

21st Century GDP: National Indicators for a New Era

BEN BEACHY AND JUSTIN ZORN JOHN F.
KENNEDY SCHOOL OF GOVERNMENT
HARVARD UNIVERSITY

PREPARED FOR CONGRESSMAN HANSEN CLARKE (MI-13)

FACULTY ADVISOR: RICHARD PARKER

EXECUTIVE SUMMARY

The US Department of Commerce has described Gross Domestic Product (GDP) as “the crowning achievement” of 20th century US economic policy.¹

This is not an exaggeration. In the eight decades since the introduction of US national income accounts, GDP has become the official barometer of business cycles, an indispensable measure of government performance, and a leading benchmark of living standards. It has, in other words, become a de facto headline indicator of economic, political, and social progress.

Yet GDP was never intended for such a role. Economists dating back to Simon Kuznets, the father of US national accounting systems, have warned that GDP is a specialized tool for measuring market activity rather than national welfare.² While the indicator achieves its stated objective of capturing aggregate economic activity, it is agnostic as to what might be described—by both Republicans and Democrats—as core elements of national welfare in the 21st Century: social capital, economic mobility, health, education, entrepreneurship, environmental quality, and public safety. Indeed, GDP actually tends to rise with societal problems such as crime, pollution, household debt, commuting time, and family breakdown. As a short-term measure of economic output, it increases with the depreciation of machinery and the extraction of finite minerals, while failing to reflect the long-term contributions of education, research, and entrepreneurship. In a poignant illustration of the indicator’s inadequacy as a measure of welfare, market analysts recently noted that the 2010 Deepwater Horizon disaster would register as a net gain in GDP due to the expenditure involved in cleanup and rescue operations.³

In light of these shortcomings, this report, commissioned by US Representative Hansen Clarke of Michigan, seeks to answer an overarching question: *How should the US government institute supplemental national accounts that better reflect the welfare of the nation’s people?* The report’s central premise is that new comprehensive indicators would lead to better-informed policymaking, and, in turn, genuine improvements in national welfare. The report presupposes that GDP still serves an important, though limited, purpose and should not be replaced, but supplemented.

The task of supplementing the national accounts is complex yet achievable. This report seeks to advise Members of Congress and their staff on the three core elements of such an effort: (1) designing new indicators, (2) attaining the operational capacity in the executive branch to produce new indicators, and (3) overcoming political obstacles to reform. The core findings and recommendations presented in this document form the basis of the attached legislation, *The 21st Century GDP Act*.

Findings and Recommendations:

The report's core findings indicate that there is a clear window of opportunity for the development of supplemental national accounts. While significant technical, institutional, and political obstacles exist, these obstacles are surmountable.

- *Indicators-* Advances in data availability and statistical methods over the last forty years have enabled the development of over two dozen viable alternative measures of social and economic progress. Unlike GDP, these new supplemental indicators take factors such as health, safety, and educational attainment into account as determinants of the nation's welfare. Several US allies and one US state have begun incorporating such alternatives into their accounting systems. New indicators, nonetheless, bring genuine methodological complexities. The potential for resolving such complexities, along with the potential for policy impact, varies with distinct indicator designs.
Recommendation: Congress should prescribe the broad parameters of new, carefully-designed supplemental national indicators; it should launch a bipartisan commission of experts to address unresolved methodological issues.
- *Institutions-* With more than 20 statistical agencies and extensive experience aggregating indicators across departmental jurisdictions, the US government is well-equipped to begin developing and integrating new comprehensive indicators. In particular, the Bureau of Economic Analysis and Interagency Council on Statistical Policy can likely help orchestrate an interagency process. Nonetheless, challenges to be resolved include data availability, timeliness, standardization, and cross-jurisdictional cooperation between statistical agencies.
Recommendation: Congress should task the new bipartisan commission with submitting an "indicator development plan" to the President, outlining ways in which new indicators might be computed, aggregated, and reported; Congress should also appropriate new funding for data collection and statistical work if recommended by the commission.
- *Interests-* Prominent critics of GDP have run the ideological gamut from Robert F. Kennedy to Reagan advisor William *Bennett*, and with good reason: new indicators could better reflect the interests of diverse groups including *social conservatives* (reflecting the value of family care), *green businesses and organizations* (reflecting the value of untapped natural resources), and *innovators* (reflecting the long-term value of R&D and entrepreneurship), among others. There are nonetheless interest groups such as the mining industry that would likely support exclusive reliance on GDP.
Recommendation: Congressional champions of reform should seek the broadest possible bipartisan coalition in support of new indicator development.

Overview of the 21st Century GDP Act

Included in Appendix 1 is draft legislation based on the recommendations in this report. Central to these recommendations is the notion that Congress should act now to modernize the US national accounts in spite of some remaining methodological questions. The *21st Century GDP Act* creates a bipartisan expert commission to address these remaining issues while recommending means of bolstering operational capacity in the US statistical services.

The goals of the legislation are to:

- Set parameters for a series of four new indicators
- Delegate technical decisions to a team of experts selected on a bipartisan basis
- Establish an action plan for implementing new indicators in federal statistical work

The centerpiece of the bill is the establishment of a “21st Century National Indicators Commission”, which draws on successful bipartisan commission models such as the Base Closure and Realignment Commission.

Mandate: The bill charges the commission with determining how US statistical agencies can compute a new series of indicators that better reflect the well-being of the nation’s people. Just as consumer price index and unemployment are reported as a series (e.g. U1-U6), GDP would be reported as a series of indicators ranging from G1 to G5. G1 would be GDP (as currently measured), while G2-G5 would be GDP adjusted for various factors related to current and future welfare such as capital investment, household labor, educational attainment, crime, consumer debt, resource depletion, volunteer labor, entrepreneurship, and others. The commission’s key responsibilities, as stipulated in the bill, are to determine the formulae and pricing guidelines for these adjusted measures of GDP. The commission is moreover responsible for recommending how the Executive agencies can acquire the data necessary to produce these new indicators and how the Bureau of Economic Analysis can aggregate the new indicators. The commission has the power to hold hearings, subpoena witnesses, and contract studies in support of its work.

Bipartisan Membership: The bill calls on the President, in consultation with the leadership of both parties in Congress, to appoint four commissioners based on expertise in relevant fields. The President may appoint eight additional qualified commissioners, no more than four of whom can be of the same party.

Process: The commission has up to two years to submit an “Indicator Development Plan” to the President, who may either approve the plan without amendment or return the plan to the commission with recommendations for revision. If the President accepts the indicator development plan, the Bureau of Economic Analysis, in consultation with the Interagency Council on Statistical Policy, must begin taking steps to implement the plan.

Funding: The bill authorizes to be appropriated such sums as may be necessary for the work of the commission. To minimize costs, commissioners would be unpaid and staff would, where possible, be seconded from the statistical agencies.

Talking Points

Messaging to a General Audience

- Advances in statistics, computing, and data collection have made it possible to meaningfully estimate national wellbeing. Eight decades after the creation of the national accounts, we now have the means to upgrade.
- While the US has led the world in the development of GDP, we are now at risk of falling behind countries including China and the UK as they modernize their national accounts to measure more than industrial output.
- GDP growth tends to accelerate with rising crime rates, air pollution levels, and average commuting distance, while slowing with more prevalent vacation days, forest preservation, and family-cooked dinners.
- Because it disregards future growth potential, GDP rises with the depreciation of machinery and the extraction of finite resources from the ground, while failing to reflect the value of education, research, and entrepreneurship.
- This bill recognizes that GDP remains effective in measuring the *economy's* output; it does not seek to change GDP but rather to supplement it.
- This bill addresses a bipartisan problem with a bipartisan solution. The new commission would feature respected statistical experts chosen by both parties. This would help take politics out of the business of statistics.

Messaging to Progressives

- As Robert F Kennedy argued, "GDP measures everything, in short, except that which makes life worthwhile."
- GDP ignores mounting inequality. A rise in GDP per capita over the last decade has masked a fall in US median incomes. New supplemental measures should reflect real economic mobility—not just raw output.
- As former World Bank economist Herman Daly has argued, "the current national accounting system treats the earth like a business in liquidation." New supplemental indicators should reflect the value of the planet's finite resources.

- GDP ignores the price of pollution. New indicators should reflect the value of clean water, clear air, a healthy climate, and vibrant ecosystems.
- As people get sick and incur medical costs, GDP rises. New supplementary indicators should value investments in preventative health.

Messaging to Conservatives

- To run the US government with the efficiency and effectiveness of a business, we need more than a mere cash flow statement, which is what GDP represents. We need a “national balance sheet” that reveals assets and liabilities.
- GDP draws no distinction between earned and borrowed money. Supplemental indicators should take into account levels of both personal and national indebtedness.
- GDP rises as the government spends more tax dollars—even if it builds a “bridge to nowhere.” We need supplemental indicators that reflect only productive uses of taxpayer dollars.
- GDP rises as daycare replaces parental care. New indicators should reflect the importance of family to the economy and society.
- As crime increases in US cities, so too does GDP. We need new measures that reflect the value of law and order for households and businesses.
- Our nation’s extraordinary spirit of volunteerism actually slows GDP growth. New measures should account for the value of charitable work.
- GDP fails to reflect the value of entrepreneurship, a core driver of national prosperity. New supplementary measures should reflect the vital role of innovation and enterprise in the US economy.

TABLE OF CONTENTS

1. INTRODUCTION.....	1
1.1 CoRE QuESTIONS.....	3
2. BACKGROUND.....	5
2.1 DoMESTIC PRoDUCT AND ITS DiscoNTENTs.....	5
2.2 OBSTACLES TO REFORM.....	8
3. WHY NEW MEASUREMENTS MATTER.....	11
3.1 DIAGNOSING THE GAPS BETWEEN GDP AND WELFARE.....	11
3.2 WE GET WHAT WE MEASURE.....	15
4. METHODOLOGY.....	17
5. INDICATORS.....	19
5.1 AGGREGATED INDICATORS vs. DISAGGREGATED DASHBOARDS.....	20
5.2 INDICATOR FRAMEWORK: SuBJECTIVE vs. CoMPOSITE vs. ADJUSTED-GD.....	25
5.3 VARIABLE SELECTION.....	35
5.4 OvERCOMING OBSTACLE 1: INDICATOR CoMPLEXITIES.....	41
6. INSTITUTIONS.....	45
6.1 RoLEs: CoMPUTATION, AGGREGATION, AND REPORTING.....	46
6.2 OvERCOMING OBSTACLE 2: INSTITUTIONAL INERTIA.....	48
7. INTERESTS.....	51
7.1 OvERCOMING OBSTACLE 3: INTEREST GROUP OPPOSITION.....	51
8. CoNCLUSION.....	57
9. APPENDICES.....	58
APPENDIX 1: THE 21sT CENTURY GDP AcT.....	58
APPENDIX 2: UNDERLYING PREMISES.....	70
APPENDIX 3: TAXONOMY OF ALTERNATIVE INDICATORS.....	74
APPENDIX 4: TIMELINE OF NEW INDICATORS AND RELATED INITIATIVES.....	80
10. SouRcEs.....	82
PERSONAL INTERVIEWS.....	82
ENDNOTES.....	84
WoRKS CoNSULTED.....	86



According to market analysts, the 2010 Deepwater Horizon disaster-the largest oil spill in history-likely registered as a net gain in GDP.

1

INTRODUCTION

GDP occupies a unique unofficial position as the “headline indicator” of political and economic progress in the United States.

In the eight decades since the federal government first instituted the measure, it has come to serve as the indispensable barometer of business cycles, government performance, and even living standards. Political leaders rely on GDP as an essential benchmark in designing and justifying fiscal, monetary, and regulatory policy; businesses employ it as a key signal for determining expenditure and investment; and journalists and voters routinely look to it as a proxy for presidential success or failure.

GDP was not designed for this all-encompassing role, and it is especially ill-suited to play such a role today. Simon Kuznets, the economist Congress tasked with developing the indicator in the 1930s, famously declared “the welfare of a nation can scarcely be inferred from a measurement of national income.”⁴ While the indicator achieves its stated objective of accurately capturing aggregate economic activity, it does not reflect essential elements of national welfare in the 21st Century: factors including social capital, economic mobility, health, education, entrepreneurship, environmental quality, and public safety.

Rising GDP, in many cases, runs counter to such important dimensions of national welfare. During the 2010 Deepwater Horizon disaster in the Gulf of Mexico, J.P. Morgan analysts noted that economic activity generated by cleanup efforts would likely outweigh losses to tourism or fishing.⁵ They concluded, in other words, that the largest oil spill in history would likely register as a net gain in national output. This is not an isolated instance. GDP growth tends to accelerate with rising crime rates, air pollution levels, and average commuting distance, while slowing with more prevalent vacation days, forest

preservation, and family-cooked dinners. Because it disregards future growth potential, GDP rises with the depreciation of machinery and the extraction of finite resource deposits, while failing to reflect the value of education, research, and entrepreneurship.

The nation can do better in measuring wellbeing. Over the past fifty years, public figures ranging from Robert F. Kennedy to Reagan adviser William Bennett have pointed to the need to move beyond GDP. Academics have developed a substantial literature on the measurement of sustainable social and economic welfare. Foreign countries and US states have experimented with new comprehensive accounting and benchmarking systems. Today in the US, growing signs of frustration with existing economic and governance frameworks-coupled with advances in data collection and statistical analysis-have opened a unique window for innovation in national accounting. Supplementing the national accounts with state-of-the-art measures of sustainable wellbeing could have profound impacts on public policy, as well as private sector and individual decision-making.

DEFINING GDP

GDP tracks total market activity in a nation by adding together the value of all the final goods and services that are produced and traded for money within a given period of time.

The measure is made up of four components: (1) household consumption expenditures, (2) government expenditures, (3) net exports (the value of the nation's exports minus the value of its imports), and (4) investment (the increase in value of a nation's stock of monetized capital goods). As opposed to Gross National Product (GNP), GDP measures economic activity within a nation's borders.

1.1 Core Questions

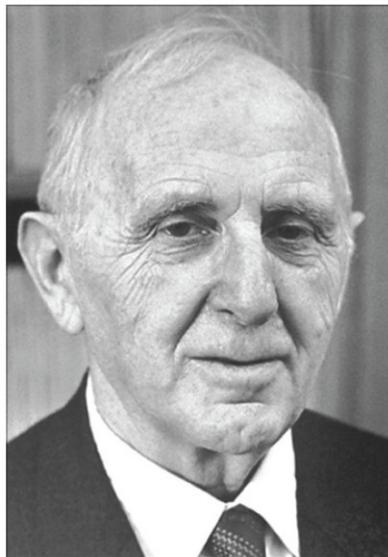
This report, commissioned by Congressman Hansen Clarke of Michigan, seeks to examine an overarching question:

How should the US government institute supplemental national accounts that better reflect the welfare of the nation's people?

The report is organized to address three major aspects of indicator development, adoption, and implementation:

- *Indicators*- designing new measures that are likely to impact policy while overcoming the methodological uncertainties and political complexities inherent to indicator design
- *Institutions* -attaining the requisite operational capacity to effectively produce new indicators in the executive branch
- *Interests*- building the political coalition needed to enact reform

As a prelude to addressing these three main areas of inquiry, the report examines the history of national accounting in the United States and efforts at reform; it moreover presents an explanation of the rationale underlying the methodology, findings, and recommendations.



Simon Kuznets, the father of US national accounting systems, famously warned, "the welfare of a nation can scarcely be inferred from a measurement of national income."



2 BACKGROUND

2.1 Domestic Product and its Discontents

As the Great Depression began, the US government had only crude estimates—derived from sources such as freight car records and mail-order purchases—to assess the depths of the economy’s downturn. In 1932, lamenting this lack of comprehensive metrics, Congress commissioned Kuznets, then of the National Bureau of Economic Research, to lead the development of the nation’s first aggregate income accounts.⁶ Both the Second World War and the ascendancy of Keynesian ideas about countercyclical economic policy heightened the need for macroeconomic measurement, and, in the early 1940s, the federal government adopted an expenditure-based model to track total national production for wartime planning.⁷ The modern system of National Income and Product Accounts (NIPAs) was born.

While the NIPAs were expanded and refined in the immediate postwar period, their salience in policymaking and the public discourse grew rapidly as a result of broader changes: namely, the emergence of economic growth as the *raison d’être* of the executive and legislative branches of government. This notion was codified in legislation, including the Employment Act of 1946, which gave the federal government both the personnel and the formal mandate to begin managing the macroeconomy. It was established further in communiqués such as NSC-68, which linked US defense and security policy to the goal of maximizing output. It was perhaps most strongly conveyed with the surfacing of GDP growth (at that time, typically reported as GNP growth) as an election issue and key marker of presidential performance. The United States government committed itself to perpetually increasing economic growth, with the President taking on a new role as “economist-in-chief.”

With the cultural shifts of the 1960s and 1970s came challenges to the paradigm of “endless growth,” and, in turn, challenges to GDP as the nation’s most prominent benchmark. The release of Rachel Carson’s *Silent Spring* in 1962 incited calls for the incorporation of externalities such as air and water pollution into the price system. Backlash against the Vietnam War provoked calls for new cost-benefit analyses regarding military expenditure. Writers including Betty Friedan raised questions regarding society’s failure to assign value to

work undertaken in the home. But crucially, criticism of GDP was not limited to the political left. In the early 1990s, William Bennett, a prominent social conservative and Education Secretary during the Reagan Administration, produced a report describing increasing social decay—including crime, family breakdown, and media addiction—in an age of rapid GDP growth.⁸ He called his report the "Index of Leading Cultural Indicators," a reference to the Commerce Department's annual Index of Leading Economic Indicators, which centers on GDP. Meanwhile, business-oriented think tanks and advocacy groups have long conveyed dissatisfaction with GDP's tendency to discount longer-term drivers of wealth such as investment and entrepreneurship.

