient mechanism of redistributing wealth than legal rules or regulations (Kaplow & Shavell 1994). Accordingly, they argue, legal regulations should be wealth-maximizing, rather than welfare-maximizing. The tax system should then be used to redistribute wealth in whatever welfare-enhancing manner the decisionmaker desires. Such critics believe that the overall effect of wealth-efficient regulation plus welfare-enhancing taxes and transfers will lead to greater wealth and welfare gains than welfare-maximizing regulation alone (Weisbach 2013). This criticism implicates WBA as well, because WBA, like distributionally weighted CBA, is a welfarist decision procedure. However, Weisbach (2013) acknowledges that WBA might nevertheless produce more accurate measures of the welfare value of various goods than CBA. As we noted above, happiness data can be used to provide more accurate prices than WTP or contingent valuation studies. Even if critics of weighted CBA are correct, WBA and its hedonic approach to welfare may well be superior to CBA.

What then of this criticism? As a theoretical matter, critics of weighted CBA are likely correct to claim that it is more efficient to redistribute wealth via the tax system than through legal regulation. But CBA is not designed or implemented in a theoretical vacuum. The power to regulate and the power to tax and transfer do not reside within the same institution. Executive branch agencies, under the control of the President, are the institutions with primary regulatory authority. Tax policy, by contrast, is made through legislation passed by Congress. The critical pragmatic question is thus whether Congress will actually respond to regulation by taxing and transferring in such a way as to increase social welfare.

Critics of distributionally weighted CBA point to the vast number of changes that are made to the tax code every year (Weisbach 2013). Yet the rate of change, by itself, provides no
information on whether those changes are social welfare enhancing or diminishing, and whether they redistribute wealth in an equitable or inequitable fashion. To our knowledge there has been no systematic study of the multitude of tax changes. But it is at least suggestive that overall inequality in the United States has been steadily increasing for much of the twentieth century, including most of the past four decades. If Congress is attempting to redistribute through the tax system, it is doing a noticeably poor job.

If Congress cannot necessarily be counted upon to enact welfare-maximizing taxes, this leaves open the question of what an executive branch regulator should seek to maximize. She could employ standard CBA and attempt to maximize social wealth. She could employ WBA and attempt to maximize social welfare. Or she could employ distributionally weighted CBA and attempt to maximize some combination of wealth, welfare, and equality. Absent any reason to believe that Congress is more likely than not to tax and transfer effectively, it is hard to find a rationale for preferring unweighted CBA. At minimum, it seems obvious that any regulatory agency should consider not only the welfare effects of its own regulations but the combined welfare effects of its regulations and whatever tax policy it expects the taxing authority to implement (Adler 2013).

Weisbach (2013) pursues this critique one step further in arguing that proponents of weighted CBA need a theory of government to explain why the President, and not just Congress, should be involved in maximizing social welfare. Weisbach argues that there is no reason to believe that the President is any likelier to arrive at the welfare-maximizing outcomes than Congress, and thus no reason for the President to assume this authority. But this critique misunderstands both the existing structure of government and the role of welfare maximization in regulation. When the executive branch regulates, it is expected to do so in the public interest. The statutes empowering executive-branch agencies to regulate the environment, health, and safety certainly do not prohibit those agencies from considering the welfare effects of their regulations. To the contrary, they are generally expected to do so. It would not serve the public interest—and thus not comport with the executive’s congressionally mandated mission—to regulate in a way that diminished social welfare. An agency decision to maximize welfare is more, not less, in keeping with its statutory mission and the theory that underpins its delegation.

