FEDERAL BUDGET REFORM: A BEHAVIORAL APPROACH

James J. Hearn and Marvin Phaup

Abstract

The budget process put in place in 1974 was a big step toward a rational and logical procedure for imposing fiscal discipline and improving allocative decisions. It likely would have been more successful if elected officials were more like what Richard Thaler has called “Econs” – fully rational, high energy, completely informed decision makers (i.e., “economic man) with well-defined preferences whose choices are always consistent with their lifetime economic objectives.” But humans, including elected officials, are not “Econs.” Humans need reminders and contextual support in making decisions consistent with their objectives. Thus, the existing budget process is not ideal for use by humans, whose behavior more often resembles Homer Simpson than *homo economicus*. For humans, information and the opportunity to choose are often not sufficient to assure that choices will be consistent with expressed intent or long-term well-being. However, human decisions can be improved, sometimes significantly, by relatively small changes in the context in which decisions are made. We use some of the relevant findings of behavioral research to develop modifications to the federal budget process that seem most likely to assist policy makers in making better budget choices.

---

1 We thank Roy T. Meyers, who suggested this topic, and Steve Redburn and Paul Posner for comments and guidance. We also thank those who commented on an early version of this paper at the 2014 Annual Conference of the Western Economic Association International, Denver, especially Timothy C. Irwin, John D. Merrifield, and Barry Poulson.
Introduction

This paper is divided into two parts. The first section makes the following points:

- A broad consensus exists among policymakers, analysts, and the public that, given current debt levels, fiscal discipline is necessary for long-term fiscal, political, and economic stability, which is strongly preferred to recurring debt crises and instability that is otherwise projected to occur.

- Some persuasive evidence for this proposition consists of legislative actions taken over the last 40 years that have attempted explicitly to establish and strengthen a federal budget process capable of achieving that objective.

- A similarly broad consensus also exists that previous reform efforts now embodied in the current federal process do not appear likely to achieve this objective. The number of possible explanations for that failure is large, and but none seems particularly compelling or dominant.

In support of these points, the paper offers:

- A brief history of U.S. fiscal outcomes, highlighting the success of the balanced-budget “norm” in achieving long-term fiscal balance and the rise of deficits and debt after the loss of that regime.

- A description of successive, relatively unsuccessful budget reforms intended as a substitute for the balanced-budget norm, beginning with the Congressional Budget and Impoundment Control Act of 1974.

The second part of the paper, drawing on the experience described in the first part:

- Casts good judgment and wise caution aside, and offers an alternative explanation for the continuing gap between the goals of the current federal budget process and its result.

- Draws on the findings of behavioral research to argue that recent budget reform efforts are based on the strong neo-classical assumption that budget decision makers are super-rational, for whom making decisions consistent with preferences requires only relevant information and an opportunity to choose.

- Describes the view of human decision making that emerges from behavioral science, which suggests that humans need carefully designed choice architectures to avoid frequent decision errors that are inconsistent with their rational objectives.
Offers some modifications to the current budget process that the Congress might adopt to more closely align its legislative decisions with its long-term goals.

**Part One: A Brief Fiscal History of the U.S.**

For some time now, there has been widespread agreement that the federal budget process is “broken.” And, there is no shortage of proposals to “improve” or “reform” it (Gregg, 2006; Meyers, 2008; Peterson-Pew, 2010; Posner et al., 2012; Capretta, 2015). Most critics single out the deterioration in the federal government’s fiscal position as a major failing. Specifically,

- Until the Great Recession in 2008, federal debt (i.e., debt held by the public) as share of GDP averaged 35 percent for the previous 50 years.

- The effects of the Great Recession, combined with legislative actions to respond to it as well as continuing to fund war operations without offsets, increased debt to 74 percent of GDP, where under current law it will remain until 2021 when it will resume its upward path (CBO, March 2015).

- The government has also come close to defaulting on the public debt, failed to agree on an annual budget resolution during 2010-2014, shut down the government, and resorted to continuing appropriations for an entire fiscal year because it was unable to decide explicitly on a more appropriate allocation of fiscal resources.

The current fiscal position and policies of the US government expose the country to significant risks of a self-inflicted debt crisis and reduced flexibility to respond to future shocks, as many authorities have warned (CBO, 2014; Yellen, 2015). And yet the U.S. appears unable to address the rising debt or the attendant economic, political and asocial risks. To understand how we got into the current state where no one in authority wants to

---


Federal Reserve Chair Janet Yellen warned lawmakers the high level of federal debt in proportion to the U.S. economy might tie their hands if the economy were to be hit by another shock like the 2008 financial crisis. . . at a Senate Banking Committee hearing on monetary policy Feb. 24, Yellen said, “I think Congress has […] stabilized for a number of years the [debt-to-GDP] ratio. But eventually debt-to-GDP will begin to rise and deficits will increase again as the population ages and Medicare, Medicaid and Social Security get to be a larger share of GDP under current programs. . . I also worry that if we were to again be hit by an adverse shock, that there is not much scope to use fiscal policy. It was used in the early years after the financial crisis. We ran large deficits, but in the course of doing that, debt-to-GDP ratio rose,” she said. . . “Were another negative shock to come along, it's questionable how much scope we would now have to put in place even on a temporary, multi-year basis expansionary fiscal policy, and I think it's important to deal with these issues, for the Congress to do so.”
be, but from which a policy exit seems difficult to negotiate, we need to consider a bit of fiscal history.