These political pleas have resonated in the academy. From the Sustainable Measure of Economic Welfare introduced by William Nordhaus and James Tobin in 1973 to the Quality of Development Index proposed by the Tellus Institute in 2011, scholars in economics, statistics, sociology, and other fields have taken part in a diligent search for supplemental measures of national welfare. The momentum of new indicator development has risen—over half of the two dozen measures proposed to date have emerged in the last decade. In 2008, France's center-right President Nicolas Sarkozy commissioned a team of top economists, led by Nobellaureates Joseph Stiglitz and Amartya Sen, to assess the increasing range of alternatives. The commission's 291-page report concluded, "the time is ripe for our measurement system to shift emphasis from measuring economic production to measuring people's well-being."⁹

In the last five years, a growing number of governments have started making this shift. Following the European Union-sponsored "Beyond GDP" conference in 2007, the national statistical agencies of Germany, France, and the United Kingdom each committed to construct measures of sustainability and welfare.¹⁰ Since then, the UK has polled 34,000 people to determine the construction of a well-being index,¹¹ while France's National Institute of Statistics and Economic Studies has launched a menu of sustainability and quality of life indicators.¹² In 2011, Bhutan unveiled a second generation of the Gross National Happiness measure designed to drive its distributional and regulatory policies. Within the US, Maryland's Department of Natural Resources recently



Reagan Era William Bennett, author of the "Index of Leading Cultural Indicators"

Beyond GDP



WWF



Logo for the 2007 "Beyond GDP" conference, convened by the European Commission

began computing a Genuine Progress Indicator (GPI) that adjusts gross state product on the basis of numerous variables, from deforestation to auto accidents. (For assessments of the relative merits of these indicators, see the Indicators section).

Efforts to supplement or reform GDP at the federal level in the US have nonetheless had mixed results. On Earth Day in 1993, President Clinton announced his intention to begin "greening" the national accounts by instructing the Bureau of Economic Analysis (BEA) to produce prototype estimates of the economic worth of environmental resources.^Y The agency made progress toward the development of a system of Integrated Environmental and Economic Accounts until Democratic Representative Alan Mollohan of West Virginia, a major supporter of the coal industry, ended the initiative through an amendment to a House appropriations bill.¹⁴ On the heels of this defeat, Redefining Progress, a nonprofit founded to contest the

role of GDP and develop and disseminate GPI, became divided over questions of indicator methodology and outreach strategy and turned its focus toward environmental advocacy. Other initiatives, including the US Human Development Project have had positive impacts on the discourse regarding social and economic progress at the state and local level but have not sought to build momentum toward the development of supplemental national measures of well-being within the federal government.

A hopeful federal-level development emerged from the Government Accountability Office (GAO) starting in 2003. The GAO's then-head, Comptroller General David Walker, hosted a forum on the need for a more consolidated and transparent system of key national indicators. The issue garnered Congressional support, particularly from GOP Senator Sam Brownback, who lamented the lack of common factual base for US political discourse. This set in motion a development process leading to formal Congressional authorization of the Key National Indicators System (KNIS) in 2010. Housed today in the National Academies, KNIS has not grown into a force to challenge GDP, however. It has reasonably opted to develop a dashboard of more than 200 indicators rather than a smaller number of "headline indicators" that could supplement or serve as a counterweight to GDP. At this time, KNIS has not received an appropriation and, accordingly, is not in operation.

2.2 Obstacles to Reform

Given the potentially broad range of constituencies with interests in reforming the system of US national accounts, why has progress been so limited? The history of US national accounting suggests three overarching obstacles.

Obstacle 1: Indicator Complexities

To effectively supplement *GDP*, a new measure needs to be built on sound methodology and reliable data. Yet decisions regarding methodology for new indicators are technically and politically fraught. The decision to incorporate variables such as household labor or water quality into a new measure demands sophisticated and generally agreed upon economic formulae to determine the value (or cost) of those factors. What is more, the relevant data must be available, produced at regular intervals, and of a sufficient scale and scope.

Obstacle 2: Institutional Inertia

Even if a government agrees upon new indicators, it is still difficult to compel numerous discrete statistical agencies to provide standardized data in a timely fashion. This is particularly the case in the US government, which lacks a central national statistical office and features notorious barriers to information-sharing between agencies. Many of the nation's various statistical operations are underfunded and, owing to an ethic of neutrality, averse to taking controversial actions that may offend stakeholders elsewhere in government, including Congressional appropriators.

Obstacle 3: Interest Group Opposition

Various factions may fear the impacts of supplementary national accounts. As the BEA example illustrates, industry groups may oppose pricing externalities such as pollution if it increases the perception that their work undercuts national welfare. Incumbent politicians may oppose alternative indicators if they diminish the extent of economic progress or even show negative net growth.

"TOO MUCH AND TOO LONG, we seem to have surrendered community excellence and community values in the mere accumulation of material things. Our gross national product ... if we should judge America by that -counts air pollution and cigarette advertising, ambulances to clear our highways of carnage. It counts special locks for our doors and the jails for those who break them. It counts the destruction of our redwoods and the loss of our natural wonder in chaotic sprawl. It counts napalm and the cost of a nuclear warhead, and armored cars for police who fight riots in our streets. It counts Whitman's rifle and Speck's knife, and the television programs which glorify violence in order to sell toys to our children.

"Yet the gross national product does not allow for the health of our children, the quality of their education, or the joy of their play. It does not include the beauty of our poetry or the strength of our marriages; the intelligence of our public debate or the integrity of our public officials. It measures neither our wit nor our courage; neither our wisdom nor our learning; neither our compassion nor our devotion to our country. It measures everything, in short, except that which makes life worthwhile. And it tells us everything about America except why we are proud that we are Americans."

Robert F. Kennedy Address, University of Kansas, Lawrence, Kansas,
March 18, 1968

3

WHY NEW MEASUREMENTS MATTER

How might more comprehensive indicators impact real lives in US communities?

Answering this question requires (1) identifying the gaps between GDP and national welfare that could be filled by new measures, and (2) describing how the enhanced measures would actually improve policies that impact lives.

3.1 Diagnosing the Gaps Between GDP and Welfare

To diagnose GDP's inadequacy as a welfare proxy, we identify three large gaps: omission of social and environmental variables that influence well-being, failure to account for changes in future growth potential, and conflation of narrowly-defined consumption with broader economic welfare.

Given the longstanding critiques of GDP as an inadequate gauge of national welfare (as outlined in the previous section), why not seek new measures to supplant rather than supplement GDP? Indeed, some GDP critics argue that, given GDP's ingrained usage, merely creating new measures alongside GDP means leaving the latter in its misplaced role as preeminent arbiter of national progress. Inserting other indicators into that role, they assert, requires finding measures that can replace the policymaking function of GDP.

Yet, trying to completely subordinate GDP is not politically or technically feasible in the short term. Finding supplemental measures to be reported alongside GDP is a more tenable prospect.

More importantly, seeking replacement of GDP constitutes a misplaced diagnosis of GDP's shortcomings. GDP remains worthwhile for its narrow original purpose. Measuring aggregate economic activity has largely proven helpful in devising countercyclical policy so as to limit unemployment and inflation. Without a means of assessing the magnitude of the 2008 recession, for example, policymakers would have had little guidance in crafting countercyclical policies in response.

Rather than viewing GDP as intrinsically flawed, a proper diagnosis would find fault with its inflated role as the headline indicator of national progress. As stated above, GDP was never designed to assume this broad purpose that the public discourse often foists onto the measure. Three critical gaps between GDP and national welfare emerge.

Three Gaps: How GDP Fails to Measure Well-Being

- **Omitting Social and Environmental Factors:** As a purely economic measure, GDP does not account for policy-relevant environmental and social factors that directly contribute to welfare. Such factors include crime, smog, commuting time, nutrition levels, leisure time, and enjoyment of parks and natural preserves. While GDP appropriately tracks some such factors, it runs counter to others. Crime boosts GDP via increased police budgets and purchases of alarm systems. Smog thickens as industrial production rises. Ascending obesity prompts increased spending on food and hospital bills. GDP serves as a perverse proxy for such environmental and social inputs to welfare.
- **Ignoring Future Growth Potential:** As a purely current measure, GDP does not account for the impact of today's policies on tomorrow's economic potential. GDP measures economic transactions, not economic stocks. Stocks are the reservoir of resources needed for future economic output: physical capital (e.g. machinery), human capital (e.g. educational attainment, health levels), social capital (e.g. interconnectedness of productive networks), financial capital (e.g. savings), and natural capital (e.g. subsoil minerals, unpolluted water). Since an increase in current-day economic activity can come at the expense of such stocks, GDP can act as a backwards predictor of sustained economic welfare. GDP also does not account for the potential technological advancement, another key to future growth, embodied in current research or entrepreneurship levels.
- **Conflating Consumption and Economic Welfare:** As a purely market-based measure, GDP does not even fully account for current-day economic welfare. Non-market goods and services that support economic standing, such as those provided for free by families, the government, non-profit organizations, or volunteers, are not counted. As these services are replaced by market exchanges (e.g. parents hiring a babysitter rather than staying at home with the

children), the resulting rise in GDP masks unchanged purchasing power, not to mention declining family and societal cohesion. In addition, not all expenditures should be counted equally in assessing economic welfare. While consumption tends to bring satisfaction, people tend to be less satisfied with consumption that requires incurring inordinate debt burdens. Similarly, most economists agree that an extra dollar spent by a poor person who has not yet attained basic standards of living will generate greater satisfaction than an extra dollar spent by a rich person who already has their needs met. To more accurately represent economic welfare, GDP would thus need to discount consumption financed by high debt levels or skewed towards high income brackets.

Figure 1:

Gaps between GDP and Welfare

These three large gaps between GDP and overall welfare are represented by the blue space in Figure 1. Overall welfare includes three primary domains (economic, environmental, and social) and two primary time periods (current and future). By contrast, GDP (represented in red) only targets one domain (economic) in one time period (current), and only partially captures that realm.

By highlighting GDP's narrowness, this report identifies the gaps that supplemental indicators should seek to fill. New metrics should provide a fuller approximation of current economic welfare, should account for future economic potential, and should incorporate environmental and social inputs.





Figure 2: Given the gaps between GDP and welfare, GDP often acts as a perverse rubric of policymaking success. It tends to rise with smog and commuting time, while Falling with public parks and leisure time. It indicates growth even when current output trends threaten Future output potential via climate instability, capital depreciation, or natural resource depletion. It registers increasing disparity and crime as contributions to prosperity while counting Family-cooked dinners as an impediment.

3.2 We Get What We Measure

Upon constructing new indicators to fill in the gaps between GDP and national welfare, how will the new statistics actually impact policymaking? This report contends that better measurement would engender better policies, bolstering living standards. First, indicators do so by influencing the policy narrative—the public discourse commonly used to define policy problems and frame the picture of policymaking success. Headlines, think-tank publications, election debates, and advocacy campaigns all perpetuate this frame. Numerical benchmarks, meanwhile, help shape the frame. New metrics could shift the discourse by changing headlines. Consider, for example, the symbolic power of headlines declaring, "Additional Years of Schooling Would Reduce GDP, Boost Well-Being," or Congressional hearings revealing that recent economic growth has masked declining R&D and entrepreneurship. Such narrative shifts have the potential to prompt corresponding policy shifts.

Second, new indicators could more directly influence policymaking via benchmarking. The U.S. federal government has historically institutionalized particularly potent metrics as targets for policymaking. For example, the Federal Reserve implicitly aims for a 2% inflation rate in setting monetary policy while many Congressional members last year sought to formalize deficit reduction targets. Such benchmarking may be on the rise. The Government Performance and Results Act (GPRA) Modernization Act of 2010 requires that the Office of Management and Budget establish "long-term, outcome-oriented goals for...crosscutting policy areas" every four years.¹⁵ New welfare indicators could inform such holistic targeting, thereby institutionalizing a redefinition of progress.



4

METHODOLOGY

In determining the best alignment of indicators, institutions, and interests to fill in the above gaps, we rely on:

1. Literature: the background for this report stems from dozens of academic articles, commission reports, governmental documents, books, magazine features, and websites that analyze GDP, supplemental measures, and federal statistical agencies. See Endnotes and Works Consulted for the most important background documents.
2. Interviews: the brunt of the following analysis and recommendations relies on substantive interviews with 40 leading economists, statisticians, federal officials, and think tank personnel, including several authors of supplemental indicators. See the Personal Interviews section for the full list of interviews. (Given a stated commitment to not directly attribute statements to interviewees, this report does not quote or reference specific interviews.) We made no attempt to elicit statistically significant aggregate feedback from this panel of experts, given the inherent sampling bias. That said, we did aim to speak with experts ranging from the political left (e.g. the Institute for Policy Studies) to the political right (e.g. the Heritage Foundation), finding supporters and skeptics of supplemental measures across the political spectrum.



5 INDICATORS

Designing new measures that are likely to impact policy while overcoming the methodological uncertainties and political complexities inherent to indicator design

The search for potent, accurate, and feasible supplements to GDP begins with appropriate indicator design. What types of indicators are most likely to deliver the desired shift in public discourse and policymaking? How can the indicator be constructed to overcome the stated obstacles of methodological uncertainty, institutional inertia, and political opposition?

CORE RECOMMENDATIONS: New national indicators should: (1) be expressed as a single, aggregate number, (2) use dollars as the unit of measurement, and (3) comprehensively quantify a precisely-defined dimension of policy-relevant welfare. Using these guidelines, this section proposes four new indicators of national well-being to parallel GDP. A bipartisan commission of experts should be tasked with settling unresolved methodological issues and determining the specific construction of each new indicator.

Over two dozen existing alternative indicators provide a diverse menu of models for supplementing GDP (see Appendix 3). To determine the best models, this section uses three primary criteria:

- **Impact:** Optimal indicators would prove effective in broadening GDP-driven public discourse and guiding better-informed policymaking.
- **Accuracy:** Optimal indicators would be precise and comprehensive in what they seek to measure, while relying on sound methodology.
- **Feasibility:** Optimal indicators would minimize

bureaucratic hurdles and invite more political support than opposition.

The section employs each of these criteria to make recommendations on three successive decision points:

1. Should supplemental indicators be singular metrics of aggregated variables or dashboards of disaggregated variables?
2. Among aggregated metrics, which indicator framework is most appropriate: adjusted-GDP, composite, or subjective?
3. Using the adjusted-GDP framework, what specific dimension of welfare should each new indicator seek to measure, and what variables should be included?

Answering these three questions leaves unresolved the overarching obstacle of methodological uncertainty. Overcoming this obstacle requires more than an appropriate indicator framework—much of the uncertainty lies in deciding the specific variables to be included in each indicator and the particular methods of assigning value to each variable. Such decisions should be delegated to a body that offers political insulation and technical expertise. To this end, the section concludes by outlining the mandate and makeup of a bipartisan commission of statistical experts.

5.1 Aggregated Indicators vs. Disaggregated Dashboards

Aggregating multiple variables into a singular indicator is preferable to creating a dashboard of disaggregated variables, given the narrative potency of a singular number.

The taxonomy of alternative indicators in Appendix 3 includes metrics that compile distinct variables (e.g. income, inequality level, carbon emissions, educational attainment, self-reported utility, etc.) into a singular number. The number of aggregated variables ranges from the three included in the United Nations' Human Development Index to the 64 distinct statistics comprising the Canadian Index of Wellbeing.¹⁶ To compile such subcomponents into a summary statistic, aggregate measures use an array of methodologies for assigning prices or weights to each variable.

Other indicator reform efforts, however, have embraced a "dashboard" approach, in which multiple variables are presented alongside GDP without any attempt at aggregation. Numerous U.S. states currently host online dashboards of disaggregated variables pertinent to state priorities. At the

federal level, the Congress-commissioned Key National Indicator System (KNIS) plans to present an online constellation of over 200 policy-relevant variables.

The dichotomy between aggregate and dashboard measures is somewhat contrived. Dashboards, for example, can include aggregate measures alongside disaggregated variables (as is currently the plan for the KNIS). Aggregate indicators, meanwhile, are typically presented alongside their disaggregated components. For example, the website for Maryland's Genuine Progress Indicator (GPI) first shows viewers a graph indicating lackluster overall performance of GPI. Viewers can then click on any of the measure's 26 subindicators to see how declining performance in some variables (e.g. income equality) has offset progress in others (e.g. consumption), thereby explaining the net stagnation.



The following are the 26 Indicators that calculate the Maryland Genuine Progress Indicator. Select an Indicator to learn more about the topic, why it is included, and how Maryland is doing.

The online home of Maryland's Genuine Progress Indicator allows viewers to explore any of the 26 variables that comprise the aggregate indicator.

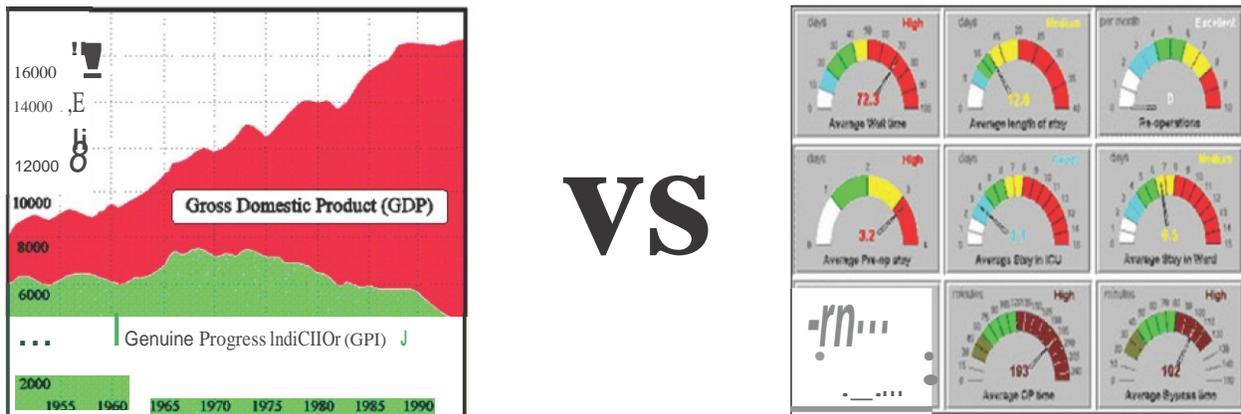
<i>Economic Indicators</i>	<i>Environmental Indicators</i>	<i>Social Indicators</i>
Personal Consumption Expenditures (1.asp)	Cost of Water Pollution (8.asp)	Value of Housework (21.asp)
Income Inequality (2.asp)	Cost of Air Pollution (9.asp)	Cost of Family Changes (22.asp)
Adjusted Personal Consumption (3.asp)	Cost of Noise Pollution (10.asp)	Cost of Crime (23.asp)
Services of Consumer Durables (4.asp)	Cost of Net Wetlands Change (11.asp)	Cost of Personal Pollution Abatement (24.asp)
Cost of Consumer Durables (5.asp)	Cost of Net Farmland Change (12.asp)	Value of Volunteer Work (25.asp)
Cost of Underemployment (6.asp)	Cost of Net Forest Cover Change (13.asp)	Cost of lost leisure time (26.asp)
Net Capital Investment (7.asp)	Cost of Climate Change (14.asp)	Value of Higher Education (27.asp)
	Cost of Ozone Depletion (15.asp)	Services of Highways & Streets (28.asp)
	Cost of Non-Renewable Energy Resource Depletion (16.asp)	Cost of Commuting (29.asp)
		Cost of Motor Vehicle Clashes (30.asp)

Yet, the distinction between the two approaches remains for those economists and policy advocates who question the merits of any attempt at aggregation. What follows are some of the key arguments for and against aggregation, categorized by the guiding principles of impact, accuracy, and feasibility.