Happiness-Based Alternatives to Gross Domestic Product

Economic tools are used not only to analyze policies but also to measure the success and quality of nations. And just as SWB data can be used to replace monetary prices in the context of policy analysis, it can also be so used in the context of monitoring national well-being. As we will explain, this process relies on precisely the same principles as does WBA. The difference is merely that calculating national well-being is simpler in certain respects than is calculating the effect of policies on well-being. For example, policy analysis requires comparing things that are difficult to commensurate, whereas national monitoring requires only a summing of individual well-being. This summing is traditionally achieved via the proxies of production and consumption, but it can be achieved more directly via data on happiness.
The most common traditional economic measure of national success is Gross Domestic Product (GDP), the total market value of the goods and services produced within a nation’s borders in one year. Although it is not calculated for the specific purpose of analyzing or crafting public policy, it is nonetheless watched closely so that if it drops, “specific policies may then be developed to ensure it rises again” (Dolan, Layard, and Metcalfe, 2011:2). Despite its ubiquity and importance, GDP has always been controversial because, among other things, it measures economic productivity rather than quality of life. Nearly fifty years ago, Robert Kennedy said that it “measures everything . . . except that which makes life worthwhile” (Kennedy, 1968).5

Not long after Kennedy’s speech, the nation of Bhutan announced that it would pursue a measure of Gross National Happiness in order to address the concerns with GDP’s limitations (Krueger, 2009:1). But outside Bhutan, only recently has the idea been discussed with the serious goal of implementation. The first notable effort in this direction came from the Dutch social scientist Ruut Veenhoven (1996). Veenhoven argued that economic indicators like GDP, and also social indicators like the Human Development Index, measure primarily means of achieving quality of life instead of life’s quality itself. To remedy the problem, Veenhoven proposed an alternative to GDP called Happy Life Expectancy (HLE). Specifically, the average quality of life in a nation would be calculated by multiplying the average life expectancy of its citizens by their average happiness as measured by survey data on subjective well-being. The survey data could be either people’s responses to a single question about their happiness, a single question about their life satisfaction, or their “Affect Balance.” Affect Balance, which Veenhoven calls “the best choice” (1996:32), is measured by the “occurrence of specific positive and negative affects in the last few weeks” (1996:22).

Veenhoven’s idea of using subjective well-being to create an alternative to GDP gained little political traction until 2009, when France’s then-President Nicolas Sarkozy endorsed it, saying: “It is time for our statistics system to put more emphasis on measuring the well-being of the population than on economic production” (Samuel, 2009). Prompted by Sarkozy’s support, a group of renowned economists led by Joseph Stiglitz were commissioned to assess whether alternatives to GDP are desirable and feasible (Stiglitz, Sen, and Fitoussi, 2009). The Stiglitz Commission produced an influential report that made three main contributions. The first was to suggest ways to improve GDP itself; the second was to propose methods for measuring quality of life more directly; and the third was to discuss sustainable development and environmental concerns. Our focus here is on the second of these contributions: measuring quality of life.

The Commission noted that quality of life could be measured by data on subjective well-being; by capabilities such as objective measures of health and education, among many other things; or by “fair allocations,” an approach that asks people to place their own values on different non-market elements of the quality of life (Stiglitz, Sen and Fitoussi, 2009:42, 145-55). Within the category of subjective well-being, the Commission stressed the importance of separating three distinct ways of measuring SWB. The first is life satisfaction, the second is the

5 Kennedy was referring to the similar measure of Gross National Product (GNP), but the points he makes in his speech are clearly meant to apply equivalently to GDP. The difference is that GNP counts all goods and services owned by a nation or its citizens, including those produced outside the nation’s borders; whereas GDP measures all goods and services produced within a nation’s borders, including those that are foreign-owned.
presence of positive moment-by-moment affect, and the third is the absence of negative moment-
by-moment affect (Stiglitz, Sen, and Fitoussi, 2009:146).

The Commission recognized that pursuing a single-component measure of life quality
would involve value judgments and might well miss important nuances (Stiglitz, Sen, and
Fitoussi, 2009:207). But it also acknowledged that “the influence of GDP” may be “proof that
such indices are essential” (Stiglitz, Sen and Fitoussi, 2009). One option it gave for creating
such a measure focuses on hedonic experience: “While this approach may still be considered as
putting a domain—hedonic experiences—at the forefront, it may also be regarded as providing a
way of weighting different experiences through a common yardstick: the intensity of the hedonic
experiences that they generate” (Stiglitz, Sen and Fitoussi, 2009: 211).