It is helpful to divide American fiscal history into three Eras: Balanced Budget, Legislated Fiscal Regimes, and the Era of the “Two Santas” (Steuerele, 2014).

1. The Balanced Budget Era (1789-1970s)

It has not always been the case that Congress and the president tolerated fiscal drift and rising debt. For more than the first century and a half of U.S. history, the distaste for deficits and debt of the founding fathers produced a constraining “norm” for budgeting: the federal budget should be balanced annually, except in the extraordinary circumstance of war. And in that exceptional case, the debt incurred should be repaid as soon as possible. The effects of this philosophy of repaying debt soon after it is incurred are evident up until the early 1970s (with a temporary dip in 1998-2001) in the following graph:

![Graph showing U.S. debt history from 1790 to 2030](image)


(Note: The alternative fiscal scenario projection incorporates the economic feedback effects of fiscal policy as well as the continuation of certain current policies, even if they are scheduled to expire, rather than current law.)
From 1789 until World War I, Congress and the president agreed, though sometimes with difficulty, to control spending via appropriation bills. The role of the federal government was consciously limited, with total outlays typically not exceeding 5 percent of GDP, except in wartime. The balanced budget norm meant that the federal government’s debt followed a tight pattern of wartime build-up followed by peacetime debt retirement, even though there was not a central budget authority outside the Congressional tax and appropriations committees. Just after World War I, with the debt still large, Congress enacted the Budget and Accounting Act of 1921 requiring, for the first time, the President to submit a single consolidated budget for all federal agencies showing the spending and taxing decisions that the President requests for the upcoming year.

For the next half century or so, this process worked as follows: the President would submit his budget requesting appropriations and revenues, and then Congress would send appropriation bills (and sometimes revenue bills) that varied somewhat from the President’s request back to the President for his signature. Although the deficit for the fiscal year was not knowable until the fiscal year was over, the balanced budget norm continued to hold, now accompanied by the executive’s responsibility for developing a consolidated budget request consistent with budget balance.

Subsequently, the demands of the Second World War drove the U.S. debt to a record high of 106 percent of GDP. After the end of the war, the debt steadily shrank until the outbreak of Korean War. However, the post-World War II era also marked the weakening and finally the end of the balanced budget norm. James Barth and Tong Li (2012, their Figure 1) have summarized the effects of that shift in terms of a change in the frequency of budget deficits:

[After] 1789, when the U.S. Treasury Department was established, to [2014], there [have been 225 annual] federal budgets. The record lists [118] deficits and 107 surpluses …Breaking down that period, however, reveals a different picture. [Up until] the beginning of the Great Depression in 1929, the government recorded 46 deficits and 94 surpluses. [And from 1930 to 2014, there have been 72 deficits and only 13 surpluses.]

After the intellectual and social upheaval of the Great Depression and WWII, the information required to make budgetary decisions became more complex, and task of budgeting became more difficult (Penner, 2014; Schick, 2007). Policymakers’ gradual acceptance of the Keynesian counter-cyclical role for fiscal policy after World War II undermined acceptance of the norm of annual budget balance. As Schick (2007) puts it:

With the rise in revenue and spending as a proportion of GDP [in the 1960s], the president increasingly used the budget to fine-tune the economy—keeping it on a high-employment, low-inflation course—by adjusting the budget to stimulate or dampen aggregate demand. The notion that the budget should [be used to] balance the economy superseded the balanced budget norm. Deficits were acceptable to spur the economy to operate at full potential and to reduce unemployment. As long as the economy was booming, small deficits were seen as a prudent means of promoting national well-being (p.17).

The previously held balanced-budget norm had allowed for needed temporary exceptions in wartime, but it was understood that public debt incurred during those episodes would have to be repaid from budget surpluses at war’s end. In this way, the balanced-budget norm did not precisely constrain deficit spending in each and every year on an annual basis, but did produce the desired effect over the longer term.
The replacement fiscal framework that emerged in the mid-20th century was interpreted to mean that large deficits were appropriate during recessions, small deficits during expansions, and surpluses only during inflationary boom periods. Analysts encouraged policymakers to focus on the current state of the economy as the key to setting the annual target deficit. Retaining this short-term horizon meant the long-term public debt would simply become the sum of a series of deficits incurred to keep the economy operating at a high rate of employment. Gone was the imperative to pay for a run of deficits with subsequent budget surpluses.

The new fiscal paradigm provided, at best, a soft constraint on annual budgets because it left to policymakers the task of making judgments about the state of the economy and the appropriate size of deficit or surplus. In this way, the new paradigm weakened the norm of annually balanced budgets and, in turn, the ability of public officials to resist the temptation of immediate political gratification that comes from increasing spending and reducing taxes.

Under the new fiscal approach, the handful of surpluses and small deficits recorded after World War II gradually turned to persistent deficits with the beginning of the Vietnam War. President Nixon, though he did not produce balanced budgets, was effective in scoring political points against the “fiscal irresponsibility” of the Congress, which he used to justify the executive practice of “impounding” (i.e., not spending monies that had been appropriated by law), fueling public interest in reducing federal deficits. By the early 1970s, Members of Congress – sufficiently concerned about the increasing power of the executive branch over fiscal policy as well as recent increases in the deficit – began to consider statutory changes in the budget process (Schick, 2007). Those concerns culminated in the enactment of the Congressional Budget and Impoundment Control Act of 1974 (CBA).