Impact

Changing the narrative that conflates GDP with national progress requires a number that will actually fit into another narrative. Dashboards of several dozen (not to mention several hundred) distinct variables will not easily translate into headlines or soundbytes. While state-level dashboards such as Virginia Performs (with 49 variables), Minnesota Milestones (60 variables), and Oregon Benchmarks (158 variables) have facilitated more informed policymaking, they have not contended with GDP in swaying headlines or election cycles. By contrast, Maryland's aggregated GPI tells a clear, potent story about the state's sustainable economic welfare—that it has not risen for three decades. In recent months, this digestible takeaway has circulated in thinktank reports, high-profile publications like the *New York Times Magazine*,¹⁸ and the website of the Nieman Watchdog Journalism Project, a Harvard-based initiative to highlight worthwhile stories for journalists.¹⁹

Further, if other states would follow Maryland's lead and develop a parallel singular indicator, cross-state or cross-district comparisons could prompt competition between elected officials to outperform each other under the redefined rubric of progress. Building on successful precedents set in the Philippines and elsewhere, the New York-based Measure of America project currently seeks to spur such a "race to the top" among policymakers by disaggregating a variant of the Human Development Index for each Congressional district. The new measure has already started to sway several Congressional elections—in



Aggregate measures, such as the Genuine Progress Indicator depicted at left, tell a simpler, more powerful story than complex dashboards of disaggregated variables.

Fresno, California, candidates seeking to unseat Rep. Costa have pushed the incumbent to explain why his district received the lowest Measure of America ranking in the country.²⁰ Complex dashboards do not allow for such simple rankings and comparisons, making them ill-equipped to spur healthy competition in governmental performance.

To some analysts, however, the policy change potency of aggregate indicators poses a risk, not a benefit. They worry that aggregation will obscure the stories that really need to be told, prompting misleading policy prescriptions. For example, several economists interviewed noted that incorporating environmental variables alongside economic output likely would have understated the depth of the 2008 recession, given the positive influence of the reduction in output-linked pollution. Such critiques, however, underscore the importance of appropriate indicator design, not the need to abandon aggregate indicators. Indeed, the design of Maryland's GPI meant that the metric dipped even lower than Gross State Product during the 2008 recession, due to appropriate discounting for inequality and physical capital depreciation, both of which rose. Furthermore, any policies for which GDP might prove more illustrative than alternative aggregate metrics (e.g. countercyclical spending) will still be able to rely on GDP. The task at hand is not to replace the prevailing metric, but to supplement it.

Accuracy

As mentioned, aggregating distinct variables underscores Obstacle 1: uncertainty over how to attribute values to non-market variables (e.g. the cost of smog, the relative worth of good health). Debates over appropriate pricing methodologies, many of which remain unresolved, have fed reams of scholarly journals. Since the summary number is sensitive to such debates, opponents argue that an aggregate indicator cannot be trusted as an accurate portrayal of welfare. The response from proponents is twofold. First, not all variables bring such uncertainty—the values of near-market goods (e.g. coal deposits not yet mined), for example, are relatively non-controversial. Second, for those variables where significant uncertainty persists, ongoing research could help lend clarity. Indeed, when GDP was first proposed, many of its subaccounts were methodologically dubious. Since then, research advances have yielded enhanced calculation methods and data sources. The Bureau of Economic Analysis routinely incorporates such methodological advances, releasing a comprehensively revamped GDP definition every five years on average.²¹ Combining these two arguments,

an aggregate indicator could be constructed today that prioritizes variables of higher methodological certitude, while further research could allow relatively uncertain variables to be iteratively improved or incorporated. (See section 5.4 for related recommendations.) As several interviewees argued, the quest for a perfect supplemental indicator should not preclude the adoption of one that, while imperfect, constitutes a marked improvement over singular dependence on GDP.

Feasibility

Aggregation also increases the challenge of Obstacles 2 and 3: bureaucratic inertia and political division. Some analysts worry that aggregation would stall adoption of supplemental indicators by inviting greater bureaucratic inertia. The concern is that the over twenty statistical agencies of the federal government have not historically collaborated as closely as would be required to consolidate data from disparate domains into one account. Yet, precedents for such interagency collaboration do exist. Indeed, GDP itself is the product of an interagency process—before calculating the metric, the Bureau of Economic Analysis must source data from the US Census Bureau, Agriculture Department, Bureau of Labor Statistics, and the Treasury Department.²² If modeled after such successful collaboration, an interagency statistical task force could lower bureaucratic hurdles to efficient aggregation (See Institutions section.)

If poorly designed, the process of assigning relative weights or prices to distinct variables could also be stalled by politically-driven fractures. Industry representatives could push for an understated cost of pollution, while environmental groups could seek to overstate the costs. Politically-motivated disputes are of course not the correct means of determining variable values. Aggregation requires logically coherent methods, such as replacement cost determination, contingent valuation, or productivity calculation. To avoid political pitfalls and employ sound methodology, aggregation design should probably be the domain of a technical body relatively insulated from political influence (See Interests section). While this solution may not eliminate political tensions, it could shrink political impediments to a surmountable size.

Summary

Overall, this report concludes that aggregate indicators hold significantly more potential than disaggregated dashboards for broadening the policymaking narrative, given their simplicity and proven messaging

potency. To be useful as a policymaking tool in addition to a messaging tool, the disaggregated variables should also be made transparent. While aggregation poses real concerns about methodological uncertainty, such concerns can be ameliorated via appropriate variable choice and continual research. While aggregation also prompts valid questions of political and bureaucratic feasibility, choosing the appropriate bodies for aggregation design and interagency cooperation, both addressed below, could loosen such constraints.

Recommendation: choose aggregate indicators over disaggregated dashboards in order to maximize impact.

5.2 Indicator Framework: Subjective vs. Composite vs. Adjusted-GDP

Of the three primary types of alternative welfare indicators, the adjusted-GDP framework is best suited to supplement GDP. Subjective measures have shortcomings in impact, accuracy, and feasibility. Composite indexes, while best equipped to avoid political contention, present irreconcilable accuracy challenges and inferior narrative potency. The adjusted-GDP framework offers a more powerful policy tool with greater accuracy potential, while political challenges remain manageable.

Among the few dozen existing aggregate indicators of welfare (Appendix 3), the Stiglitz Commission on the Measurement of Economic Performance and Social Progress identified three main categories:²³

- 1) Adjusted-GDP Measures (e.g. Genuine Progress Indicator): these indicators begin with GDP (or some component thereof), altering the dollar figure to account for a wide range of welfare-relevant variables excluded from the traditional measure. Common adjustments include discounting GDP for income inequality, subtracting components of GDP seen as not contributing to welfare (e.g. defensive expenditures associated with increasing crime), and imputing prices for nonmarket variables that either contribute to welfare (e.g. household labor) or detract from it (e.g. pollution). Adding the benefits and subtracting the costs from GDP yields a dollar-denominated expression of policy-relevant welfare that internalizes market externalities.
- 2) Composite Indexes (e.g. Human Development Index): these measures assess relative performance on a series of welfare-relevant variables to arrive at a final number between zero and

one. The indexes are assembled by first picking the variables (e.g. educational attainment, income, pollution levels) and then creating a zero to one scale of relative performance for each variable, based on a range of plausible achievement. Countries receive scaled grades for all included variables, which are typically then averaged (with equal weighting) to determine the final index number.

- 3) Subjective indicators (e.g. Gross National Happiness): these measures employ the logic that the best judges of a people's welfare are the people themselves. Relying on surveys that ask respondents to rate their quality of life, these measures intend to capture a nation's average well-being, typically expressed as an index. Surveys range in complexity from a single question (e.g. All things considered, how would you rank your quality of life?) to the 5.5-hour survey used to compute Bhutan's Gross National Happiness. While aggregation methodologies differ, all subjective measures rely on self-reported well-being.

Each of these formats offers inherent strengths and shortcomings in the realms of impact, accuracy, and feasibility. The analysis below discusses these tradeoffs, as identified by economists, statisticians, bureaucrats, indicator authors, policy advocates, and other experts.

Accuracy

None of the three competing frameworks completely avoid Obstacle 1, methodological uncertainty, though the distinct designs overcome this challenge to differing degrees.

Four key problems inhibit the accuracy of the subjective framework in measuring policy-relevant welfare.

1. Self-reported happiness may not fully capture policy-relevant changes in well-being. This concern arises from two key findings of happiness studies. First, a person's happiness tends to remain fairly static despite changes in external factors, including policy-relevant ones. As a dramatic example of this tendency, both lottery winners and those who suffer spinal cord injuries tend to return to a nearly static level of self-reported happiness after the elating or tragic event.²⁴ Second, even if external changes can durably influence happiness, relative changes in living standards seem to matter significantly more than absolute changes. Studies have demonstrated that while happiness rises when a person's income surpasses that of peers, happiness remains fairly constant if the

person's income rises in tandem with peers.²⁵ This finding may explain the Easterlin paradox, an empirical observation that countries experiencing broad-based income growth have not seen a corresponding rise in happiness.²⁶ Such relativity makes happiness indexes problematic if absolute living standards are assumed to be important for well-being. These findings suggest that comparisons of subjective measures across time will tend to show static results, even if policy shifts deliver meaningful changes in welfare.

2. Self-reported happiness may capture factors outside the realm of policy impacts. Happiness levels can vary with weather, sports outcomes, spousal infidelity, genetics, and other factors not directly influenced by policy. As such, analysis of variance in subjective metrics could offer a distorted assessment of policy impacts.
3. Even if subjective metrics could fully and exclusively capture policy-relevant welfare, humans' irrational judgment would likely cloud the results of surveys. For example, humans tend to disproportionately weigh recent events, making survey results dependent on the day and time of administration. Some scholars, including current Council of Economic Advisers (CEA) Chairman Alan Krueger, have proposed innovative survey methods to surmount such measurement difficulties^P though self-reporting will likely remain less accurate than objective measures for the foreseeable future.
4. Even if all of the above problems could be overcome, subjective measures would still tend to overlook changes affecting future welfare potential. When asked to assess their quality of life, few people will account for machinery depreciation, carbon emissions, changes in subsoil mineral stocks, or other trends that are likely to affect future quality of life. Given this myopic tendency, subjective indicators will prove inadequate as measures of sustainable welfare.



Bhutan has made large strides in developing its innovative Gross National Happiness indicator. Such subjective measures, however, may be ill-suited for US policymaking purposes.

Composite and adjusted-GDP measures both avoid most of these accuracy shortcomings, given that both frameworks allow for the ex-ante choice of objective, policy-relevant variables. If a variable is not policy-relevant, it can be excluded. If a variable is likely to influence future welfare despite minimal bearing on current welfare, it can be included. Instead, the accuracy hurdle that comes with both adjusted-GDP and composite measures is how to reflect the relative importance of each included variable. Take, for example, two variables that negatively impact welfare: an additional hour of commuting time and an additional acre of deforestation. Adjusted-GDP measures use imputed prices to determine the relative importance of each: what is the dollarized opportunity cost of sitting in traffic for an hour versus the dollarized natural capital loss of an acre of forest? Composite measures, meanwhile, bring the option of assigning weights: should decreased performance in forestation be given less, equal, or greater emphasis than decreased performance in commuting time? Both methods bring challenges.

Some economists and statisticians doubt the reliability of adjusted-GDP variables due to uncertainty associated with imputing prices for non-market variables. Such uncertainty tends to increase the farther that a variable stands from the market. To use the same exemplary acre of forest, the untapped value of the trees as a source of timber can be readily determined by multiplying the quantity of trees by the current price of timber. The value of the forest's service of reducing atmospheric carbon concentration, meanwhile, cannot be imputed from an active carbon market (since there is not one currently in the United States), though ample research and carbon markets abroad offer a menu of competing carbon costs. Moving a step further afield, the existence value of the forest (the value humans derive from knowing that the forest has not been decimated) has received less academic attention and may never be priced in a market, requiring reliance on a scarcer and wider range of scholarly estimates. Thus, while using adjusted-GDP indicators implies risk of uncertainty, that risk ebbs with prioritization of well-researched and near-market variables.

While composite indicators avoid the unresolved debates of appropriate pricing methodologies, they must determine a weight to place on each component variable. Indeed, such a choice is arguably more arbitrary than imputing prices from market values, given that it is unclear what empirical rationale could be used to justify, for example, the relative importance of health versus education versus income. To avoid making such controversial decisions, the authors of most composite indicators tend to assign equal weights to all included variables. However, the choice of equal weights,

while more palatable, is no less arbitrary. To use the example above, declaring commuting time and deforestation to be of exactly equivalent importance for human welfare seems specious.

In sum, while the methodological uncertainty of adjusted-GDP measures can be attenuated via appropriate variable choice and further research, the structural design of composite indicators leaves few options for mitigating uncertainty. By assessing the worth of all variables with a common logic (dollars imputed from market values), the adjusted-GDP framework allows for an internally consistent assessment of the relative importance of distinct factors in contributing to overall welfare. By contrast, composite measures assess each factor on a separate scale, providing no cross-cutting rubric for assessing relative welfare importance. As such, of the three frameworks, overcoming accuracy challenges seems most plausible for adjusted-GDP measures.

Impact

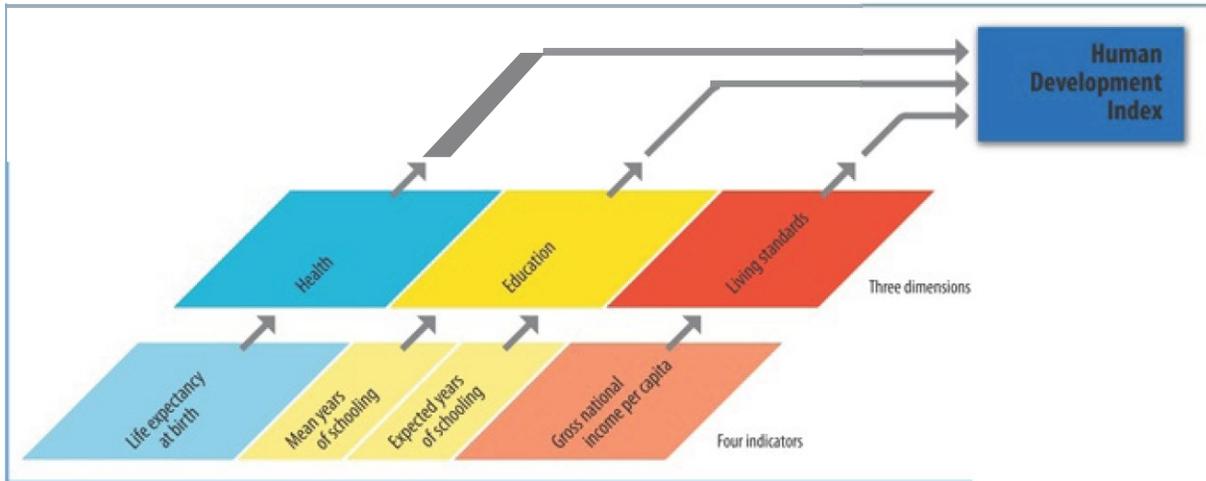
Among the three types of indicators, the subjective framework would likely have the least policy salience within the U.S. context. The multiple accuracy issues noted above would likely cause policymakers to discount the worth of the measure. As such, a "happiness" indicator should not be expected to sway the policymaking narrative, nor should federal agencies be expected to adopt the measure as a formal policy benchmark.

Among the remaining two categories, some authors of composite indexes have argued for their superiority in changing the policy narrative, given their unique ability to actually place GDP in a subordinate role as one of several component variables. The Human Development Index, for example, incorporates GDP alongside health and education, reducing the traditional measure's role to one-third of the definition of progress. By contrast, they argue, adjusted-GDP measures further validate GDP's primary role by starting with and merely augmenting the headline indicator. Such an alteration, they believe, is too marginal to really shift the public discourse on what constitutes policymaking success.

However, the similarity of adjusted-GDP measures to GDP itself may be precisely the strength of the adjusted framework. Proponents of indicators such as Maryland's Genuine Progress Indicator argue that they have successfully influenced the policy debate by expressing welfare in the same unit of measurement that GDP uses to express economic activity. Consider, for example, that opponents of drilling for oil in the

Components of the Human Development Index

The HOI—three dimensions and four indicators



M>rdheond<••00 Pesefl d in<r...t.gure folkM <henevmetlodolog(<Sd<f>nro onoo•ll
S:...l(t:HOAO.

Figure 3: The UN's Human Development Index assigns equal weights to health, education, and GDP in ranking countries' aggregate performance.

Arctic National Wildlife Refuge want to contest an assertion that doing so will boost GDP by \$350 million. Which is the better counterargument: 'the ensuing loss of natural capital will actually cost a net \$147 million,' or 'the project will lower our welfare index from a 0.7 to a 0.63'? The analog of dollars provides a more readily comprehensible, and thus more powerful, argument in U.S. policy decisions.