Specifically, the Commission suggested that a metric known as the U-Index could be
used to create a GDP-style measure of a nation’s hedonic well-being (Stiglitz, Sen and Fitoussi,
2009:212). The U-Index is a metric created by Alan Krueger, Daniel Kahneman, and several co-
authors that “measures the proportion of time an individual spends in an unpleasant state”
(Krueger et al., 2009:9, 18). Using self-assessments of affect, the index divides a person’s time
into two categories — positive and negative — based on the strongest emotion a person reports
while doing the activity she was engaged in during that time.

The binary nature of the U-Index addresses a major concern about subjective well-being:
the concern that different people may use differently the numerical scale by which they report
how they feel. One person’s reported “8” out of 10 may not mean the same thing as another
person’s 8. Because the U-Index relies on an ordinal measure rather than a cardinal one, such a
problem is eliminated. Still, this strength is also a weakness. As Richard Layard has noted
(2009:145, 146-51), and as the creators of the U-Index have acknowledged (Krueger et al.,
2009:20), well-being itself is cardinal and non-binary: two activities may each be unpleasant (or
pleasant), but one may be substantially more unpleasant (or pleasant) than the other one. In our
view, this makes the U-Index significantly less desirable for use in WBA than simpler SWB
measures such as cardinal data on affect and life satisfaction. Moreover, Layard argues that the
danger of differential scale usage is overstated because the survey data behave, when subjected
to statistical analysis, in an orderly way that would be “inconceivable with a . . . scale that varies
widely across individuals” (2009:149).

Nevertheless, the U-Index may still make sense for use in GDP-style measures. The
reason is that such measures are often used for comparing different countries (Krueger et al.,
2009:69-77), and there is evidence that scale usage may vary across countries, perhaps due to
differences in culture or even language. For example, the French appear to use the top end of the
scale less than do Americans (Krueger et al., 2009:69-77), a problem that the U-Index is
uniquely suited to overcoming. In addition, despite the fact that the U-Index does not measure
degrees of pleasantness or unpleasantness, it does “take[]” on useful cardinal properties” due to its
aggregation of time (Krueger et al., 2009:20). In other words, although it does not say how
unpleasant an episode is, it says something meaningful about how unpleasant a day or a life is by
showing how much of the day or the life is spent in experienced unpleasantness. The upshot of
these strengths and weaknesses is that the U-Index is a useful tool that may be most fitting for
measures such as GDP that involve inter-country comparisons.
The Stiglitz Commission praised the U-Index not only because it solves the problem of differential scale usage, but also because it “naturally focuses attention on the most miserable in society” (Stiglitz, Sen, and Fitoussi, 2009:212). This is a theme in the literature on using subjective well-being measures to create a GDP-like indicator for the quality of life. Veenhoven, the first to try to create such a measure, more recently suggested a way to make the measure sensitive to inequality in happiness (Veenhoven and Kalmijn, 2005). His measure, Inequality-Adjusted Happiness (IAH), gives equal weight to the amount of happiness in a country and to the extent to which that happiness is evenly distributed (Veenhoven and Kalmijn, 2005:428), resulting in a composite score for each country between 0 and 100. Malta, Denmark, and Switzerland scored very highly, whereas several African and Eastern European countries scored the lowest, and the United States was toward the high-middle of the list (Veenhoven & Kalmijn, 2005:430). Interestingly, these results have many similarities to — but also notable differences from — the ranking of countries by Gini coefficients, which measure income inequality (CIA 2014). Countries like Malta, Denmark, and Switzerland do very well by either measure and several African countries do very poorly, but the United States does much better on the IAH measure than on the measure of pure income inequality, whereas Eastern European countries tend to have the opposite results.

In 2011, Paul Dolan and Robert Metcalfe followed up on the Stiglitz Commission report by surveying British citizens’ opinions of the extent to which government policy should aim to improve subjective well-being (2011). What they found was that people have “a strong preference for targeting the SWB of the worst-off as opposed to maximising the sum total of SWB in society” (2011:50). Reducing suffering was considered the most important aim of policy (2011:48-49). Dolan and Metcalfe thus recommended, among other things, that Britain “measure more about the experiences of daily lives, especially negative affect” (2011:49), and that it “estimate efficiency-equality trade-offs” (2011:50) so as to compare the value of helping the least happy people with the value of maximizing overall happiness.