The CBA was a rational, logical approach to improving budget decisions for macro-economic stabilization. It defined a process that would enable the Congress to consciously choose, with the assistance of budget analysts and economists, the total amount of federal spending, revenues, and budget deficit. It established responsibility for Who does What and When, by creating the House and Senate Budget Committees, CBO, and a logical, sequential decision process set out in a calendar schedule. With the CBA, Congress could rightly claim that it had acquired the capability to budget responsibly. The new process provided the Congress with access to much of the information that a rational, energetic policymaker would require to make informed fiscal choices for short-

---

3 “Nixon regarded changes in his budget proposals by the Democratic-controlled Congress as ‘undisciplined’ and ‘fiscally irresponsible.’ To a greater extent than his predecessors, he took it upon himself to impound funds that Congress had appropriated for programs that he considered unwise or unnecessary.”


4 For a detailed statutory history of the federal budget process see Lee, Johnson, Joyce (2008) Ch.9 or Mikesell (2014) Ch. 3.
run economic stabilization. In hindsight, the CBA appears to have been the first in a
series of statutory process-related efforts intended to help decisionmakers reach fiscally
prudent budget decisions. It was the first step in developing a process replacement for the
balanced budget norm and to reduce the chances of a reversion to ad hoc, sequential
fiscal decision making.

As noted by many, the new process was layered over the pre-existing process
(Dauster, 1993; and Rivlin, in Joyce, 2008). It also aimed to minimize changes in
existing practice in order to avoid alienating opponents who might have otherwise
prevented its enactment or could have hampered its launch. There were advantages to
this approach. For example, while providing an overall budget framework for fiscal
decisions, the CBA retained the practice of delegating specific decisions to committees
and subcommittees. This delegation of authority permits Members to specialize and
become expert in specific program areas.

But there were also disadvantages of layering the CBA on top:

- Mixing the old and the new processes added complexity and multiple decision
  points to the new process;

- The CBA also left committee jurisdictions balkanized by fiscal instrument (tax
  expenditures versus spending) and by spending account, rather than policy
  objective. As a consequence of such fragmentation, it is hard for legislators to
  trade off resources across more and less effective policies; and

- It took only modest steps toward lengthening the planning horizon. Over the first
two decades, the “budget window” incrementally grew from one year, to three
years, to five years, to seven years (only once), to the ten years that decision
makers have used intermittently in recent years. Although budget resolutions are
sometimes written for only five years (as that is all that the CBA requires), CBO
baseline projections and cost estimates of legislation are for 10 years.

While the CBA created a logical and orderly process for budgeting, it failed to
resolve ongoing concern about deficits and debt. Indeed, deficits continued to average
nearly 5 percent of GDP per year for 1982-1985 – levels not seen since 1946. Concern
about this failure prompted the Congress to enact the Balanced Budget and Deficit
Control Act of 1985, or Gramm- Rudman-Hollings. This Act attempted to reduce the
deficit directly by setting a series of annual deficit targets intended to reduce the deficit to
zero by 1991. Enforcement was to be provided by the now infamous, year-end
sequesters, or across-the-board cuts in spending, sufficient to reduce the deficit to the
target level.
Because actual annual deficits are affected by uncontrollable variables including the pace of economic activity, the deficit targets of Gramm-Rudman-Hollings were replaced as the primary fiscal enforcement tool in 1990 in the bipartisan enactment of the Budget Enforcement Act (BEA). That law set statutory caps on discretionary (or appropriated) spending and created a requirement that legislated increases in mandatory spending and reductions in tax revenues be paid for by offsetting changes in law (PAYGO); both disciplines were enforced by sequestration. BEA was designed to lock in deficit reduction agreed to in the Omnibus Budget Reconciliation Act (OBRA) of 1990 by the Republican president and the Democratic Congress.

The BEA’s PAYGO mechanism and discretionary spending caps, which were extended by subsequent legislation to 2002, combined with a robust economic expansion, resulted in steadily declining deficits over fiscal years 1993-1997. In 1998, the U.S. recorded the first of four consecutive budget surpluses. With the expiration of PAYGO and discretionary spending caps in 2002 and the onset of the 2001-2002 recession, however, the budget returned to deficit, where it has since remained.

3. The Era of Two Santas, 1997-present

Success in shrinking the deficit and ultimately achieving a budget surplus turned out to be a mixed blessing for fiscal discipline. On the one hand, the growth in public debt came to a halt and actually fell for the first time since the late 1960s. On the other hand, the problem of what to do with the surplus proved to be not much of a challenge for Congress and the Administration. Tax cuts and spending increases were embraced quickly with such enthusiasm that Eugene Steuerle in his book, Dead Men Ruling (2014), dubbed the period – the Era of Two Santas.

The prospect of budget surpluses also spurred creative energies to discover new applications for the process of reconciliation, which, until 1996, Congress had only considered using for deficit reduction. That is, the “elastic clause”\(^5\) of the CBA allowed its broad framework to subsequently be interpreted as allowing reconciliation instructions in the budget resolution and subsequent reconciliation legislation that would either (a) reduce spending in an omnibus fashion (1981) or (b) reduce taxes, thereby increasing the deficit (1996, 2003, and 2005) or decreasing the surplus (2001). \(^6\)

\(^5\) See section 301(b)(4) of the Congressional Budget Act: “(b) ADDITIONAL MATTERS IN CONCURRENT RESOLUTION.—The concurrent resolution on the budget may— . . .(4) set forth such other matters, and require such other procedures, relating to the budget, as may be appropriate to carry out the purposes of this Act”.