Given these competing advantages of the composite and adjusted-GDP frameworks, which is likely to have the bigger impact as a narrative-swaying metric and policymaking tool? A supplemental indicator's mode of expression seems more critical to this question than its means of construction. That is, while design of composite indexes can appropriately convey the message that GDP is but one subcomponent of overall welfare, this construction-derived message will be relatively subtle and short-lived. While technocrats and policymakers are likely to know and care about the subcomponents of any new indicator, media outlets and the politically-active public are likely to be more influenced by the conclusions of an indicator (e.g. welfare would decline by \$147 million) than its components. Similarly, while the design of a composite indicator might capture policymaking interest in the first year or so following its release, the weakness of a non-analogous unit of measurement will persist indefinitely. As such, an adjusted-GDP framework holds greater and longer-term policy change potential.

Some analysts who acknowledge the policy change potency of adjusted-GDP measures fear that the resulting policy changes will be the wrong ones. This argument hinges on the fact that the framework assigns prices to nonmarket environmental variables. Some environmental advocates have argued that forests, clean air, and other environmental resources are "invaluable," and thus "priceless." Driving this viewpoint seems to be the fear that picking a price, whatever the methodology, will facilitate the destruction of natural resources because the case can then be made that said destruction would enable economic production of a greater dollar amount. The common retort from many environmental economists is that, in the realm of policymaking, not assigning a dollar value means an implicit price of zero. In assessing the worth of a proposed suburban development in virgin forestland, policymakers using cost-benefit analysis would conclude that the development constitutes pure progress unless the forest is attributed monetary value. In the same way, welfare indicators require the imputation of price to begin valuing environmental protection on par with economic factors. By explicitly counting the worth of environmental resources, adjusted-GDP measures are far more likely to reverse policies that destroy the environment, not ones that seek its protection.

Feasibility

How much would each type of indicator cost the federal government? While aggregation of existing data implies some financial outlays (e.g. compensation for staff time), the cost associated with new indicators is largely a function of the degree to which they require the gathering of new data (e.g. survey design, administration, follow-up, and analysis). Subjective measures would definitively require new data collection, given the lack of official U.S. surveying on subjective welfare. Complex subjective measures, such as Bhutan's Gross National Happiness, would require substantial costs for new, standalone surveys, while simpler subjective measures could plausibly be achieved for a lower-cost option of adding several questions to existing surveys, such as the U.S. Census Bureau's monthly Current Population Survey.

The cost of adjusted-GDP and composite metrics, meanwhile, would depend on the number and type of variables included. Costs would be negligible if variables were restricted to those for which reliable data already exists. The 2010 launch of Maryland's Genuine Progress Indicator, for example,

had zero budgetary impact for the state. After an interagency group established the computing methodology and identified existing data sources, a single state official needed only two weeks to solicit and aggregate the 26 variables that became GPI. However, the relevant data is not so readily available at the federal level, where calculating the same 26 variables would likely require more thorough study and new data gathering initiatives. According to statistical experts interviewed, creating a federal-level GPI with a similarly expansive list of 26 variables would likely cost the federal government approximately \$50 million. (A composite indicator of equivalent breadth would likely cost slightly less, given the avoided costs of determining appropriate price imputations.) That sum is just under half of the Bureau of Economic Analysis's budget request for fiscal year 2012.²⁸ While the amount is significant, revamping the country's national accounts at a cost of 0.60% of next year's Department of Commerce budget falls within the scope of feasibility.²⁹

More than cost, the primary feasibility-related threats to indicator adoption are the aforementioned Obstacles 2 and 3 that have hindered past reform efforts: institutional inertia and opposing political interests.

The challenge of institutional inertia rises with the degree to which new indicators require interagency coordination. Subjective indicators present less of a challenge, since the requisite survey could likely be administered and analyzed by a single agency. Composite and adjusted indicators, by contrast, could require the cooperation of a couple dozen data-gathering agencies if variables span an array of environmental, social, and economic domains. (See the Institutions section for recommendations to enable such cooperation.)

Some analysts argue that adjusted-GDP measures pose a particularly high bureaucratic hurdle, given that they purport to alter the standardized National Income and Product Accounts (NIPA) that have been in place for over seven decades. Attempting to change that which has resisted change for so long, they argue, is senseless. Yet, this characterization of the NIPA is misleading. As stated, the BEA makes significant adjustments to the NIPA methodology at least once every five years. In addition, the BEA has consistently sought to augment the NIPA with satellite accounts, such as the current effort to better account for R&D within GDP.³⁰ The adoption of new adjusted measures would accelerate the pace of augmentation, but would not radically depart from the BEA's own historic initiatives.

Beyond institutional challenges, how might the indicator framework influence the likelihood of overcoming political obstacles? Subjective indicators may be politically objectionable, given their distance from policy impact. Legislators would be rightly fearful of adopting subjective indicators that do not correlate with demonstrable improvement in policy objectives (e.g. educational attainment). Even high-performing Congressional representatives could be politically injured by likely reports that happiness levels did not increase during their term.

Table A. Comparison of Average Annual Real GDP Growth Rates

[Pereola]

	2010 vintage		2007 vintage	
	Adjusted real GDP ¹	Real GDP ²	Unadjusted real GDP ³	Adjusted real GDP ⁴
1959-1973	4.20	4.33	4.20	4.33
1974-1994	3.02	3.03	3.02	3.03
1974-1994	3.76	3.93	3.53	3.72
1995-2001	2.75	2.87	n.a.	n.a.
1005-2001	3.37	3.39	3.33	3.40
1005-2001	3.32	3.39	n.a.	n.a.

Figure 4: The BEA recently adjusted the way it measures R&D within GDP, the latest in a long series of BEA modifications to the NIPAs.

While adjusted-GDP measures, if well-designed, would avoid this concern of policy irrelevance, the need to determine prices for politically-charged non-market variables (e.g. the social cost of carbon) poses potential for significant conflict. Two viable solutions could help in circumventing such conflict. First, assigning the task of imputing prices to a bipartisan technical commission would help depoliticize the process (see Section 5.4). Second, by incorporating numerous variables that appeal to different political constituencies, an adjusted-GDP measure brings the opportunity to build a wide coalition with sufficient common interests to overcome political opposition. (See Interests section for details.)

Even so, composite indexes are more politically promising than adjusted measures, given that the opportunity for such coalition-building remains, while some of the controversy of imputed prices does not. Unlike adjusted-GDP measures, composite indicators hold the option of using equal weights across all variables, making no attempt to determine the degree to which education versus carbon emissions affects overall welfare. While this approach remains methodologically spurious, it would likely obviate some political spats over the relative welfare impact of particularly fractious variables.

Summary

Of the three types of indicators, the subjective category ranks lowest on all three overall criteria. Subjective measures categorically lack accuracy because their aim is less policy-driven than our own. The loose connection to policy would also tend to limit the impact of such measures in shaping policy debates. In addition, the

political distrust of such measures limits the feasibility of the subjective framework.

The remaining two categories offer similar strengths and shortcomings. Perhaps the largest shortcoming for both lies in the uncertainty of pricing or weighting subcomponent variables. For adjusted-GDP measures, this challenge decreases with variables' proximity to a market and degree of academic consensus. For composite indexes, this challenge appears largely irreconcilable, even in the arbitrary scenario of equal weighting.

While accuracy challenges are more mitigable under the adjusted-GDP framework, the composite framework edges out the adjusted one as the more feasible type to implement. This greater feasibility stems primarily

TABLE 1: SUMMARY OF INDICATOR FRAMEWORKS

	Accuracy	Impact	Feasibility
Subjective	Happiness indexes are not designed to fully or exclusively capture the policy-relevant component of well-being.	The notion of aggregate happiness is not widely trusted, hindering its policy impact.	The bureaucratic cost of indicator assembly is low, though political hurdles to adoption would likely be high.
Composite	Assigning weights to component variables is a generally arbitrary process that challenges the accuracy of the overall indicator.	The indicator structure can place GDP in an appropriately subordinate role, but the unit of measurement (a decimal) lacks narrative potency.	Assigning equal weights to component variables could reduce political obstacles, while bureaucratic and financial costs remain manageable.
Adjusted-GDP	Determining prices for non-market variables is difficult, though the uncertainty declines for near-market and well-researched variables.	The expression of values in dollars offers the potential to contend with GDP in swaying policy debates.	The need to determine prices for non-market variables presents political obstacles that, while significant, are surmountable.

from the simplicity with which composite indicators reduce each variable to a scaled index. Doing so largely avoids the politically sensitive pricing debates that an adjusted indicator seems likely to spur.

However, when it comes to impact, adjusted-GDP measures pose greater potential for shifting the policymaking narrative than composite measures. The advantage lies primarily in that the adjusted measures' unit of measurement remains in GDP-analogous dollars. While the design of composite indicators can more clearly underscore the subordinate role of *GDP*, the ability of adjusted-GDP measures to express welfare in dollarized losses or gains signifies a wider and longer-lasting impact on the public discourse that guides policymaking priorities.

As such, an adjusted-GDP measure offers a tradeoff: in comparison to the composite framework, it offers a more powerful policy tool with greater accuracy potential, but threatens to spark greater political controversy. There is potential for circumventing such political obstacles, in addition to confronting methodological challenges, by delegating the design of an adjusted-GDP indicator to a bipartisan technical commission (see Section 5.4). In contrast, there are no evident means of overcoming the impact or accuracy shortcomings of composite indicators. As such, this report ranks the adjusted framework as preferable to the composite form.

Recommendation: choose adjusted-GDP as the indicator framework best suited to parallel GDP.

5.3 Variable Selection

Four new indicators should be created to assess four precise definitions of national progress: G2 for current prosperity, G3 for sustainable prosperity, G4 for current well-being, and G5 for sustainable well-being. Each new indicator includes a tailored list of GDP adjustments. Standing alongside GDP (which would become G1), this series of 21st-century indicators offers the opportunity to redefine policymaking success.

Within the adjusted-GDP framework, one must decide which welfare-relevant variables to include. Included below are the 26 most common variable types, culled from the two dozen existing alternative indicators. Each of these variables is impacted by policy, and in turn, impacts welfare.

VARIABLES

Consumption:	the standard aggregate consumption measure used in GDP (Personal Consumption Expenditures)
Inequality:	discounting for income and/or wealth inequality
Household:	assigning a value to household labor
Informal:	assigning a value to informal economy activity (beyond household labor)
Volunteer:	assigning a value to the free services performed by volunteers
NGO:	assigning a value to the free services provided by NGOs
GovtFree:	assigning a value to roads and other government services consumed for free
Crime:	assigning costs to the relative incidence of crime
Finance:	removing or discounting financial sector activity included in GDP (given that it's an intermediate good)
FamilySep:	discounting doubly-purchased items due to family separation
Accidents:	assigning costs to motor vehicle, industrial, and other accidents
Underemp:	assigning a cost to unemployment and underemployment rates
ConsDebt:	discounting for unsustainable consumer debt levels
Pollution:	assigning a cost to air, water, noise, or other sorts of pollution
Biomes:	assigning a cost to the destruction of distinct biomes
Subsoil:	assigning a cost to nonrenewable energy sources used
Emissions:	assigning a cost to ozone, carbon, and other emissions with long-term consequences
Health:	assigning a value to the stock of health available to the economy
Education:	assigning a value to the stock of education
NatlDebt:	accounting for unsustainable national debt/surplus levels
Social:	assigning a value to social capital
Resilience:	assigning a value to the capacity to recover from disaster
Knowledge:	assigning a value to intellectual capital
Netinvest:	accounting for net capital investments
Leisure:	assigning a value to leisure time
Commuting:	assigning a cost for commuting time

To avoid skewed welfare assessments, the process of picking variables from this list should be neither arbitrary nor subjective. Two steps are essential but often overlooked prerequisites for choosing variables:

1. Naming the particular welfare function that the indicator seeks to measure, so as to determine the range of variables that count as valid inputs. For example, the degree of leisure time would be a valid input for assessing general welfare, but not for assessing economic welfare.
2. Naming the criteria used to assess the worth of including each valid variable. For example, will variables be included to the extent that data is available, to the extent that their calculation is

methodologically sound, or both? (See section 5.4 for a listing of such criteria.)

To address the first step, alternative indicators too often name their objective functions in imprecise terms: "human development," "genuine progress," "well-being." Without more concrete definitions of what constitutes success or failure, indicators risk including variables that cannot be defended as inputs for the desired output (or excluding other variables that would indeed be valid inputs), resulting in a muddled aggregate number with unclear implications.

Among the several dozen existing indicators, there seem to be two main cleavages in the implicit objective functions: current vs. sustainable welfare, and economic vs. general welfare. Such distinctions are important. Sustainable welfare should include assessments of changes in resource "stocks" that indicate future potential (accounting for depreciation of the capital stock, depletion of subsoil mineral stocks, etc.). To include such variables in an assessment of current welfare would conflate purposes. Similarly, indicators should offer a coherent answer to the question, "Are social and environmental welfare considered only insofar as they contribute to economic welfare, or are they to be treated as ends in and of themselves?" For example, if the existence of a particular species of salamander makes no ostensible contribution to economic activity, the extinction of the species should probably not be discounted from a measure of economic welfare, though its existence value could well be subtracted from a metric of overall welfare.

Table 2 presents a two-by-two matrix based on these two primary distinctions, resulting in four quadrants with four different welfare objectives: current economic welfare, sustainable economic welfare, current general welfare, and sustainable general welfare. The assumed welfare function is stated for each quadrant. The chart also considers which welfare and policy-relevant variables (from the list above) would count as a valid contribution to the particular type of welfare being measured. For each qualifying variable, the adjustments that would be performed are listed in the relevant quadrants. Here is an overall summary of the adjustments made in each quadrant:

- **Current Economic Welfare:** This type of indicator comes closest to what GDP attempts to measure. Indeed, it starts with personal consumption expenditures, a critical component of GDP. Yet, it makes three types of adjustments to provide a more accurate

gauge of an average person's felt economic experience. First, the measure adds in goods and services that contribute to economic welfare but are not counted as consumption because they are free of charge, such as those provided in the home (family dinners), by the government (toll-free roads), or by NGOs (soup kitchens). Second, the measure subtracts goods and services that count as consumption but do not contribute to economic welfare, such as damaged property, stolen items, or defensive expenditures (e.g. locks, alarm systems) that rise with crime rates. Third, this indicator adjusts for the fact that some types of consumption typically offer lesser degrees of satisfaction. Most economists *agree*, for example, that an extra \$100 spent by a low-wage worker contributes greater satisfaction than an extra \$100 spent by a millionaire. As such, the measure adjusts aggregate consumption for inequality so that increasing income/wealth in the top quintile does not count equally toward progress as increasing wealth in the bottom quintile. Similarly, consumption that must be financed by an onerous degree of debt is thought to contribute lesser satisfaction than debt-free spending. As such, the measure discounts consumption financed by debt levels above a given debt sustainability threshold.

- **Sustainable Economic Welfare:** This type of indicator makes all of the adjustments under current economic welfare, but accounts for changes in various stocks necessary for future economic welfare. In the same way that Net Domestic Product adds investment in the physical capital stock but subtracts capital depreciation, this category accounts for net marginal changes in stocks critical for sustained economic activity: physical capital (e.g. machinery), natural capital (e.g. subsoil minerals), human capital (e.g. higher education levels), financial capital (e.g. national debt/surplus), etc. This type of measure also subtracts the net present value of future inhibited economic activity caused by current-year factors such as crime (reduced future sales) and carbon emissions (future damages from climate change).
- **Current General Welfare:** This indicator includes all the adjustments made to current economic welfare, but also values social welfare and environmental welfare insofar as they contribute to current well-being. As such, the measure adds positive social variables (e.g. leisure time), subtracts negative social variables (e.g. commuting time), adds positive environmental variables (e.g. the existence value of a forest), and subtracts negative environmental variables (e.g. smog). This type of indicator probably comes closest to measuring policy-relevant happiness, given that humans tend to depend on current economic, social, and environmental inputs for happiness, but tend to myopically discount prospects for future welfare in estimations of well-being (few people consider the extent of subsoil mineral supplies in assessing their welfare). As such, this quadrant offers a more policy-relevant companion to subjective welfare measures.

-
- **Sustainable General Welfare:** This broadest category includes all of the adjustments in the three prior categories. It also includes adjustments for the net present value of current-year, policy-relevant factors that affect future years' social and environmental welfare. For example, consider the destruction of an acre of forest. Sustainable economic welfare already captures the natural capital loss of timber and ecosystem services, while current general welfare captures the loss of that forest's existence value for the current year. Sustainable general welfare would additionally capture the net present value of the lost existence value for all future years.

Each of these four categories presents a type of welfare relevant to policymaking. The best indicator for shaping the narrative depends on the narrative being shaped. If the city of Baltimore, for example, wanted to assess this year's quality-of-life impact of a sprawl-reducing urban renewal project, the best indicator would include current social and environmental factors (e.g. commuting time, smog), but probably not changes in assets (e.g. debt or depreciation incurred for the project). Current general welfare would suit best. By contrast, if the city wanted to assess the contribution of the project to sustainable economic growth, the appropriate gauge would subtract asset depletion (e.g. additional debt) and add asset creation (e.g. increased social capital), but not account for non-economic variables like commuting time. Sustainable economic welfare would prove the best indicator in this case.

Given such distinct indicator needs, four new national indicators-G2, G3, G4, and GS-should be created to assess four definitions of progress: G2 for current economic welfare, G3 for sustainable economic welfare, G4 for current general welfare, and GS for sustainable general welfare. This series of welfare benchmarks would stand alongside GDP, which would become *GI*, in the same way that the federal government uses U1 through U6 as complementary measures of unemployment.