The current state of the literature on this matter is that more data, especially on negative affect and (perhaps to a slightly lesser degree) on positive affect, should be collected. As the Stiglitz Commission wrote:

[T]he overwhelming conclusion that we derive from existing analyses of various aspects of subjective well-being is that these measures tap into QoL [quality of life] in meaningful ways . . . . The type of questions that have proved their value within small-scale, unofficial surveys should start being included in larger-scale surveys undertaken by official statistical offices. (Stiglitz, Sen and Fitoussi, 2009:151).

Once collected, the data can be used to create GDP alternatives based either on the U-Index, Veenhoven’s HLE or IAH, or other measures along those lines. We think that the best measure of a nation’s aggregate welfare — putting to the side normative issues about the possible wrongfulness of inequality — would be one like HLE that simply adds together all individuals’ subjective well-being throughout their lifetimes, preferably defining well-being as affect and measuring it via ESM or DRM. This is the same way that WBA calculates the costs and benefits of a regulation.
Conclusion

The main targets of policy analysis and of assessments of national success are the overall quality of human life and the way in which that quality is distributed among people. Traditional economic measures such as Gross Domestic Product and cost-benefit analysis have been criticized for ignoring distributional considerations and for being only weak proxies for quality of life. New advances in the study of subjective well-being or happiness have made it possible to create alternative devices for monitoring national success and for analyzing prospective policies.

We proposed well-being analysis as a method of policy assessment that holds several advantages over traditional economic measures. The chief advantage of WBA is that its data come from asking people a question they can answer relatively easily and credibly: how do you feel right now?\(^6\) By contrast, CBA relies on the assumption — contradicted by psychological evidence — that people accurately predict how much something will affect their well-being when they say how much money they would pay or accept for it.

WBA is the latest step in a progression toward using SWB data in policy analysis. Intermediate steps included using SWB data to determine “prices” for non-market goods, and weighting CBA to account for the distribution of wealth on welfare and equality. Although we see both of these as improvements upon traditional economic analysis, neither possesses all of the virtues of WBA.

Finally, the idea behind WBA is the same as the idea that has lately been animating the project of replacing GDP with a measure based on SWB data. The simplest approach — adding every citizen’s lifetime SWB — mimics one of the steps of how WBA assesses prospective policies.

Self-reports of happiness are the most valid proxies for the quality of human life. As such, it makes sense to use them as replacements for traditional economic indicators in both policy analysis and monitoring of national success. WBA is our contribution to that project.

\(^6\) Even the evaluative question of “How satisfied are you with your life?” is a far more valid gauge of well-being than is the assumption that someone uses money efficiently to increase his or her well-being.
Bibliography


Readers with comments should address them to:

Professor Jonathan S. Masur
jmasur@uchicago.edu
For a listing of papers 1–600 please go to Working Papers at http://www.law.uchicago.edu/Lawecon/index.html