\(^6\) “Nobody dreamed of doing it through reconciliation then.” See page 102, and also page 88 of http://www.budget.senate.gov/republican/public/index.cfm?a=Files.Serve&File_id=3ef00980-b6ef-4a0b-b804-df74ef7b1486
The emergence of surpluses also played a surprising role in an intra-party dispute that prevented the adoption of a budget resolution in 1998 (for FY 1999) for the first time since the enactment of the Congressional Budget Act. The Republican majority in the House wanted to start enacting tax cuts immediately. But the Republican majority in the Senate wanted to wait and see how the Balanced Budget Act of 1997 would work out. They were not prepared after the fiscal path to balance by 2002 that was laid out in reaching agreement with the Clinton Administration. As a result, even though both bodies were controlled by the same party, 1998 was the first instance of a failure to adopt a budget resolution.

After the Treasury recorded a surplus of $69 billion for fiscal year 1998, CBO’s March 1999 baseline projected surpluses totaling $2.7 trillion over 1999-2009. With surpluses in the bank for the previous year and seemingly in hand for years to come, the Senate in 1999 agreed with the House on a budget resolution for fiscal year 2000 that included reconciliation instructions for a $778 billion tax cut bill over the 2000-2009 period. Because their former negotiating partner from 1997, who opposed such tax cuts, remained in the White House through 2000, the Republicans were not able to translate their budget resolution assumption of tax cuts into enacted law until 2001 – when the Republican Congress adopted a budget resolution for fiscal year 2002. Tax cuts, proposed by the new President, George W. Bush, were enacted (on a mostly partisan basis) via a reconciliation bill that amounted to an estimated $1.4 trillion over 2001-2010.

Similarly, when Democrats took control of Congress in 2007 and the presidency in 2009, they increased non-defense discretionary spending by about 10 percent each year without offsets. The onset of the financial crisis in late 2007 and the Great Recession of 2008 sharply reduced revenues and increased spending (e.g., unemployment insurance and other automatic stabilizers). Eventually, Congress enacted a counter-cyclical response of fiscal stimulus legislation. These efforts resulted in the deficit increase of nearly $1 trillion from 2008 ($459 billion) to 2009 ($1.413 trillion).

By early 2010, Senate Budget Committee Chairman Kent Conrad and other Democrats were sufficiently concerned about deficits and debt to insist that, in exchange for votes to increase the debt limit, the President agree to (1) sign into law a new version of statutory PAYGO and (2) appoint a National Commission on Fiscal Responsibility and Reform (Bowles-Simpson Commission). (See Appendix 1 for more details on the Statutory Pay-As-You-Go Act of 2010 and the Budget Control Act of 2011.)

As this review of budget process history indicates, the federal budget process for the past 40 years has been an on-going work-in-progress aimed at adjusting the process to

---

8 http://www.fiscalcommission.gov/
help decision makers improve their ability to limit federal deficits and debt. The success of this effort has been mixed, while the current fiscal outlook seems bleak, with a chance that condition will not get much worse for a few years.

At many points during the last 40 years, budget analysts have become frustrated with the effort. After all, the CBA provides decision makers with the information they need to make decisions and it affords them the timely opportunity to do so. The BEA, in combination with the three enacted deficit-reduction reconciliation bills (1990, 1993, and 1997) it was associated with, also serves as a model for how to “lock in” deficit reduction. It is understandable then, that some, including former CBO Director Rudy Penner, who is among the most quotable, would conclude the “process is not the problem, the problem is the problem.” And perhaps sotto voce, some might have muttered: “Just do it, dammit! What more could you need?”

In more reflective moments, however, we know that the current budget outlook and the potential losses from the risks the country is taking give us every reason to continue the search for a better process.

We now turn to that effort.

Part Two: Introduction to Behavioral, Experimental Economics

Around the same time the Congress was attempting to strengthen its budget processes, a seemingly unrelated line of research was flourishing in academia. Building on work from the 1930s, a number of scholars were engaged in the systematic empirical study of human decision making. They were especially interested in whether people acted rationally in pursuit of their interests as widely assumed, the nature of deviations from this assumption, and, ultimately, the circumstances under which people are likely to make better choices, defined in terms of consistency with their expressed preferences.