The resulting indicators would look as such, listed in the order in which they expand on *GDP*'s narrow original purpose:

- G1: Gross Domestic Product
- G2: Current Prosperity
- G3: Sustainable Prosperity
- G4: Current Well-being
- GS: Sustainable Well-being

Recommendation: four new national indicators-G2, G3, G4, and GS-should be created to provide four supplementary measures of progress

TABLE 2: 21 51-CENTURY WELFARE INDICATORS

	Economic Welfare	General Welfare
Current	<p>G2: Current Prosperity</p> <p><i>Measuring this years purchasing power satisfaction</i></p> <ul style="list-style-type: none"> • Start with pe-sonal consumption expenditures(+) • Discount for income inequality(-) • Discount for debt-financed consumption above a critical threshold (-) • Add the value of unpaid goods/services from household labor and volunteer labor(+) • Add the value of unpaid goods/services from the non-profit sector and the public sector(+) • Add the value of informal sector goods/services(+) • Subtract the loss of stolen goods(-) • Subtract property damage from accidents(-) • Subtract defensive expenditures (e.g. locks)(-) 	<p>G4: Current Well-being</p> <p><i>Measuring this years satisfaction/rom purchasing power, societal health, and the environment</i></p> <ul style="list-style-type: none"> • Start with G2 • Add the value of willful leisure time(+) • Add the utility value of community cohesion (+) • Add the spillover gains of preventative health (+) • Subtract the utility loss of environmental degradation (e.g. smog) (-) • Subtract the utility loss of corrrnuting time (-) • Subtract the psychological/physical injury of crime or accidents(-) • Subtract/add the existence value change of biome loss/gain(+/-)
	<p>G3: Sustainable Prosperity</p> <p><i>Measuring this years purchasing power satisfaction, while accounting for future potential satisfaction</i></p> <ul style="list-style-type: none"> • Start with G2 • Add/subtract changes in stocks of physical capital, human capital(health/education), financial capital, intellectual capital (R&D), and social capital(+/-) • Add/subtract changes in stocks of natural capital: land/air/water rendered unusable, biome loss/gain, and net changes in stocks of subsoil minerals(+/-) • Add/subtract changes in entrepreneurship (+/-) • Subtract future damage (net present value) from current-year carbon emissions(-) • Subtract potential business losses from crime (-) • Subtract productive capacity loss from accidents(-) 	<p>G5: Sustainable Well-being</p> <p><i>Measuring this years satisfaction from purchasing power, societal health, and the environment, while accounting for future potentia/satisfaction</i></p> <ul style="list-style-type: none"> • Start with G2 + G3 + G4 • Add/subtract the future effect (net present value) of current changes in biome existence value (+/-) • Add/subtract the future utility effect (NPV) of current changes in community cohesion(+/-) • Add/subtract the future spillover effect (NPV) of current changes in preventative health(+/-) • Add/subtract the future psychological/physical damage (NPV) of current changes in crime or accidents(+/-)
Sustainable		

Some analysts may argue that keeping GDP as G1 and placing the supplemental indicators in a succeeding series will undermine the narrative or policymaking importance of the new indicators. The fear is that press and policy debates would continue focusing exclusively on G1, while G2 through G5 would remain largely academic ideas. Yet, such has not been the case with the unemployment measures that serve as the model for this series. Since the Bureau of Labor Statistics introduced the broader unemployment measures U4-U6 in 1995,³¹ the new measures have regularly captured headlines and influenced policy debates, particularly when the official unemployment measure (U3) falls out of sync with the figure capturing full under-and-unemployment (U6). Such usage is not confined to the halls of think-tanks and Congress-while campaigning in Nevada this February, Presidential hopeful Mitt Romney criticized President Obama for taking credit when U3 dropped to 8.3%, arguing that "the real employment rate is over 15 percent."³² Romney was citing U6. If broader unemployment indicators are being publicly employed as a rubric of policymaking success, broader welfare indicators can fill the same purpose. It is not difficult to imagine the symbolic power of headlines like "GDP Grows, Prosperity Falls," or blog-disseminated graphs that reveal a growing gap between G1 and G3, sparking the question of why current growth is coming at the expense of future growth. As with U6, the incorporation of these new indicators into media and policymaking debates could catalyze a shift to a broadened definition of progress.

5.4 Overcoming Obstacle 1: Indicator Complexities

While the indicator design recommendations thus far have sought to limit methodological inaccuracy, the choice of the adjusted-GDP framework and broad recommended parameters of G2-G5 leave much room for methodological uncertainty. Diminishing that uncertainty will require carefully choosing the particular variables and pricing formulae to include within each of the four new indicators. Congress cannot be expected to make such detailed decisions. Members and staff have neither the technical means nor the time to devote to determining and justifying all the specifications of new indicators. Thus, Congress should seek to delegate these tasks to a "21st Century National Indicators Commission": comprising a range of bipartisan economic and statistical experts and modeled after the most successful bipartisan commissions, including the Base Closure and Realignment Commission.

In general, commissions can be categorized as policy commissions, investigatory commissions, or commemorative commissions. The 21st Century National Indicators Commission would be a policy commission tasked with performing specific duties including the following:

- Determining the optimal composition of each indicator
- Determining the optimal pricing formulae for assigning value to each included variable
- Identifying opportunities for making available the requisite data
- Identifying opportunities for enabling Executive agencies to compute, aggregate, and report the series

The commission would be bipartisan in membership. The President, in consultation with the leadership of both parties in Congress, would be empowered to appoint four commissioners based on expertise in relevant fields. The President could moreover appoint eight additional qualified commissioners, no more than four of whom can be of the same party.

The commission would have up to two years to submit an "Indicator Development Plan" to the President, who could either approve the plan without amendment or return the plan to the commission with recommendations for revision. If the President accepts the indicator development plan, the Bureau of Economic Analysis would, in consultation with the Interagency Council on Statistical Policy, begin taking steps to implement the plan.

A significant share of the commission's work would be to identify what variables should be included for each new measure, G2 through GS. Identifying a specific welfare objective for each of these new indicators (see previous section) helps to narrow the field of valid variables. Yet, further narrowing of variables within any well-defined indicator would be necessary based not only on methodological certainty, but funding requirements, policy relevance, and other considerations. Thus, a second critical step of variable choice is identifying ex-ante criteria to assess the respective merits of each possible variable. Doing so is important to directly address three oft-expressed critiques of the composition of alternative indicators:

- 1) "The choice over an indicator's variables lacks an objective rationale, clouding the meaning of the indicator itself." Indicators should justify the choice of variables based on evidence or outside expert opinion that makes clear the variables' importance for the indicator's objective function. Without such substantiation, indicator authors may misjudge the relative importance of competing variables and include lesser variables at the expense of more relevant ones.
- 2) "The range of included variables is arbitrarily constrained by the availability of data." While data constraints are significant considerations in deciding which variables to include, the cost of gathering the required data may be outweighed by the relative importance of the variable as a welfare input. The outcome of this cost-benefit analysis cannot be known, however, if benefits are not weighed alongside costs.
- 3) "Indicators can be reverse-engineered to convey whatever message their authors desire." Those pursuing a narrow political agenda could certainly cherry-pick variables to create an indicator likely to advance their agenda. To avoid this outcome, or even this critique, it is necessary to set broadly-shared, logical, and transparent criteria before picking the variables, letting the criteria rather than partisan agendas guide the choice.

This report proposes nine such criteria, three each for the three guiding principles of impact, accuracy, and feasibility (see below for explanations of the criteria). While the commission would need to develop its own criteria and use its own chart, Table 3 may serve as a starting point. In whatever rubric is ultimately used, the chosen criteria should at minimum address each of the overarching principles guiding sound variable choice: impact, accuracy, and feasibility.

TABLE 3: MATRIX FOR APPROPRIATE VARIABLE CHOICE

	Impact			Accuracy			Feasibility			S U M
	Dollar Effect	Narrative Strength	Widely Appropriate	Price Certitude	Quantity Certitude	Policy Relevance	Politically Tenable	Affordable Data	Regular Data	
Weight										
Consumption										
Inequality										
Household										
Informal										
Volunteer										
NGO										
GovtFree										
Crime										
Finance										
FamilySep										
Accidents										
Underemp										
ConsDebt										
Pollution										
Biomes										
Subsoil										
Emissions										
Health										
Education										
NatlDebt										
Social										
Resilience										
Knowledge										
NetInvest										
Leisure										
Commuting										

CRITERIA:

Dollar Effect: To what extent is GDP likely to change (in dollars) by including the variable?

Narrative Strength: To what extent would the variable assist in explaining why new measures of progress are needed?

Widely Appropriate: Is this variable readily computable in states, districts, or other countries? How likely are bodies like the UN or IMF to one day recommend calculation of this variable?

Pricing Certitude: What degree of academic consensus exists over the appropriate pricing methodology for the variable?

Quantity Certitude: How reliable do statisticians believe the underlying quantity data to be (e.g. time-use hours, subsoil mineral supplies, etc.)?

Policy Relevance: How direct and exclusive is the causal link between policy and this variable?

Politically Tenable: How likely is it that indicator legislation passes and gains adoption with this variable included?

Affordable Data: To what extent would data gathering and analysis impose new costs?

Regular Data: How regularly and reliably can data on the variable be gathered?



6

INSTITUTIONS

Attaining the requisite operational capacity to effectively produce new indicators in the executive branch

With more than 20 diverse statistical agencies and substantive data-sharing arrangements with state and local authorities, the federal government is well-equipped to begin developing the range of indicators prescribed in the previous section. Yet, given the multijurisdictional nature of the new set of indicators and the absence of an integrated US national statistical agency, there are considerable coordination and standardization challenges inherent in this undertaking.

This section will outline the specific roles that would be involved in producing the prospective G2-G5 indicators, describe the government actors that may take part, and explain how Congress, the Executive, and the commission prescribed in the previous section can play important roles in bringing this vision of supplemental national indicators to fruition.

CORE RECOMMENDATIONS: (1) The Bureau of Economic Analysis (BEA) should play a primary role as the only federal agency capable of aggregating new multijurisdictional indicators; (2) the Interagency Council on Statistical Quality, a multi-agency steering committee housed in the Office of Management and Budget (OMB) can help translate the Commission's findings, if accepted, into interagency standards and rules and assist with cross-departmental coordination; (3) if necessary, Congress should act on the Commission's findings by appropriating funding for new data collection and/or removing barriers to interagency information-sharing.

6.1 Roles: Computation, Aggregation, and Reporting

Computation

The first general task in the production of new indicators--computation--refers to the processes of collecting data and calculating the various sub-indicators that will form component parts of the overall indicator. This entails compiling the relevant data and making computations in accordance with the pricing methodologies stipulated by the commission. No single agency can reasonably be expected to compute all of the included sub-indicators. Indeed, it is expected that a wide range of agencies including but not limited to those listed in table 4 would take part in the computation process.

TABLE 4: COMPUTING ROLES

Computing Subindicators	
Agency	Potential Sub-indicator Area
Agency for Healthcare Research and Quality	Healthcare
Bureau of Economic Analysis (BEA)	National income, investment, nonmarket activity
Bureau of Justice Statistics (BJS)	Crime, incarceration
Bureau of Labor Statistics (BLS)	unemployment, consumer prices, average hourly earnings
Bureau of Transportation Statistics (BTS)	Commuting, Energy
Census Bureau	Income, demographics, business, education, health, well-being, nonmarket activity, household income
Centers for Disease Control and Prevention (CDC)	Health outcomes
Centers for Medicare and Medicaid Services (CMS)	Healthcare
Economic Research Service (ERS)	Agriculture, land
Energy Information Administration	Energy efficiency, air pollution
Environmental Protection Agency	Air and water quality, emissions
Federal Deposit Insurance Corporation	Banking
Federal Reserve	Household debt
Health Care Financing Administration (HCFA)	Healthcare
Internal Revenue Service (IRS)	Taxation
National Agricultural Statistics Service (NASS)	Agriculture, land
National Center for Education Statistics (NCES)	Education quality and attainment
National Center for Transportation Statistics and Analysis (NTSA)	Transportation
National Climatic Data Center (NCDC) - NOAA	Climate change
National County Administration Utilization Survey (NCAS)	Activity net, aging, disabilities

Aggregation

Unlike computation, a single agency would likely need to be responsible for the second task, aggregation of the various indicator components according to frameworks set by the commission. There are several reasons why BEA should play this role. First, BEA is widely viewed as having the highest level of technical capacity related to aggregate indicators. It compiles the NIPAs, which are themselves aggregates, and has periodically spearheaded the development of new prototype "satellite accounts." Second, BEA is widely perceived as politically insulated and therefore highly credible in its statistical work. This is essential for ensuring the indicators are salient in policy debates and public discourse. Third, as the G2-GS indicators would form a set of related indicators (existing GDP would become G1), it stands to reason that this set should be aggregated in the institution where GDP is currently compiled. Fourth, as an agency that largely relies on other government institutions for data collection, it has strong existing data-sharing linkages to offices such as the Census Bureau that would be involved in computation.

It is important to note that, in spite of its unique fit for the role of aggregator, there are clear reasons why BEA itself should not determine the indicator specifications in lieu of the commission. BEA is justifiably concerned about undertaking the design of initiatives that, owing to their inherent subjectivity, might undercut its reputation for independence. BEA is also understandably concerned with taking on greater responsibilities than its budget and staff can handle. Given that major statistical agencies including BEA lack direct relationships with Congressional appropriators (funding needs are managed by departmental-level officials), it would be less likely than a congressionally-ordered commission to effectively shepherd the indicator design and development process.

Reporting

The final task, reporting the indicator once aggregated, can fall on multiple agencies including the BEA, which currently reports GDP through its quarterly Survey of Current Business. The Office of Management and Budget (OMB) is also an opportune place for reporting, given that the Government Performance and Results Act (GPRA) and 2010 GPRA Modernization Act require the office to establish outcome-oriented whole-of-government goals every four years.³³ The new indicators prescribed in this report could serve as tools for outcome-based interagency performance measurement. Also in the executive branch, the President's Council of Economic Advisers could potentially report new G-series data as part of its annual Economic Report of the President, which reports GDP as part of its effort to survey the economic landscape.

6.2 Overcoming Obstacle 2: Institutional Inertia

There are challenges inherent in producing new supplemental national accounts, even assuming the commission reaches agreement on methodological standards that are approved by the President. The most central of these questions is how to ensure that numerous discrete statistical agencies provide standardized data in a timely fashion. Unlike many industrialized country governments, the US lacks a single national statistical office. While this arrangement is advantageous insofar as it enhances accountability and makes the nation's statistical apparatus less vulnerable to capricious budget cuts in times of austerity, it nonetheless makes standardization and coordination more difficult. This difficulty may manifest in at least three ways: (1) agencies, citing a lack of staff or funding, may fail to produce crucial sub-indicators, (2) sub-indicators may not conform to the standards set forth by the commission, (3) legal barriers may prevent computing agencies from sharing information with one another or with BEA.

In addition to setting standards for indicator methodology, the commission designated in the attached legislation is charged with investigating new data requirements, funding needs, and legal barriers to information sharing. It is moreover charged with reporting these needs to the President. While the President should act to make as many of the relevant changes as possible through administrative action, it would ultimately be up to the Congress to take necessary follow-up actions, including appropriations. Doing so would help ensure that agencies are compelled to produce the information in a timely fashion (in order to receive funding) and are capable of producing the information (both in terms of resources and regulations).

As the aggregating agency, BEA would naturally play a major role in coordinating the work of the computing agencies; nonetheless, the President may designate an interagency authority to formally manage some aspects of the process. The Interagency Council on Statistical Policy (ICSP), established in 1995 and housed in the Office of Management and Budget (OMB), would likely be effective in such a role. The Council, whose members include the heads of all the principal statistical agencies, typically coordinates cross-jurisdictional statistical work and advises the White House on statistical matters related to the federal budget. As with the development of GDP and related NIPAs, it may take years for the right information infrastructure for supplementary accounts to emerge. ICSP can help steer and accelerate this process. Ultimately, as the Council is comprised solely of statistical officers but bears the White House imprimatur, it may help manage two key objectives that often stand in tension: expediency and independence.



7

INTERESTS

Building the political coalition needed to enact reform

Important stakeholders in the US political system have vested interests in maintaining the status quo in national accounting. The Clinton Administration's failure to overcome coal industry pressure to implement its plans for "greening" the NIPAs in the early 1990s illustrates this fact. International precedents also suggest that incumbent politicians may oppose alternative indicators if they diminish the extent of economic progress or even show negative net growth.³⁴ Yet there is evidence to suggest that such opposition is surmountable: The section suggests the potential composition of a reform coalition and reasserts the political benefits of delegation to an expert commission.

CORE RECOMMENDATIONS: (1) keep indicator parameters intentionally broad in order to ensure the development of a diverse political coalition, (2) delegation of the most contentious political questions to the expert commission, and (3) identify a strategic time for introducing reform.

7.1 Overcoming Obstacle 3: Interest Group Opposition

Strength in Numbers: Building a large and diverse coalition for reform

The clearest reason for the failure of the aforementioned Clinton Administration case was the hollowness of the coalition for reform. Only one constituency, environmentalists, had a sense of ownership over the issue of green accounting, whereas multiple well-funded interest groups including power generation and mining stood in opposition. Yet, it is possible to imagine an alternate scenario in which multiple

reforms to national accounting-including the addition of sub-indicators on R&D and family breakdown, as well as the environment-would be simultaneously advanced. That is the proposal inherent in the prospective G2-GS indicators. The electricity and mining interest groups would now be forced to contend with a far larger constituency of potential supporters of reform, including science research firms and some social conservatives. Beating this coalition would require far greater lobbying resources, reducing the opponents' chances of success.

The proposed G2-GS supplemental indicators may appeal to an extraordinarily diverse range of constituencies. These include:

- *Social Conservatives:* In 1993 Reagan Administration official and prominent conservative activist William Bennett penned a Wall Street Journal article that began, "Is our Culture declining? I have tried to quantify the answer to this question with the creation of the Index of Leading Cultural Indicators."³⁵ Bennett's attempt to codify performance on social values was widely supported by social conservatives, but never instituted. While G2-GS do not explicitly aim to track social mores, several of the component variables would indeed accomplish much of the monitoring that Bennett envisioned. By accounting for household services, for example, the new indicators would reverse GDP's tendency to rise as society outsources family values to the market (e.g. choosing babysitters over parental care). By subtracting the various costs of crime, adding the worth of volunteerism, and placing a value on community cohesion, G2-GS would further appeal to those concerned with cultural decline.
- *Innovation-Oriented Firms:* GDP currently does not account for levels of research and development, the degree of observed entrepreneurship, or other indicators of growth-driving innovation. G3 and GS, by contrast, include such variables. Usage of these indicators would encourage more innovation-friendly policies (e.g. public incentives for private R & D), benefitting firms from diverse sectors that invest in cutting-edge designs and products. Doing so could also attract the support of pro-business think tanks and advocacy groups that have been pushing for better innovation indicators for years.
- *Green Businesses:* Given the emergence of Corporate Social Responsibility and "green business" models, industry is no longer monolithically opposed to the pricing of environmental externalities. To the contrary, firms that have gone above and beyond regulatory mandates in reducing greenhouse gas emissions or other pollutants may actually desire national accounts such as G3, G4, and GS to

monitor their progress. For instance, PSEG, a major power-provider, has actually complained that its emissions-reducing investments went unnoticed by government after a range of EPA standards turned out more lenient than anticipated.