602. Saul Levmore, Harmonization, Preferences, and the Calculus of Consent in Commercial and Other Law, June 2012
603. David S. Evans, Excessive Litigation by Business Users of Free Platform Services, June 2012
604. Ariel Porat, Mistake under the Common European Sales Law, June 2012
608. Lior Jacob Strahilevitz, Absolute Preferences and Relative Preferences in Property Law, July 2012
611. Joseph Isenbergh, Cliff Schmiff, August 2012
613. M. Todd Henderson, Voice versus Exit in Health Care Policy, October 2012
615. William H. J. Hubbard, Another Look at the Eurobarometer Surveys, October 2012
616. Lee Anne Fennell, Resource Access Costs, October 2012
617. Ariel Porat, Negligence Liability for Non-Negligent Behavior, November 2012
618. William A. Birdthistle and M. Todd Henderson, Becoming the Fifth Branch, November 2012
620. Rosa M. Abrantes-Metz and David S. Evans, Replacing the LIBOR with a Transparent and Reliable Index of Interbank Borrowing: Comments on Wheatley Review of LIBOR Initial Discussion Paper, November 2012
621. Reid Thompson and David Weisbach, Attributes of Ownership, November 2012
626. David S. Evans, Economics of Vertical Restraints for Multi-Sided Platforms, January 2013
627. David S. Evans, Attention to Rivalry among Online Platforms and Its Implications for Antitrust Analysis, January 2013
632. Adam B. Cox and Thomas J. Miles, Policing Immigration, February 2013
633. Anup Malani and Jonathan S. Masur, Raising the Stakes in Patent Cases, February 2013
637. Lior Jacob Strahilevitz, Toward a Positive Theory of Privacy Law, March 2013
639. Lisa Bernstein, Merchant Law in a Modern Economy, April 2013
640. Omri Ben-Shahar, Regulation through Boilerplate: An Apologia, April 2013
641. Anthony J. Casey and Andres Sawicki, Copyright in Teams, May 2013
643. Eric A. Posner and E. Glen Weyl, Quadratic Vote Buying as Efficient Corporate Governance, May 2013
646. Stephen M. Bainbridge and M. Todd Henderson, Boards-R-Us: Reconceptualizing Corporate Boards, July 2013
647. Mary Anne Case, Is There a Lingua Franca for the American Legal Academy? July 2013
651. Maciej H. Kotowski, David A. Weisbach, and Richard J. Zeckhauser, Audits as Signals, August 2013
652. Elisabeth J. Moyer, Michael D. Woolley, Michael J. Glotter, and David A. Weisbach, Climate Impacts on Economic Growth as Drivers of Uncertainty in the Social Cost of Carbon, August 2013
656. Evidentiary Privileges in International Arbitration, Richard M. Mosk and Tom Ginsburg, October 2013
658. The Impact of the U.S. Debit Card Interchange Fee Regulation on Consumer Welfare: An Event Study Analysis, David S. Evans, Howard Chang, and Steven Joyce, October 2013
659. Lee Anne Fennell, Just Enough, October 2013
662. Have Inter-Judge Sentencing Disparities Increased in an Advisory Guidelines Regime? Evidence from Booker, Crystal S. Yang, March 2014
665. Lee Anne Fennell and Eduardo M. Peñalver, Exactions Creep, December 2013
666. Lee Anne Fennell, Forcings, December 2013
674. Tom Ginsburg and Thomas J. Miles, The Teaching/Research Tradeoff in Law: Data from the Right Tail, February 2014
676. Nuno Garoupa and Tom Ginsburg, Judicial Roles in Nonjudicial Functions, February 2014
681. Yun-chien Chang and Lee Anne Fennell, Partition and Revelation, April 2014
682. Tom Ginsburg and James Melton, Does the Constitutional Amendment Rule Matter at All? Amendment Cultures and the Challenges of Measuring Amendment Difficulty, May 2014
688. Robert Cooter and Ariel Porat, Disgorgement for Accidents, May 2014
690. Robert Cooter and Ariel Porat, Lapses of Attention in Medical Malpractice and Road Accidents, June 2014
691. William H. J. Hubbard, Nuisance Suits, June 2014
692. Saul Levmore & Ariel Porat, Credible Threats, July 2014
694. Adam Chilton and Mila Versteeg, Do Constitutional Rights Make a Difference? August 2014
695. Maria Bigoni, Stefania Bortolotti, Francesco Parisi, and Ariel Porat, Unbundling Efficient Breach, August 2014
697. David A. Weisbach, The Use of Neutralities in International Tax Policy, August 2014
699. Saul Levmore, Inequality in the Twenty-First Century, August 2014
702. Dhammika Dharmapala, What Do We Know About Base Erosion and Profit Shifting? A Review of the Empirical Literature, September 2014
705. Thomas J. Miles and Adam B. Cox, Does Immigration Enforcement Reduce Crime? Evidence from 'Secure Communities', October 2014
706. Ariel Porat and Omri Yadlin, Valuable Lies, October 2014
707. John Bronsteen, Christopher Buccafusco and Jonathan S. Masur, Well-Being and Public Policy, November 2014