---

9 The frustration of many budget analysts and support staff with the failure of a political institution to produce results from what had seemed to be a more than adequate process was captured in Philip Joyce’s “Does more (or even better) information lead to better budgeting?: A new perspective” (2008), Journal of Policy Analysis and Management, 27, 4, 945-960). Joyce’s largely negative assessment was accompanied by comments of five respondents: Ken Apfel, Don Moran, Rudy Penner, Paul Posner, and Alice Rivlin. The commenters cited instances where information and analysis had mattered (e.g., Social Security reform in 1983 and tax reform in 1986, and the Budget Enforcement Act of 1990); argued for limited expectations from a rational process in a political arena where entitlement and tax reform is difficult, except “when Congress is giving away money”; and lamented the “layering” of the 1974 process atop pre-existing procedures, which resulted in more information but also excessive complexity with “too many actors and decisions points.” Regarding the potential usefulness of further budget process changes, some analysts seemed to share the sentiment expressed by Will Parker, in Rogers and Hammerstein’s Oklahoma!, about the state of progress in Kansas City: “they’ve gone about as far as they can go.”
In 1974, psychologists Daniel Kahneman and Amos Tversky caught the attention of many social scientists when they published “Judgment under Uncertainty: Heuristics and Biases” in *Science*. They argued, using experimental laboratory results, that people often use rules of thumb, or heuristics, to estimate probabilities of uncertain events, rather than systematic analysis. Subjects resort to crudely informed heuristics even in cases where they are given relevant information that could improve their estimates and increase their rewards from greater accuracy. Their finding that human decision making is inconsistent with the rational choice assumption was not the first study to reach this conclusion, but it extended that result by showing that errors in human decisions are systematic, allowing decision heuristics to be identified and generalized.

The influence of Kahneman and Tversky on the social sciences, especially economics, increased further with their 1979 publication of “Prospect Theory: An Analysis of Decisions under Risk” in *Econometrica*. That paper used experimental evidence to argue against the then generally-accepted theory of expected utility as a descriptive model of economic decision making under uncertainty. They also advanced an alternative, Prospect Theory. In their descriptive framework, people undervalue risky expected outcomes relative to certain ones and prefer to avoid losses more than they value equal size gains. This latter condition is often referred to as “loss aversion.”

Concurrently, economists Vernon L. Smith and Charles R. Plott were pioneering the use of laboratory studies of economic decision making by individuals. In successfully publishing the results of their research in leading economics journals, Smith and Plott established the legitimacy of experimentation as a useful methodology for testing economic hypotheses. In 2002, Kahneman and Smith shared the Nobel Prize in Economic Science.¹⁰

The work of Kahneman and Tversky, which contributed to the development of a sub-discipline, behavioral economics, has been extended and popularized by many, including Dan Ariely, Cass Sunstein, Richard Thaler, Reid Hastie as well as Kahneman’s own *Thinking, Fast and Slow*, included on the *New York Times* 2011 best seller list.¹¹ In various combinations (Thaler and Sunstein, 2008; Sunstein, 2013; and Sunstein and Hastie, 2015), these researchers have succeeded in developing many implications of behavioral economics for public policy. Their focus has been on identifying modifications to decision processes that support choices consistent with the authentic, ¹⁰ Kahneman’s *Thinking, Fast and Slow* was written “In memory of Amos Tversky,” who died in 1996; Smith praised Plott’s work and acknowledged his contributions in his Nobel lecture.

¹¹ Herbert A. Simon (1916-2001) also has a claim to partial parenthood of behavioral economics. Simon’s studies of decision making by individuals and firms led him to conclude that decisionmakers do not optimize across all choices because they have limited cognitive skills and information which render optimizing too costly to attain. Instead of optimizing, decisionmakers stop searching for the optimum solution when they identify one that is satisfactory, a process he called “satisficing” (1956). Simon also coined the phrase “bounded rationality” to refer to the limited human ability to make optimal decisions. He was awarded the Nobel Prize in Economics in 1978.
often long-term preferences of an individual making the choice, such as saving for retirement and practices that promote good health. These process modifications do not constrain or regulate behavior; rather they operate through simple reminders, information frames, and contextual cues, without coercion.

In this paper, we use some results of behavioral research to diagnose human failings associated with use of the current federal budget process – defined as decisions that are inconsistent with the deliberate and reflective intent of lawmakers – and to propose a series of changes in “choice architecture” of the budget process that Congress might adopt to improve its decisions.

**Key Findings of Behavioral Economics**

Kahneman (2011) uses the fiction of two separate human systems of cognition (System 1 and System 2) to summarize recent efforts “to identify and understand errors of judgment and choice.” System 1 is automatic, quick, requires little mental energy, includes a working model of our world, and is the origin of much error; but it also is responsible for much that we do well, such as recognizing danger or carrying out repeated actions like driving a car or playing a skill sport. System 1 is not effective at self-control, planning, or complex analysis. System 2 is effortful, energy intensive, slow, better at making important choices and judgments and is essential for self-control. But System 2 often endorses or rationalizes ideas that originate in System 1. While System 2 can improve decisions and behavior, its analytical ability is limited and it often lacks important, relevant information.

Humans frequently make poor decisions when System 1 makes a decision that is really better suited to System 2. For example, when on my way home from a meeting with a financial planner, I – consistent with my long-term savings plan – refuse to buy a car for $60K, but then change my mind after a salesman explains that I will make no payments for the first year of ownership (including regular service costs such as oil changes), and then pay only $19.99 per day (for the next 11 years) – less than I would pay for an inexpensive rental car.

Humans also make poor decisions when System 2 is turned on while System 1 should be working instead. For example, Sunstein retells the story of Larry Bird asking an opponent if he has “changed his shot because it looks different” and mentioning that the balls used at this site seem to “feel slippery.”

The fundamental result of behavioral science is that humans are neither “Econs” (*homo economicus*), nor the uber-logical, science-fiction equivalent, Mr. Spock. At the very least, they are not reliably rational in decision making. This result is at odds with the assumption of rational choice, found in virtually all introductory economics
textbooks. It thus suggests the extent to which that model is an idealized simplification that needs to be modified and extended to explain many human decisions.