- *Economic Progressives:* The U.S.'s falling degree of economic mobility and increasing degree of consumer debt have been longstanding concerns of economic progressives. GDP ignores indebtedness and masks mobility concerns, showing increasing per capita incomes over the last decade even as median incomes declined.³⁶ G2-GS would be the first aggregate economic indicators to reverse these measurement fallacies by accounting for inequality and onerous debt. Doing so would provide economic progressives with a more powerful tool to assert their case for policies to facilitate upward mobility and downward debt levels.
- *Environmentalists:* Since its origins, the environmental movement has argued that current growth should not come at the expense of future growth via ecological destruction. G3 and GS take that same perspective, discounting current economic activity for depreciation of natural capital (e.g. subsoil minerals, forests) and environmental degradation that jeopardizes future growth (e.g. water pollution, carbon emissions). The new measures allow environmentalists to overcome the false dichotomy between environmental protection and economic growth, providing a measure of economic welfare that is finally sustainable.

Delegation to trusted experts

A second reason for the failure of the GDP-greening efforts of the 1990s was the unilateral nature of the action by a center-left administration to change the rules of national accounting in a controversial manner. By contrast, the proposed commission (see section 5.4) would be strictly bipartisan in its composition. It would also have a highly technical mandate to determine the appropriate sub-indicators and the appropriate pricing formulae for each aggregate indicator. As described in previous sections, this would absolve Congress of responsibility for wrestling with the most difficult questions.



PSEG is one of a number of green businesses that may well support measuring environmental externalities.

Overall

New metrics bring the opportunity to surmount the polarization that defines the current U.S. political scene. While the Tea Party and Occupy movements represent profoundly different political ideologies and policy prescriptions, they are in agreement on one point: that the system is broken. On the right, concerns are mounting as headlines tout growth alongside rising national debt and falling social capital, while entrepreneurship, crime, and family breakdown are omitted from the story. On the left, unmitigated carbon emissions and resource depletion have accompanied a precipitous fall in economic mobility even as the prevailing metric declares policymaking success. As new, comprehensive accounting measures, G2-GS hold the power to break through tired left-right divides in a common reassertion of what constitutes progress.





8

CONCLUSION

The Time Is Ripe

Had Simon Kuznets been tasked with designing a system of national accounts in 2012 rather than 1932, he would have likely reached rather different conclusions. Advances in statistics, computing, and data collection have made meaningful estimation of national welfare possible. While Kuznets lamented the inability to measure and incorporate factors such as natural capital and nonmarket labor, we are today free of such limitations.

The attached legislation seeks to enable the United States to harness state-of-the-art thinking regarding indicator development. By delegating methodological questions to technical experts who are trusted across the ideological spectrum, it seeks to remove politics from the business of statistics. By seeking to supplement rather than replace GDP, it seeks to promote sound incremental change. By stipulating the new indicators that emphasize both sustainability and welfare, the legislation seeks to restore US leadership in the science of measuring progress.



DISCUSSION DRAFT
H R
112TH CONGRESS
2D SESSION

To establish a commission to develop alternatives to the gross domestic product for measuring National welfare.

IN THE HOUSE OF REPRESENTATIVES

Mr. CLARKE of Michigan introduced the following bill; which was referred to the Committee on

A BILL

To establish a commission to develop alternatives to the gross domestic product for measuring National welfare.

1 *Be it enacted by the Senate and House of Representa-*
2 *tives of the United States of America in Congress assembled)*

3 SECTION 1. SHORT TITLE.

4 This Act may be cited as the "21st Century GDP
5 Act".

6 SEC. 2. FINDINGS.

7 Congress finds the following:

8 (1) Since the Federal government first insti-
9 tuted national income accounts

1
2 eight decades ago,
3 gross domestic product (referred to in this Act as
4 "GDP") has come to serve as the unofficial barom-
5 eter of not just business cycles, but also of govern-
6 ment performance, living standards, and overall na-
7 tional welfare.

8 (2) GDP was not designed for this all-encom-
9 passing role, and it is especially ill-suited to play
10 such a role today. Simon Kuznets, the economist
11 who Congress tasked with developing GDP's pre-
12 cursor in the 1930s, notably declared "the welfare of
13 a nation can scarcely be inferred from a measure-
14 ment of national income."

15 (3) While GDP is an effective measurement of
16 aggregate economic activity, it fails to take into con-
17 sideration core elements of national welfare in the
18 21st Century, such as social capital, educational at-
19 tainment, environmental quality, health, and eco-
20 nomic mobility.

21 (4) GDP ignores future economic potential be-
22 cause it only counts current-day transactions. **It**
23 overlooks important determinants of growth, such as
24 the degree of entrepreneurship and the depth of con-
25 sumer debt. Further, GDP actually rises with the

1 depletion of physical or natural capital necessary for
2 future prosperity.

3 (5) GDP growth tends to run counter to impor-
4 tant dimensions of national welfare because it omits
5 non-market values. GDP rises as crime and result-
6 ing policing expenditures increase, as extended com-
7 mutes boost gasoline consumption, and as parental
8 care is outsourced to hired services.

9 (6) Advances in statistical methods and data
10 availability have opened new opportunities to develop
11 supplemental indicators that better reflect national
12 welfare. Academics have proposed over two dozen
13 such indicators in the last two decades.

14 (7) Allies of the United States, including the
15 United Kingdom, have taken steps to develop new
16 comprehensive accounting and benchmarking sys-
17 tems. Several States have also already adopted indi-
18 cators that supplement gross State product.

19 SEC. 3. ESTABLISHMENT.

20 There is established a national comm1sswn to be
21 known as the "21st Century National Indicators Commis-
22 sion" (in this Act referred toast eh "Commission").

23 SEC. 4. MEMBERSHIP.

24 (a) NUMBER AND APPOINTMENT.-The Commission
25 shall be composed of 12 members. Not later than 30 days

1 after the date of the enactment of this Act, the President
2 shall appoint-

3 (1) eight members, not more than four of whom
4 shall be from the same political party;

5 (2) one member following consultation with the
6 Speaker of the House of Representatives;

7 (3) one member following consultation with the
8 minority leader of the House of Representatives;

9 (4) one member following consultation with the
10 President pro tempore of the Senate; and

11 (5) one member following consultation with the
12 minority leader of the Senate.

13 (b) QuALIFICATIONS.-Members of the Commission
14 shall meet the following qualifications:

15 (1) Members of the Commission shall be ap-
16 pointed from among persons who possess—

17 (A) an advanced degree in a field relevant
18 to developing social, technological, enViron-
19 mental, or economic indicators of national wel-
20 fare; or

21 (B) significant ex:penence developing so-
22 cial, technological, environmental, or economic
23 indicators of national welfare.

24 (2) Each member shall be a United States cit-
25 izen and shall reside in the United States.

1 (3) A member may not hold any paid position
2 in any local government, State government, or Exec-
3 utive agency (as defined in section 105 of title 5,
4 United States Code), or have held any such position
5 during the 5 year period preceding such member's
6 appointment to the Commission.

7 (4) A member may not be a party to an ongo-
8 ing contract with any local government, or State
9 government, or Executive agency, or be an employee
10 of an entity that is a party to such a contract.

11 (5) A member may not be a lobbyist, as defined
12 by State or Federal law, at the time that the mem-
13 ber is appointed to the Commission.

14

15

16 (c) TERMS.-

17 (1) IN GENERAL.-Each member shall be ap-
18 pointed for the life of the Commission

19 (2) VACANCIES.-A vacancy in the Commission
20 shall be filled in the manner in which the original
21 appointment was made. The appointment of the re-
22 placement member shall be made not later than 30
23 days after the date on which the vacancy occurs.

24 (3) MANDATORY RESIGNATION.-In the event
25 that a member accepts a position as an officer or

1 employee of any local government, State govern-
2 ment, or Executive agency, the member shall resign
3 from the Commission not later than 30 days after
4 the date on which the member accepts such position.

5 (4) COMPENSATION.-Members shall serve
6 without pay, except that members shall receive travel
7 expenses, including per diem in lieu of subsistence,
8 in accordance with applicable provisions under sub-
9 chapter ■ of chapter 57 of title 5, United States
10 Code.

11 (5) QuoRUM.-Six members of the Commission
12 shall constitute a quorum for the transaction of
13 business, but a lesser number may conduct hearings.

14 (6) PROFESSIONAL STAFF.-The Commission
15 may employ, pursuant to laws and regulations gov-
16 erning the civil service, an executive secretary and
17 any clerical, professional, and technical assistants as
18 may be necessary.

19 (7) CHAIRPERSON.-The President shall des-
20 ignate one member to serve as Chairperson of the
21 Commission.

22 (8) MEETINGS.-The Commission shall meet
23 not less than once per month at the call of the
24 Chairperson.

1 SEC. 5. DUTIES, RESPONSIBILITIES, AND POWERS.

2 (a) NEW ECONOMIC INDICATORS.-

3 (1) IN GENERAL.-The Commission shall deter-
4 mine how to compute, aggregate, and report a new
5 dollar-denominated set of indicators to be included
6 in the national income and product accounts, includ-
7 ing the following indicators:

8 (A) GDP.-An indicator, to be known as
9 G1, shall reflect GDP in its current form.

10 (B) PRESENT-DAY ECONOMIC WELFARE.-
11 An indicator to be known as G2, which shall
12 measure present-day economic welfare, includ-
13 ing market and non-market consumption, ad-
14 justed for differences **in** consumption-related
15 utility.

16 (C) SUSTAINABLE ECONOMIC WELFARE.-
17 An indicator to be known as G3, which shall
18 measure sustainable economic welfare. Building
19 on the G2 indicator, the G3 indicator shall ad-
20 just for changes in capital stocks and other de-
21 terminants of economic growth.

22 (D) PRESENT-DAY OVERALL WELFARE.-
23 An indicator to be known as G4, which shall
24 measure present-day overall welfare. Building
25 on the G2 indicator, the G4 indicator shall ad-

1 just for social and environmental factors that
2 help or hinder current welfare.

3 (E) SUSTAINABLE OVERALL WELFARE.-

4 An indicator to be known as G5, which shall
5 measure sustainable overall welfare. Incorporating the G2, G3, and G4 indicator adjust-
6 ments, the G5 indicator shall account for
7 changes in social and environmental stocks.
8

9

10

11

12 (2) COMPOSITION OF INDICATORS.-In order to
13 determine how to compute, aggregate, and report
14 the new indicators described in paragraph (1), the
15 Commission shall identify and address—

16 (A) the optimal composition of each indi-
17 cator, using criteria such as methodological cer-
18 tainty, welfare impact, and policy relevance to
19 determine the appropriate components of each
20 indicator;

21 (B) the optimal pricing formulae to assign
22 a dollar value to the components of each indi-
23 cator;

24 (C) recommendations for how and where
25 the Director of the Bureau of Economic Anal-

1 ys1s, and any other Executive agency that the
2 Commission determines that the Director
3 should consult under paragraph (3), can ac-
4 quire the data necessary to compute, aggregate,
5 and report the new indicators described in para-
6 graph (1).

7 (3) EXECUTIVE AGENCY COOPERATION.-The
8 Commission shall determine-

9 (A) which Executive agencies the Director
10 of the Bureau of Economic Mfairs should con-
11 sult in order to compute, aggregate, and report
12 the new indicators; and

13 (B) methods to foster cooperation among
14 any such Executive agencies.

15 (b) HEARINGS.-The Commission shall conduct three
16 public hearings across the United States on the compo-
17 nents to consider to measure national welfare and on how
18 to compute such measurements. The witnesses at such
19 hearings shall include statisticians and representatives
20 from relevant interest groups. The final hearing shall be
21 held in Washington, DC, not later than two years after
22 passage of this Act.

23

24 (c) ADDITIONAL POWERS.-The Commission may-

1 (1) meet at additional times and places that it
2 may consider appropriate;

3 (2) issue subpoenas reqmrmg the attendance
4 and testimony of witnesses and the production of
5 any evidence relating to any matter on which the
6 Commission is empowered to hold hearings under
7 subsection (b);

8 (3) administer oaths; and

9 (4) contract, as it considers appropriate, for the
10 provision of services, facilities, studies, and reports
11 that will assist the Commission in carrying out its
12 duties, responsibilities, and powers.

13 (d) INDICATOR DEVELOPMENT PLAN.-Not later
14 than the date that is two years after the President ap-
15 points the last member to the Commission, the Commis-
16 sion shall submit to the President an indicator develop-
17 ment plan that describes its findings under paragraphs (2)
18 and (3) of subsection (a) and makes recommendations for
19 how the Director of the Bureau of Economic Analysis
20 should use such findings in order to compute, aggregate,
21 and report the new indicators.

22 **SEC. 6. PRESIDENTIAL ACTION ON INDICATOR DEVELOP-**
23 **MENT PLAN.**

24 (a) PRESIDENTIAL CONSIDERATION.-Not later than
25 60 days after receiving the proposed indicator development

1 plan from the Commission under section 5(d), the Presi-
2 dent shall approve the plan without amendment, or return
3 the plan to the Commission for review. If the President
4 returns the plan to the Commission, the President shall
5 include any recommendations to revise the plan that Presi-
6 dent considers appropriate.

7 (b) REVIEW AND REVISION.-If the President re-
8 turns the proposed indicator development to the Commis-
9 sion for revision, the Commission shall-

10 (1) review the President's recommendations
11 submitted under subsection (a); and

12 (2) revise its plan, incorporating, at its discre-
13 tion, any recommendations to the plan that the
14 President proposed.

15 (c) RESUBMISSION.-Not later than 180 days after
16 the receipt of the President's recommendations submitted
17 under subsection (a), the Commission shall resubmit the
18 revised indicator development plan to the President.

19 (d) IMPLEMENTATION OF PLAN.-If the President
20 approves the indicator development plan, the Director of
21 the Bureau of Economic Affairs shall compute, aggregate,
22 and report the new indicators in accordance with the Com-
23 mission's recommendations in the indicator development
24 plan.

1 SEC. 7. TERMINATION.

2 The Commission shall terminate not later **than-**

3 **(1)** the date on which the President accepts the
4 first indicator development plan that the Commis-
5 sion submits under section 5(d); or

6 **(2)** the date on which the Commission submits
7 its revised indicator development plan under section
8 6(c).

9 SEC. 8. AUTHORIZATION OF APPROPRIATIONS.

10 There is authorized to be appropriated such sums as
11 may be necessary for each of fiscal years 2012 through
12 2015 for purposes of carrying out this Act.

This report's central premise is that new welfare indicators would lead to better-informed policymaking, yielding an improvement in welfare. Underlying this premise is a theory of policy change depicted at right. Each arrow represents a critical, but non-exclusive, causation link in the presumed chain of policy change. While other factors are certainly at play, this appendix focuses on justifying the links (Figure 5's numbered arrows) critical to this project:

- 1. Policy influences welfare.** Few will contest this assertion; many policy debates hinge on the degree to which the policy will help or hinder welfare. For example, the Environmental Protection Agency estimates that the 1990 amendments to the Clean Air Act saved over 160,000 lives in 2010.³⁷ Meanwhile, industry representatives have claimed since 1990 that the regulation has meant thousands of lost jobs each year.³⁸ Of course, policy impacts are not the only determinants of human welfare. While weather, genes, and other non-policy factors also play a role, this report is concerned with components of welfare impacted by policy decisions.
- 2. Indicators can capture policy-relevant welfare.** Given the focus on policymaking, this report seeks an indicator that exclusively measures policy-relevant components of welfare. In Figure 6, the pink box represents the confluence of policy impact and welfare input that this effort seeks to measure. While no indicator will wholly and exclusively capture policy-relevant welfare,

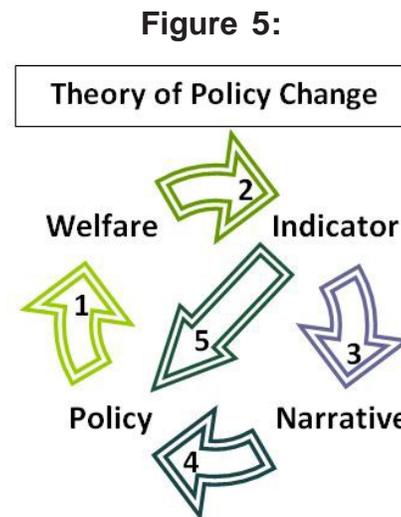
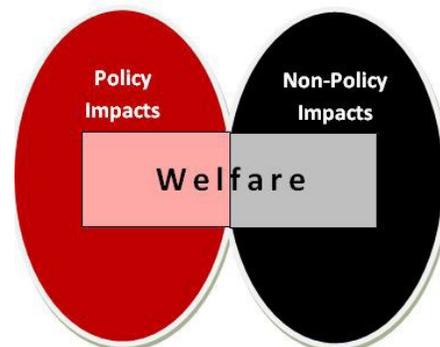


Figure 6: Policy-Relevant Welfare



carefully constructed indicators can provide a reasonable approximation.