In general, humans often act impulsively with little regard for their longer term consequences and in ways that are inconsistent with their previously-expressed preferences and plans. We consume more and less healthy food than is consistent with our dietary goals; we save less than is needed to achieve our long-term financial objectives; we make purchases without regard to the larger benefit per dollar of cost we could obtain from alternative goods and services. When confronted with complex decisions, instead of thinking systematically, we tend to rely on recent and limited experience with similar choices and events, or follow the example of others. Our decisions also vary with our emotional state or level of fatigue. In a phrase, our decisions vary with the context in which we are making decisions, and “everything” affects context.

Arguably, these findings confirm what we already knew. For millennia, humans have been aware of the experiential gaps between intent and action. And despite our many other failings, we have often managed to adopt practices and processes that can better align decisions with well-being. Avoiding temptation, habitually taking stairs rather than elevator, saving the seed corn, and entering into forced saving contracts (e.g., cash-value insurance and holiday savings accounts) are examples of means by which humans have attempted to resist self-defeating behavior by limiting their range of future choices and modifying the context in which they make decisions.

**Individuals – Systematic Behaviors and Decision Errors**

The gain from better understanding human decision making is not so much its novelty as its potential to improve decisions by identifying systematic errors, their associated causes, and conditions under which those errors tend to be avoided. Here our interest is specifically limited to decision errors that appear to occur in public budgeting and that are inconsistent with the goal of long-term fiscal stability and efficient use of public resources. For that purpose, we focus on three individual behaviors that lead to related decision errors: (1) limited self-control and an associated present bias toward “gain now, pain later,” (2) limited information and processing (cognitive) capability, and (3) a desire for social acceptance that inclines us to seek and retain membership in a tribe or “herd.”

---

12 Behavioral science has produced numerous consumer innovations: the connection that prevents gas caps from being left atop the pump; the beep that signals a card left in an ATM; smaller plates in better restaurants; and, instantaneous feedback on the Prius dashboard about the effect of our driving speed and style on gasoline consumption. The prevalence of prices that end with “99” appears to have developed without the aid of randomized controlled trials, however. Relatedly, behavioral patterns and corrections are often identified by considering how “ordinary people” think, according to Raj Chetty, a Harvard economist. Kahneman (2011) also attributes many insights to conversations with Tversky that were essentially mental experiments.
We emphasize that these behaviors are closely related, can be interactive, and at times are difficult to distinguish from one another.

1. **Limited Self-Control and Present Bias.** Widespread agreement exists among researchers and their human subjects that people have “… a tendency to pursue immediate gratification in a way that their ‘long run selves’ do not appreciate” (O’Donoghue and Rabin, 2001). That is, we have “present-biased preferences.” Especially when System 1 is making decisions, people tend to choose short-term benefits at a disproportionately high future cost and to defer actions now that would produce relatively large future benefits. We are especially prone to overconsumption now if the cost of that action is not salient – for example, when the easy availability of credit enables us to shift cost into the distant, heavily discounted, optimistically anticipated future. In time, sometimes quickly, we recognize that those decisions are inconsistent with our goal of long-term well-being.

Some scholars have attempted a deeper understanding of this welfare-reducing bias. One prevalent view is that inter-temporal choice is heavily influenced by the degree to which a person feels connected to his or her future self. “To people estranged from their future selves, saving is like a choice between spending money and giving it to a stranger, years from now” (Hershfield et. al. 2013). Following this perspective, Hershfield and his associates found that by showing subjects age-progressed photographs of themselves, they could induce an increase in willingness to “accept later monetary rewards over immediate ones.” (Also, see Sunstein, 2013, pp. 55-57.)

The factors that affect inter-temporal choice decisions, including those made in System 2, are multiple and complex and appear to include the linguistic structure of the language we speak. For native speakers of English and other languages that require the distinctive marking of future events from the present (e.g., “it is raining today” vs. “it will be raining tomorrow”), the challenge of acting today consistent with our future plans is made more difficult by the forced separation between now and later. By contrast, “…languages that grammatically associate the future and the present [such as German, Finnish, Japanese, Mandarin: e.g., “it is raining today” and “it is raining tomorrow”], foster future-oriented behavior” (Chen, 2013). “Empirically, I find that speakers of such languages: save more, retire with more wealth, smoke less, practice safer sex, and are less obese. This holds both across countries and within countries [where two grammatically different languages are spoken] when comparing demographically similar native households.”

2. **Limited Information and Processing Capability.** A central tenet of behavioral studies is that humans have a limited ability to focus their attention and exert mental effort. Both are scarce resources with alternative beneficial uses (Kahneman, 2011, Ch.2). Many people, much of the time “…apparently find cognitive effort at least mildly unpleasant and avoid it as much as possible.” Instead, we rely mostly on our automatic, or System 1,
thinking, which is effortless and quick in producing “impressions, intuitions, intentions and feelings” but is unsuited for solving complex problems or exercising self-control.

On its face, effective budgeting is one of the more complex series of decisions that humans are called upon to make, which may explain why so few of us do it explicitly. If we take the simplest case of a single individual attempting to solve the problem of allocating her scarce resources during some planning period so as to obtain something in the neighborhood of maximum benefit, the task is a daunting one. She must simultaneously decide how much income and other resources she is likely to have available; the expected value to be obtained from consuming each alternative good or service at varying levels of intensity, the cost of units of each good and service, and how much consumption she should shift to or borrow from the future.

Public budgeting, done well, is more difficult. The first-order choice is to determine the share of national output to be devoted to public rather than private use. A substantial literature in public finance offers conceptual guidance as to how resources should be allocated to address market failures for public goods, external costs and benefits, departures from the ideal of highly competitive markets, and common pool resources, among others. But these guides are mostly abstractions: the public-private choice is made by a collection of elected individuals whose behavior does not match the ideal assumed in the public finance literature. There is also the closely related, thorny problem of allocating the burden of financing government across the current population.

The second big set of decisions to be made, simultaneous with the first is: how shall the portion of national product allocated publicly (i.e., by government) be distributed across various beneficial uses, including redistribution of income? In concept, systematic cost-benefit analysis can be helpful, but the task of compiling and then comparing those values across a large number of potential alternatives tests the analytical prowess of many, even the best and brightest. The third allocative set of decisions is no easier: what share of the costs and benefits of the government’s use of national output today shall be allocated to future generations? We might have hopeful expectation that future generations will have higher incomes than present ones, so that some share of current costs might be fairly shifted forward, but that is not a certainty.

Budgeting is inherently a System 2 activity. But System 2 has limited analytical capabilities and requires scarce mental energies and costly information. One general rule therefore for the design of any decision process is: “Reduce the strain on System 2” (Sunstein and Hastie, 2015). Make the process easier by simplifying wherever possible:

- Frame important relevant information so that it is salient and vivid to reduce the cost of processing. This includes using descriptive names for budget variables, e.g. deficits are costs incurred in the current period to be paid in future periods. A direct link exists between more precise language and better decisions (Kahneman, 2011, p. 418).
• Provide budget makers with the most relevant, summary measures of costs – those that are not already sunk and are controllable in the budget planning period. Do not expect policymakers to read the caveats in the text of cost estimates or to understand and see through potential distortions from prevailing budgetary accounting rules. Make the bottom-line numbers the relevant numbers for the function of budgeting.

• Summarize the results of program analyses and evaluations in easy-to-understand language and make use of this information essential to the decision process. One way of getting people to switch from System 1 to System 2 is to ask them a question that can’t be answered in System 1. For example, how does the dollar cost of this proposed activity compare with the estimated value of benefits? (Sunstein. 2013 pp.151-5 endorses the use of cost-benefit language as a means of triggering the switch to System 2.)

• Use defaults to facilitate “temporary,” reversible decisions. For example, in 2001 there were serious proposals to include in the tax cut law a provision that would “trigger” the tax cuts off if it turned out that the future budget surpluses that were projected at the time did not actually materialize; instead that law fixed the lower tax rates in place for 10 years regardless of the size of surpluses or deficits over that period. One way to use a default would be to add such a reversal feature to all spending increases and tax cuts.

3. Inclination to Herd. A classic behavioral case study is the Robbers Cave experiment, conducted in the early 1950s by Muzafer Sherif and his associates at the University of Oklahoma. This carefully documented experiment\textsuperscript{13} demonstrates the tendency of humans (in this instance, 11-12 year old boys from white, middle class families, randomly assigned to two separate groups at a camp in Robbers Cave State Park, Oklahoma) to:

• Bond into named groups through shared experiences, with their own colors, mascots and leaders.

• Develop an intense hostility, verging on violence, toward any other group of like individuals with whom they compete for scarce resources.

• Reconcile with the “enemy” under the mediating influence of shared experiences and challenges, which required the cooperative efforts of both groups to resolve.\textsuperscript{14}

\textsuperscript{13} Details at http://psychclassics.yorku.ca/Sherif/
\textsuperscript{14} Full disclosure: this reconciliation required adult leaders to provide contrived external threats and joint challenges of appropriate difficulty.
Membership in a group often modifies the decisions of individuals, and not always for the better. The idea that desire for continued acceptance influences our judgments and decisions has been robustly confirmed in a variety of experiments, many of which are summarized in Thaler and Sunstein (2008), Sunstein (2013), and Sunstein and Hastie (2015). The desire to maintain one’s standing with the group inclines the individual toward conformity and following the lead of others. The observed tendency away from independence and critical thinking by individuals is regarded by Sunstein and Hastie as among the most destructive behaviors for public and private institutions.

**Groups – Systematic Behaviors and Decision Errors**

Public budgeting is group decisionmaking. Some grounds exist for expecting that a group of elected officials would arrive at better decisions than individual members acting alone. Kahneman (2011), for example, suggests that organized groups, because of their reliance on routine processes designed to require the accumulation of relevant information, reflection, and the review of decisions by others, ought to be able to improve on the judgment and choices made by individuals. He notes, however, that few organizations make much effort at putting quality-control mechanisms in place and often neglect training in the conduct of efficient, decision-making sessions. Sunstein and Hastie further develop the idea of a lack of attention to process design and argue there is scant evidence that group decisionmaking avoids the errors observed in individual decisions (Sunstein and Hastie, p.13). Indeed, they conclude that groups are prone to worse decisions.