- 3. Indicators influence the policy narrative.** The policy narrative is the public discourse commonly used to frame policy problems and define policymaking success. Headlines, think-tank publications, election debates, and advocacy campaigns all shape and reinforce the lens through which policy decisions are commonly seen. The performance of numerical benchmarks, particularly GDP,

strongly influences the policy narrative. New metrics could spur new, discourse-shifting headlines (e.g. "Rising Indebtedness Boosts GDP, Reduces Well-Being". There are two common critiques of this premise:

- a. "Qualitative arguments are more persuasive than quantitative ones in shifting policy discourse." Proponents of this critique assert, for example, that climate change activists propelled their concerns to the national stage via warnings of flooded coastlines and pictures of polar bears on melting ice caps. While these arguments certainly helped highlight the problem, the debate over policy solutions has been driven by numbers. Congressional hearings and lobbying campaigns have focused on the social cost of carbon on one side, and the GDP impact of carbon mitigation on the other. Not attempting to quantify the cost of carbon emissions would leave a lopsided debate in which projected GDP losses would handily trump the most compelling melting ice cap picture. While qualitative and quantitative arguments both influence policy, the critical role of numbers should not be overlooked.
- b. "Among quantitative measures, GDP is not that important in swaying policy debates." Those who make this argument point *out*, for example, that the primary statistic stalling passage of the U.S.-Colombia free trade agreement (FTA) in 2009-2010 was the number of unionists murdered in Colombia. While such metrics do play a role, GDP tends to trump them in policy influence. Indeed, the U.S.-Colombia FTA eventually passed in 2011 not because the number of union deaths decreased (indeed, the death toll remained at a constant level⁹), but because an increasing number of policymakers from both parties emphasized the FTA's expected boost to exports, and thus, GDP.⁴⁰ In pushing for passage of the FTA last *year*, the Obama Administration sidestepped labor concerns and reduced the cost-benefit analysis to a single rubric: a projected \$2.5 billion addition to GDP.⁴¹

4. **The policy narrative influences policymaking.** Guiding policy frames are transmitted by media outlets, employed by policy advocates, adopted by voters, and echoed by politicians themselves. As such, it is difficult for legislators to pass policy changes that do not fit into a leading narrative, given that doing so carries the risk of political isolation and lost votes. While other factors certainly also

influence the design of policy (e.g. policymakers' personal beliefs, partisan pressure, etc.), indicator-informed policy narratives play an important role.

5. Indicators can directly inform policymaking. The U.S. federal government has historically institutionalized particularly potent metrics as benchmarks for policymaking. For example, the Federal Reserve implicitly targets a 2% inflation rate in setting monetary policy while many Congressional members last year sought to formalize deficit reduction targets. Such benchmarking may be on the rise. The 2010 Government Performance and Results Modernization Act requires that the Office of Management and Budget establish “long-term, outcome-oriented goals for... crosscutting policy areas” every four years.⁴² New welfare indicators could inform such holistic targeting.

Accepting these causal links should lead one to accept the central premise that new welfare indicators could improve national well-being by enabling better-informed policymaking.

One additional premise underlying our approach is that new measures of progress can be created without a new general theory of progress. Some GDP critics argue that GDP, despite its shortcomings, remains a theoretically coherent indicator by measuring not just the ends policy should seek to achieve, but the means of achieving them. Designed to fit Keynesian demand management theory, GDP's disaggregated form tracks the degree of government spending and investment, the two primary policymaking levers for altering GDP itself. As such, GDP intends to offer policymakers a direct recipe for success—if government spending is increased by X, GDP will increase by Y.

Available alternative indicators make no such attempt. While they may measure inequality levels, for example, they do not directly name or track inequality's policy inputs, largely because this report does not present a theory to suggest a surefire policy recipe for inequality reduction. Some argue that such theory should precede the creation of new headline indicators so as to mimic the diagnostic strength of GDP. Without a theory describing the policy levers that influence our expanded policy objectives, they assert, we will be left with indicators that measure ends but provide no guidance on how to influence those ends.

While the desire for such a theory is understandable, its development does not seem imminent. The array of policy objectives this report seeks to measure are far more numerous and complex than gross economic activity. The corresponding theory would need to describe the means by which many interrelated policy levers could simultaneously achieve a constellation of policy aims. By contrast, the Keynesian theory behind GDP sought to maximize a singular aim of economic output. Rather than waiting to see if such a complex theory emerges, it seems better to create indicators that at least broaden the scope of policy objectives. Indeed, doing so may act as a step toward developing a theory of how policies should be constructed to deliver the complex objectives. Merely waiting for a theory, however, would mean maintaining the status quo on both policy means and ends during the long interim.

Appendix 3: Taxonomy of Alternative Indicators

We borrow from the Stiglitz-Sen-Fitoussi Commission three categories for grouping new welfare indicators:⁴³

1. **Corrected GDP:** eliminating components that should not count and adding ones that should
2. **Composite Indexes:** combining other indicators with GDP to reach a synthesized number
3. **Subjective well-being measures:** aggregating self-reported utility

1. Corrected GDP: these might be more politically palatable for adoption, since they start with the GDP formula. Yet, they are difficult to calculate, in part due to the difficulty of imputing market values for non-market goods.

- **Sustainable Measure of Economic Welfare** (Nordhaus/Tobin, 1973):
The indicator starts with household consumption. It subtracts elements not conducive to welfare (commuting) and adds ones that are (leisure time). This yields the MEW. To convert to the SMEW, they reduce the MEW by the amount that would need to be devoted to keeping the capital stock constant (i.e. accounting for depreciation). Their capital stock includes reproducible capital, land, net foreign assets, educational capital, and health capital (but not natural resources beyond land).
 - The SMEW differed from GNP in magnitude (higher due to leisure), but paralleled GNP in fluctuations. Nordhaus/Tobin thus concluded that GNP was a fine benchmark for policy.
- **Index of Sustainable Welfare** (Daly/Cobb, 1989) and **Genuine Progress Indicator** (Talberth/Cobb/Slattery, 2006):
 - ISEW is very similar to SMEW construction, but with three key differences:
 - Natural resource depreciation is included in the adjustments to the indicator, measured as the amount of extra money that would need to be invested in renewable substitutes.
 - Income distribution is accounted for by multiplying consumption by $1 - \text{GiniCoefficient}$.
 - Leisure time isn't monetized.
 - Genuine Progress Indicator: this is almost identical to ISEW, with some differences:

- GPI accounts for leisure time (discounting for hours worked above a threshold)
- GPI discounts the degree of social ills like divorce and crime.
- Both ISEW and GPI tend to be lower than SMEW (due to accounting for leisure time) and GDP (due to environmental depreciation, inequality discounting, etc.). Growth is not nearly as correlated between ISEW/GPI and GDP as between SMEW and GDP. Max-Neef (1995) posits that GDP contributes to growth in these welfare indexes up to a certain threshold, and then loses its correlation.
- **The Fleurbaey/Gaulier Indicator** (Fleurbaey/Gaulier, 2007): This French measurement is similar to ISEW/GPI. It tries to incorporate even more monetary values of non-monetary goods (job security, healthy life expectancy, environmental sustainability), using the amount that people value these goods to create adjusted "equivalent incomes" which are then aggregated and reduced by the degree of inequality in the equivalent incomes. We don't have info on how this measure has changed over time, but the static indicator (from 2004) shows significantly different country rankings than when ordered by GDP.

2. Composite Indexes: these are easier to compute than corrected GDP indicators, and are still somewhat palatable for policy given that the subindexes are objective. *Yet*, the weighting of different subindexes is typically arbitrary (as is the overall criteria for welfare in some instances).

- **Human Development Index** (UNDP, 1990): HDI combines three assumed components of development (presumably both causes and consequences of development). Each one is indexed according to the current probable range of values.
 - Here are the three indexes, combined without weights to produce the final number:
 - Wealth: GDP per capita (PPP) at \$100 = 0, \$40K = 100
 - Health: average life expectancy at 25 = 0, 85 = 100
 - Education: a composite of adult literacy rate (2/3 weight) and school enrollment (1/3 weight) at 0% = 0, 100% = 100
 - HDI-based country rankings are significantly different than GDP-based rankings.
- **Gender-related Development Index** (UNDP, 1995): GDI modifies HDI

by calculating the relative position of women vs. men in each of the three subcomponents. Inequality between the genders in health, education, or wealth is discounted. (GDI also makes the education component more weighted towards enrollment than literacy—the reverse of HDI.)

- **Index of Economic Well-Being** (Osberg/Sharpe, 2002): This measure combines diverse indexes into one aggregate. Like HDI, each is indexed according to the probable range of values (though IEWB takes into account historical values).
 - Each of these components gets an equal weight in the final index:
 - Consumption flows: measured the same way as the C component of GDP
 - Net capital accumulation: RandD, human capital contribution, costs of CO₂
 - Inequality: Gini and a poverty intensity index
 - Social factors: unemployment, illness, single parent poverty, old-age poverty (cost of these conditions times the probability of getting into the condition)
 - IEWB-based rankings are also different from GDP-based lists, both in static rankings and in performance over time (diverging in France since the 1980s)
- **Canadian Index of Well-being** (CIW Network (an NGO), 2011): This composite index is comprised of eight dimensions of welfare: democratic engagement, community vitality, education, environment, healthy populations, leisure and culture, living standards, and time use. Each categorical subindicator is in turn the combination of a dozen or so objective measures (for example, the environment indicator includes GHG emissions, ozone levels, a water quality index, a timber sustainability index, etc.). Each of the eight subindicators is combined with equal weights to create an average CIW, largely because of the lack of a rationale at this point for particular weighting (though such rationale may be forthcoming). Since 1994, three of the subindicators have decreased (environment, time use, and leisure and culture), and none of the five that have increased have risen as much as GDP, meaning that the CIW index shows much lower growth than GDP.
- **Quality of Development Index** (Tellus Institute, 2011): This composite index incorporates the following six, equally-weighted, metrics: GDP per capita, average annual hours of work (to measure leisure), Gini coefficient, share of population that is hungry, carbon emissions per capita, and degree of forested or protected land.
- **BIP40** (French NGO network, 2002): This includes variables for work, health,

education, and housing quality and uses a methodology similar to HDI to create an aggregate value.

- **Regional Quality of Development Index (QUARS)** (Italian NGO network, 2000): This is a composite index of 45 variables pertaining to environment, economy, rights, gender equality, education, culture, working conditions, health, and political participation.
- **Sustainable Development Indicators** (U.S. Interagency Working Group on SDI, 1997): This is a composite index of 32 social, economic, and environmental indicators proposed by a federal-level task force comprised of twelve departments of the U.S. government.

3. Subjective Measures: these are based on people's own evaluation of their well-being. This obviates the need to predefine what well-being is, assuming that each person is in the best position to name her/his well-being. These indicators are, however, less politically palatable given a distrust of subjective measures. Also, subjective measures can be influenced by variables irrelevant to policymaking, such as the weather.

- **Gross National Happiness** (Bhutan, 1972): GNH is based on a big household survey, which has been conducted twice. The 2010 survey assesses nine different dimensions of welfare, which are comprised of 33 distinct indicators (e.g. literacy is one under education, spirituality is one under psychological wellbeing), each of which is assessed by several survey questions. Some questions are objective (e.g. income earned) while others are subjective (e.g. self-reported health status). GNH employs a threshold approach, judging the minimum satisfactory point score for each of the 33 indicators. For each of the nine dimensions, they aggregate threshold satisfaction for the included indicators to determine whether the overall wellbeing dimension was satisfactorily achieved (in this aggregation, objective indicators tend to get greater weight than subjective indicators). Aggregating the nine dimension results for all survey respondents, they calculate:
 - **Headcount:** a headcount of those who do not achieve sufficiency for six of the nine domains (and do not achieve two-thirds sufficiency in the 33 weighted indicators). As such, each of the nine domains is weighted as equally important, though the indicators within each domain are differently weighted.
 - **Intensity:** an assessment of the average number of domains not satisfied among those who are relatively deprived (those who did not meet the threshold for at least six of the nine domains).
 - **Overall Index:** The overall index is 1 minus the headcount times the intensity

degree, creating a number between 0 and 1 that rises as the headcount or intensity level falls. As such, the GNH isn't really an assessment of the relative welfare of the country, but the relative deficiency in welfare. (However, it seems that an assessment of the relative welfare could be calculated by simply aggregating the point scores for each of the 33 indicators.)

The government can use the survey results to see which wellbeing dimensions are least satisfied and which districts and demographic groups are most deficient, so as to allocate resources accordingly. Also, many regulatory policies are scored on their probable impact on dimensions that are similar to (but not exactly the same as) the GNH dimensions.

- **Gross National Happiness, 2nd Gen** (Med Jones, 2006): This proposal, out of Jones' International Institute of Management, aims to measure a generalized GNH by assessing seven different areas of wellness: economic, environmental, physical, mental, workplace, social, and political. Most wellness areas include both subjective results (via survey) and objective data.
- **Inequality-Adjusted Happiness** (Veenhoven/Kalmijn, 2005): takes into account the mean/variance of a basic well-being survey question. It rests only on this subjective indicator, not incorporating any objective accounts.
- **Advanced Quality of Life Index** (Diener, 1995): it combines the answers to subjective well-being questions with objective indicators of health, income, environment, etc.
- **Happy Life Expectancy** (Veenhoven, 1990): this similarly combines subjective responses with objective measures. Yet, the rankings have produced counterintuitive results, like a positive correlation between the index and the unemployment rate.
- **Happy Planet Index** (New Economics Foundation, 2006): This is actually a combination of the three types of indicators, combining life expectancy (used in the composite measures), ecological footprint (a partial GDP adjustment), and life satisfaction (a subjective indicator). It is computed as (life expectancy x life satisfaction)/ecological footprint. The object is not to measure gross happiness but to see how efficiently (in environmental terms) a country is able to achieve relative well-being. They find that a) no country in the world does good at all three subindicators, and b) countries tend to see the index rise with GDP per capita up until \$5000, after which the index tends to decline (due to life satisfaction not increasing much more and ecological footprint increasing).

Systems of Indicators

In addition to proposals for specific indicators, there have been proposals for series of indicators grouped together as an accounting system, to be used on par with the UN's System of National Accounts (which includes GDP). Most of these series have been suggested to supplement, not replace, the System of National Accounts. By being compatible with the System of National Accounts, these numbers allow for easy comparison with GDP—both can be decomposed in similar ways to look at the various impacts, say, of a target sector. Here are some proposed new systems:

- **SEEA:** The System of Economic Environmental Accounts is the best known one. It would have four categories of indicators: flows of pollutants/resources/energy, expenditures on environmental protection, stocks of natural resource assets, and environmentally-adjusted macroeconomic indicators. The UN has not fully formalized the method to use for creating these SEEA indicators. However, using the conceptual proposal of the SEEA, Germany has proposed German Environmental-Economic Accounting, a supplemental accounting system that will measure economic activity after depreciating for environmental impacts.
- **NAMEA:** Similar to SEEA is the National Accounting Matrix including Environmental Accounts (NAMEA), which prescribes indicators that allow governments to determine which sectors have the biggest contribution to environmental problems (e.g. air emissions accounts by economic activities). It does not take into account social concerns.
- **SESAME:** The System of Economic and Social Accounting combines social data (from Socio-Economic Accounts) and environmental data (from NAMEA).

- 1972:** Bhutan first announces Gross National Happiness
- 1973:** Nordhaus and Tobin propose the Sustainable Measure of Economic Welfare
- 1989:** Daly and Cobb propose the Index of Sustainable Welfare
- 1990:** UNDP releases the Human Development Index
Veenhoven proposes Happy Life Expectancy
- 1995:** UNDP releases the Gender-related Development Index
Diener proposes the Advanced Quality of Life Index
- 1997:** An interagency task force of the U.S. federal government (SDI Group) proposes the Sustainable Development Indicators. (See below.)
- 2000:** An Italian coalition of NGOs propose the Regional Quality of Development Index (QUARS)
- 2002:** A French non-governmental network proposes the BIP-40
Osberg and Sharpe propose the Index of Economic Well-Being
- 2004:** Alan Krueger, current chair of the Council of Economic Advisers, publishes a methodology for calculating self-reported welfare.
- 2005:** Veenhoven and Kalmijn propose Inequality-Adjusted Happiness
- 2006:** Talberth, Cobb, and Slattery propose the Genuine Progress Indicator
Med Jones proposes a second generation of Gross National Happiness
The New Economics Foundation releases the Happy Planet Index
- 2007:** The European Commission hosts the Beyond GDP conference.
Fleurbaey and Gaulier propose an adjusted GDP indicator
- 2008:** The Stiglitz-led Commission on the Measurement of Economic Performance and Social Progress begins its investigation of GDP alternatives. (See below.)
- 2009:** As follow-up to the Beyond GDP conference, the European Commission publishes the EU Roadmap to develop more holistic indicators. (See below.)
- 2010:** President Obama signs into law a bill to create a “Key National Indicators System” to create a more consolidated and accessible short list of national-level indicators. (See below.)
- 2011:** The Tellus Institute proposes the Quality Development Index
The OECD hosts a post-Stiglitz Commission conference to discuss the road ahead, entitled “What Well Being and Sustainability Measures?”

Notes on Indicator-Advancing Initiatives

U.S. Interagency Working Group on Sustainable Development Indicators

Formed in 1994, this group was comprised of representatives from ten governmental departments, the *EPA*, and *NASA*. Initially the group reported to the President's Council on Sustainable Development, but later started reporting to the Council on Environmental Quality of the Executive Branch. The mandate of the group was to 1) select and organize a set of sustainable development indicators, 2) provide public access to the compiled indicators and disaggregated data, 3) report on progress in the indicators, and 4) create an inclusive strategy for ongoing development of the indicators. They only accomplished part of the first goal—in 1997 they proposed a list of 32 social, economic, and environmental indicators that should be tracked and integrated into policymaking. The working group did not publicize data on these indicators, track progress, or propose a strategy for future indicator evolution and implementation.

Beyond GDP Initiative

The Beyond GDP Initiative began as a 2007 conference organized by the European Commission, the European Parliament, *OECD*, *The Club of Rome*, and World Wildlife Fund. With delegates from over 50 countries, the conference indicted *GDP*'s shortcomings, surveyed existing alternative indicators, and called for national-level progress toward adopting such supplemental measures. As a follow-up to the conference, the European Commission began working on a road map for reforming *GDP* within the EU. The road map identified the following as key next steps: 1) Developing aggregate environmental and social indicators to complement *GDP*, 2) Increasing the flow of data for on such indicators for policymaking purposes, 3) Developing better reporting on inequality. 4) Creating a European Sustainable Development Scoreboard, and 5) Integrating environmental and social indicators into the National Accounts system. For examples of progress made toward these goals, see the Background section.

Stiglitz Commission

The Commission, chaired by Stiglitz, *Sen*, and Fitoussi, was convened by French President Sarkozy with the mandate to investigate *GDP* alternatives for a broad (non-targeted) audience of policymakers, politicians, academics, and civil society actors. After a survey of existing *GDP* alternatives, several plenaries on competing theories, and new welfare research, the commission issued its final report in September 2009. It recommended changes to a) better measure purely material welfare (by focusing on income rather than production, discounting inequality, imputing prices for non-market activities), b) incorporate non-material components of welfare with a focus on maximizing capabilities rather than outcomes (by using both objective indicators like health/education and subjective self-reporting), and c) begin measuring the sustainability of economies as separate from a current welfare indicator (by assessing changes in economic and physical stocks). The report brought greater publicity to the alternative indicator debate and propelled further research. *Yet*, concrete implementation of the commission's recommendations has thus far been limited (though the French statistical bureau has begun to collect the recommended data). In October of 2011, the *OECD* hosted a follow-up conference entitled "What Well Being and Sustainability Measures?"

Key National Indicators System

In 2003 the *GAO* hosted a forum on the need for a more consolidated and transparent system of key national indicators. After the issue garnered Congressional and interagency support, the *GAO* formally recommended that Congress create the system in 2006. In 2008, Representatives Ted Kennedy and Michael Enzi submitted legislation for the creation of a Key National Indicators Commission to establish and oversee a system of indicators that would inform the public and policymakers on relative progress towards national priorities. The bill passed and became law in 2010, and the bipartisan commission was selected later that year. In June of 2011, the *GAO* published a report on existing systems of indicators used by local and foreign governments to instruct the Commission as they determine the content and form of the U.S. indicator system. Due to a lack of appropriations to date, the Commission has yet to begin concerted work on designing the new dashboard of indicators.

PERSONAL INTERVIEWS

We would like to thank the following individuals, all personally interviewed, for lending insights that inform this report's analysis:

Frank Ackerman, Director, Climate Economics Group, Stockholm Environment Institute

Dean Baker, Co-Director, Center for Economic and Policy Research

William Beach, Director, Center for Data Analysis, The Heritage Foundation

Sarah Burd-Sharps, Co-Director, American Human Development Project

John Cavanagh, Director, Institute for Policy Studies

Clifford Cobb, Co-creator of the Index of Sustainable Economic Welfare and the Genuine Progress Indicator

The Congressional Research Service

Lew Daly, Director, Sustainable Progress Initiative, Demos

Karen Dolan, Fellow, Cities for Progress, Institute for Policy Studies

Nicholas Eberstadt, Henry Wendt Scholar in Political Economy, American Enterprise Institute (and appointee to the Key National Indicators Commission)

Michael Eddy, Consultant, Gross National Happiness Commission of Bhutan

Nancy Folbre, Professor of Economics, University of Massachusetts Amherst (and member of the Commission on the Measurement of Economic Performance and Social Progress)

Jeffrey Frankel, James W. Harpel Professor of Capital Formation and Growth, Harvard Kennedy School of Government

Richard Freeman, Herbert Ascherman Chair in Economics, Harvard University

Jeff Gillenkirk, Co-creator of the Genuine Progress Indicator

Neva Goodwin, Co-Director, Global Development and Environment Institute at Tufts University

Lucas Hitt, Congressional and Public Relations, Bureau of Economic Analysis

Christopher Hoenig, Senior Adviser to the Presidents, The National Academies (and President and CEO of The State of the USA)

David King, Lecturer in Public Policy, Harvard Kennedy School of Government

Steve Landefeld, Director, Bureau of Economic Analysis

Robert Lawrence, Albert L. Williams Professor of International Trade and Investment,
Harvard Kennedy School of Government

Sean McGuire, Director, Sustainability Policies, Maryland Department of Natural Resources

Chris Mihm, Managing Director for Strategic Issues, U.S. Government Accountability Office

Leticia Miranda, Associate Director, Economic and Employment Policy Project, National
Council de la Raza

Chad Moutray, Chief Economist, National Association of Manufacturers

Susan Offutt, Chief Economist, U.S. Government Accountability Office

Noel Ortega, Coordinator, New Economy Working Group

Bob Parker, former Chief Statistician, U.S. Government Accountability Office

Richard Parker, Lecturer in Public Policy, Harvard Kennedy School of Government

Robert Pollin, Co-Director, Political Economy Research Institute, University of
Massachusetts Amherst

Robert Putnam, Peter and Isabel Malkin Professor of Public Policy, Harvard Kennedy School
of Government (and member of the Commission on the Measurement of Economic
Performance and Social Progress)

Juliet Schor, Professor of Sociology, Boston College

Joel Segal, Senior Advisor, Office of Congressman John Conyers

Samuel Sherraden, Associate Program Director, Economic Growth Program, New America
Foundation

Elizabeth Stanton, Senior Scientist, Stockholm Environment Institute

Chad Stone, Chief Economist, Center on Budget and Policy Priorities

John Stutz, Senior Fellow, Tellus Institute

John Talberth, Senior Economist, World Resources Institute (and Co-creator of the Genuine
Progress Indicator)

ENDNOTES

1. Haque, Umair. 2011. *Betterness: Economics for Humans*. Boston: Harvard Business Press.
2. U.S. Senate. 1934. "National income 1929-1932." 73rd Congress, 2nd Session, Document No. 124. Washington, DC: Government Printing Office.
3. Di Leo, Luca. 2010. "Oil Spill May End Up Lifting GDP Slightly." *Wall Street Journal*, June 15. <http://blogs.wsj.com/economics/2010/06/15/oil-spill-may-end-up-lifting-gdp-slightly>.
4. U.S. Senate. 1934. "National income 1929-1932." 73rd Congress, 2nd Session, Document No. 124. Washington, DC: Government Printing Office.
5. Di Leo, Luca. 2010. "Oil Spill May End Up Lifting GDP Slightly." *Wall Street Journal*, June 15. <http://blogs.wsj.com/economics/2010/06/15/oil-spill-may-end-up-lifting-gdp-slightly>.
6. National Accounts Review Committee. 1958. "The National Economic Accounts of the United States: Review, Appraisal, and Recommendations." <http://www.nber.org/chapters/c6359.pdf>
7. Marcuss, R. D. and R. E. Kane. 2007. "US National Income and Product Statistics: Born of the Great Depression and World War II." *Bureau of Economic Analysis: Survey of Current Business* 87(2): 32–46.
8. Bennett, William. 1999. *The Index of Leading Cultural Indicators: American Society at the End of the 20th Century*. New York: Touchstone Publishing.
9. Stiglitz, Joseph, Amartya Sen, and Jean-Paul Fitoussi. 2009. "Report by the Commission on the Measurement of Economic Performance and Social Progress." http://www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf.
10. European Commission. 2007. "Beyond GDP." Conference Proceedings. http://www.beyond-gdp.eu/proceedings/bgdp_proceedings_intro_ses1.pdf
11. BBC News. 2011. "US Considers Factors for a National Wellbeing Index." BBC News Online. <http://www.bbc.co.uk/news/uk-14272038><http://www.bbc.co.uk/news/uk-14272038>
12. French National Institute of Statistics and Economic Studies. 2011. "The French National Statistical Agenda." http://www.insee.fr/en/publications-et-services/dossiers_web/stiglitz/stiglitz_agenda_anglais_080212.pdf
13. The White House. 1994. "Green GDP." http://clinton1.nara.gov/White_House/EOP/OVP/html/Enviro_GDP.html
14. Moss, David and Sarah Brennan. 2002. "National Economic Accounting: Past, Present, and Future." Harvard Business School Case 9-703-206. <http://www.ceibs.edu/ibmcep/schedule/images/20100919/27667.pdf>
15. Dodaro, Gene L. 2011. "Managing for Results: GPRA Modernization Act Implementation Provides Important Opportunities to Address Government Challenges." Washington D.C.: Government Accountability Office, May 10. <http://www.gao.gov/assets/130/126150.pdf>.
16. Canadian Index of Wellbeing. 2012. "Composite Index." <http://ciw.ca/en/CanadianIndexOfWellbeing/CompositeIndex.html>
17. Cohn, Jonathan. 2011. "Measuring for the Future." New York: New Economics Institute.
18. Posner, Stephen. 2010. "Letters: The Rise and Fall of GDP." *The New York Times Magazine*, May 27. http://www.nytimes.com/2010/05/30/magazine/30letters-t-THERISEANDFA_LETTERS.html
19. Daly, Lew. 2012. "Growth Does Not Equal Progress: Why GDP Is (Increasingly) Obsolete." *Nieman Watchdog*, February 21. <http://www.niemanwatchdog.org/index.cfm?fuseaction=background.view&backgroundid=611>
20. The American Human Development Project. 2011. "About the American Human Development Project." <http://www.measureofamerica.org/the-measure-of-america-2010-2011-book/faq/>
21. Fixler, Dennis. US Bureau of Economic Analysis. "Revisions to GDP Estimates in the US." Presentation to OECD Workshop on Revisions. 7 October 2004.
22. Bureau of Economic Analysis. 2011. "Concepts and Methods of the U.S. National Income and Product Accounts." <http://www.bea.gov/national/pdf/chapters1-4.pdf>
23. Stiglitz, Joseph, Amartya Sen, and Jean-Paul Fitoussi. 2009. "Report by the Commission on the Measurement of Economic Performance and Social Progress." http://www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf.

24. Brickman, Philip, Dan Coates, and Ronnie Janoff-Bulman. 1978. "Lottery winners and accident victims: Is happiness relative?" *Journal of Personality and Social Psychology*, Vol 36(8). 917-927.
25. American Sociological Association. 2005. "Money Can Buy You Happiness, but Only Relative to Your Peer's Income." <http://www.asanet.org/press/20050814.cfm>
26. Easterlin, Richard, et al. 2010. "The happiness-income paradox revisited." *PNAS*, December 13.
27. D. Kahneman et al. 2004. "Toward National Well-Being Accounts," *The American Economic Review* 94 (2). 429-434.
28. United States Department of Commerce Economics and Statistics Administration. 2011. "Fiscal Year 2012 Budget Estimates." http://www.bea.gov/about/pdf/bea_fy_2012_congressional.pdf
29. Office of Management and Budget. 2012. "Fiscal Year 2012 Budget for the Department of Commerce." <http://www.whitehouse.gov/sites/default/files/omb/budget/fy2013/assets/commerce.pdf>
30. Bureau of Economic Analysis. 2006. "Preliminary Study Shows Research and Development Contribution to Economic Growth." <http://jwww.bea.gov/newsreleases/generaljrd/2006/rdspond06.htm>
31. Bregger, J. and Steven Haugen. 1995. "BLS introduces new range of alternative employment measures." *Monthly Labor Review* 118: 19-26.
32. Boxer, Sarah. "Romney to Obama: 'Not so Fast'" CBSNews. CBS Interactive, 05 Feb. 2012. Web. 10 Mar. 2012. <http://www.cbsnews.com/8301-503544_162-57371646-503544/romney-to-obama-not-so-fast/?tag=contentMain;contentBody>.
33. Dodaro, Gene L. 2011. "Managing for Results: GPRA Modernization Act Implementation Provides Important Opportunities to Address Government Challenges." Washington D.C.: Government Accountability Office, May 10. <http://www.gao.gov/assets/130/126150.pdf>.
34. Costanza, Robert, Maureen Hart, Stephen Posner, and John Talberth. 2009. "Beyond GDP: The Need for New Measures of Progress." Boston University Pardee Center Working Paper No.4. <http://jwww.bu.edu/pardee/files/documents/PP-004-GDP.pdf>
35. Bennett, William. 1993. "Quantifying America's Decline." *The Wall Street Journal*. <http://www.columbia.edu/cu/augustine/arch/usadecline.html>
36. DeNavas-Walt, Carmen, Bernadette D. Proctor, and Jessica C. Smith. 2011. "Income, poverty, and health insurance coverage in the United States: 2010." *Current Population Reports*. Washington, DC: United States Census Bureau.
37. Environmental Protection Agency. 2011. "The Benefits and Costs of the Clean Air Act from 1990 to 2020." <http://www.epa.gov/oar/sect812/feb11/summaryreport.pdf>
38. Kilborn, Peter. 1996. "East's Coal Towns Wither In the Name of Cleaner Air." *The New York Times*, February 15. <http://jwww.nytimes.com/1996/02/15/us/east-s-coal-towns-wither-in-the-name-of-cleaner-air.html?pagewanted=all&src=pm>
39. International Trade Union Confederation. 2011. "Colombia." Annual Survey of Violations of Trade Union Rights. <http://survey.ituc-csi.org/Colombia.html?lang=en>
40. US Senate Republican Policy Committee. 2011. "Trade Agreements Finally Poised to Create Jobs." http://rpc.senate.gov/public/index.cfm?p=PolicyPapers&ContentRecord_id=77583506-624b-4c7e-bcd9-9fcbe705b809
41. White House. 2011. "Fact sheets: U.S.-Colombia trade agreement and action plan." Washington, DC: Office of the Press Secretary, 6 April.
42. Dodaro, Gene L. 2011. "Managing for Results: GPRA Modernization Act Implementation Provides Important Opportunities to Address Government Challenges." Washington D.C.: Government Accountability Office, May 10. <http://jwww.gao.gov/assets/130/126150.pdf>.
43. Commission on the Measurement of Economic Performance and Social Progress. 2008. "Survey of Existing Approaches to Measuring Socio-economic Progress." http://jwww.stiglitz-senfitoussi.fr/documents/Survey_of_Existing_Approaches_to_Measuring_Socio-Economic_Progress.pdf

WORKS CONSULTED

- Abraham , Katherine and Christopher Mackie (eds.). 2005. *Beyond the Market: Designing Nonmarket Accounts for the United States*. Washington, DC: National Academies Press.
- Bennett, William. 1999. "The Index of Leading Cultural Indicators: American Society at the End of the 20th Century." New York: Touchstone Publishing.
- Cobb, Cliff., Ted Halstead, et al. 1995. "If the GDP is Up, Why is America Down?" *The Atlantic Monthly*. 276: 59–78.
- Cobb, Clifford, Slattery, Noah, and John Talberth. 2007. "The Genuine Progress Indicator 2006." Oakland, CA: Redefining Progress.
- Cohn, Jonathan. 2011. *Measuring for the Future*. New York: New Economics Institute.
- Costanza, Robert, Maureen Hart, Stephen Posner, and John Talberth. 2009. "Beyond GDP: The Need for New Measures of Progress." Boston University Pardee Center Working Paper No.4.
- Daly, Herman and John Cobb. 1989. *For the Common Good: Redirecting the Economy Toward Community, the Environment, and a Sustainable Future*. Boston: Beacon Press.
- Daly, Lew and Stephen Posner. 2011. *Beyond GDP: New Measures for a New Economy*. New York: Demos.
- "Economist Debates: GDP." 2010. *The Economist*, April. <http://www.economist.com/debate/overview/171>
- Gertner, Jon. 2010. "The Rise and Fall of GDP," *New York Times*. <http://www.nytimes.com/2010/05/16/magazine/16GDP-t.html>.
- Government Accountability Office. 2011. "Key Indicator Systems." Report to Congressional Attendees, March.
- Hamilton, K., G. Ruta, et al. 2006. *Where Is the Wealth of Nations? Measuring Capital for the 21st Century*. Washington, DC: The World Bank.
- Haque, Umair. 2011. *Betterness: Economics for Humans*. Boston: Harvard Business Press.
- Jorgenson, Dale W., J. Steven Landefeld, and William D. Nordhaus, (eds.) 2006. *A New Architecture for the U.S. National Accounts*. Chicago: University of Chicago Press.
- Landefeld, J .S., et. al. 2010. "GDP and Beyond: measuring economic progress and sustainability," *Survey of Current Business* 91(4):12-25.
- Landefeld, J .S. and Shaunda Villones. 2009. "National Time Accounting and National Economic Accounting," in Alan B. Krueger (Ed.), *Measuring the Subjective Well-Being of Nations: National Accounts of Time Use and Well-Being*, National Bureau of Economic Research Conference Report Chicago: University of Chicago Press.

Kahneman, D. et al. 2004. "Toward National Well-Being Accounts," *The American Economic Review* 94 (2). 429-434.

Marone, Heloisa. 2011. "Measuring Economic Progress and Well-Being: How to Move Beyond GDP?" Oxfam America.

McCulla, S. H. and S. Smith. 2007. *Measuring the Economy: A Primer on GDP and the National Income and Product Accounts*. Washington, DC: Bureau of Economic Analysis.

Marcuss, R. D. and R. E. Kane. 2007. *US National Income and Product Statistics: Born of the Great Depression and World War II*. Bureau of Economic Analysis: *Survey of Current Business* 87(2): 32-46.

Moss, David and Sarah Brennan. 2002. "National Economic Accounting: Past, Present, and Future." Harvard Business School Case 9-703 206. <http://www.ceibs.edu/ibmcep/schedule/images/20100919/27667.pdf>

National Accounts Review Committee. 1958. "The National Economic Accounts of the United States: Review, Appraisal, and Recommendations." <http://www.nber.org/chapters/c6359.pdf>

Nordhaus, William D., and James Tobin. 1973. "Is Growth Obsolete?" In Milton Moss (ed.) *The Measurement of Economic and Social Performance*. New York: Columbia University Press. 509-534.

Nordhaus, William D., and Edward C. Kokkelenberg, (eds.) 1999. *Nature's Numbers: Expanding the National Economic Accounts to Include the Environment*. Washington, DC: National Academy Press.

Organization for Economic Cooperation and Development. 2009. *Measuring the Progress of Societies*. <http://measuringprogress.org/>

Stiglitz, Joseph, Amartya Sen, and Jean-Paul Fitoussi. 2009. "Report by the Commission on the Measurement of Economic Performance and Social Progress." http://www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf.

Stutz, John. 2011. "The Quality of Development Index: A New Headline Indicator of Progress." *Journal of Futures Studies*, 15 (3): 73 – 102.

Talberth, J. 2008. *A New Bottom Line for Progress*. 2008 *State of the World: Innovations for a Sustainable Economy*. Starke. New York: WW Norton & Company.

The Centre for Bhutan Studies. 2011. "GNH Survey Findings 2010." *Gross National Happiness*. <http://www.grossnationalhappiness.com/>

United Nations, Commission of the European Communities, International Monetary Fund, Organisation for Economic Co-operation and Development, and World Bank. 1993. *System of National Accounts*. New York: United Nations.