They offer several related group dynamics to explain this result: susceptibility to polarization; the power of the group, and especially early speakers, to restrict the expression of alternative views by creating a flow toward quick agreement ("cascades" of opinion); and the associated reliance on commonly-held ("shared") information. Kahneman (2011) also observes a group tendency toward overoptimism, which he refers to as the "planning fallacy." In particular, the phenomenon of "cascades" occurs in many of these group behaviors. Its importance makes it worthwhile to include Sunstein and Hastie’s definition (p.63) here:

A cascade occurs when people influence one another, so much so that participants ignore their private knowledge and rely instead on the publicly stated judgments of others. . . .[T]here are two kinds of cascades: informational and reputational. In informational cascades, people silence themselves out of respect for the information conveyed by others. In reputational cascades, people silence themselves to avoid the opprobrium of others.

1. Polarization. Even within a single herd, people are likely to hold different initial preferences. Those differences might include dimensions that are unimportant in a public
policy setting, such as the intensity of their like or dislike of broccoli, and those that are more critical to public budgeting, such as the appropriate reach of government into private life or markets. Yet, as may be obvious from recent experience, when members of rival herds are sorted and selected on the basis of an initial preferred view on such an issue, polarization can be a serious threat to effective and efficient public decisions.

Sunstein and Hastie also identify conditions that can lead to extreme polarization. They conclude, based on the “Colorado” experiment (Schkade, Sunstein, and Kahneman, 2000) (summarized in Sunstein and Hastie. 2015, pp 81-83) and others, that when like-minded individuals deliberate or “caucus” together, they are likely to adopt a more extreme view than they had as individuals before they talked together. In this experiment, a small sample of residents were recruited in two cities, (a right of center group from Colorado Springs in one group and a left of center sampling from Boulder in another). Each group was asked to first record their anonymous views on three contested policy topics and then to talk together to reach a group decision. After the group deliberations were concluded, individuals were asked again to record their views. Both the group decision and the second set of anonymous statements of individuals were more extreme than the initial statements.

Sunstein and Hastie offer three explanations for this. The first is that when like-minded people talk together on a policy question, the number of speakers supporting the view to which the group was predisposed will vastly outnumber those opposed, and some of the arguments offered will be new to the group. The frequency and novelty of the “evidence” on the question will move the group toward a more extreme position. Second, having heard the views of other members, individuals seeking the approval of the group are likely to adopt a stronger position along the spectrum favored by the majority. Third, when others corroborate our view, we are likely to hold it with greater confidence. This reduction in doubt about the validity of our position disposes us to accept a more extreme version of the policy.

Group affiliation can also blind members to the substantive content of a policy proposal (System 2 yields to System 1). Cohen (2003) reports a series of experiments in which self-described liberal and conservative students are randomly assigned to read one of two alternative, detailed proposed welfare reform policies, which are summarily characterized as either stringent or generous. Subjects were also randomly assigned to receive one of two different pieces of information: either that mostly Democrats favor the proposal (95 percent of House Democrats and 10 percent of House Republicans favor it) or that mostly Republicans favor the proposals (95 percent of House Republicans and 10

---

percent of House Democrats favor it). Each student was then asked to express support of or opposition to the policy.

The result: liberal students supported the assigned policy, whether stringent (less generous that any current proposal) or generous, if told that Democrats favored it, and opposed it otherwise. Similarly, conservative students supported both generous and strict versions, when told Republicans supported that version. The substantive content of the two proposals had little effect on their assessment of the policy.

2a. Loss of diversity of observation and dissent. The power of the group over individual judgment is not restricted to undergraduates. In an earlier study, Sherif placed adult subjects in a darkroom where they were asked to estimate the distance traveled by a stationary pinpoint of light. Even though the light does not move, most people experience the illusion of movement. Estimates ranged from two to six inches. After a number of individual trials, two or more subjects were added to the same room and asked to agree on an estimate. Both high and low estimators tended to compromise toward the average. Subjects were then asked if their estimate was affected by the views of the other observers. They denied such an effect. However, when they were retested alone, their estimate moved toward the previously established average of individual trial observations.

Group pressure can also change an individual’s expressed view, even if that view is inconsistent with the subject’s verifiable observation. Some instances are reminiscent of the Hans Christian Andersen tale of the “Emperor’s New Clothes,” where the sober, commonsense judgment of many individuals is replaced by faulty “group think” fostered by actors whose pitch to the group is essentially – “Who are you going to believe? Me, or your own eyes?”

Solomon Asch (1955) tested for answers to this question using an experiment, ostensibly dealing with “visual judgment.” In the test, a subject is shown two cards: one card has three lines of unequal length, and the second card has a single vertical line that matches the length of one of the lines on the other card. The task is to match the lines of equal length. The differences in length were significant so that ordinarily the error rate of subjects in identifying the matching line is less than 1 percent. The test is administered to the subject, along with six to eight others who, unbeknownst to the subject, are collaborators with the researcher. Their role is provide answers aloud before the subject and, initially, to provide unanimously correct answers. After two rounds of

16 http://www.intropsych.com/ch15_social/sherif_1936_group_norms_and_conformity.html
answering correctly, however, the collaborators unanimously report false answers for the next 16 rounds. The object is to test the extent to which the subject will report what his eyes tell him or conform to the overwhelming majority view.

On average, subjects tend to yield to the judgment of the majority and report a false answer about 36 percent of the time. About one-fourth of subjects never conform and continue to report correctly through the 18 rounds of the test. Those who begin to follow the majority almost never revert to their own judgment. After the test, all subjects who yielded to the majority’s judgment underestimated the frequency with which they followed the majority. In a controlled variation, experimenters inserted a “truthful partner” onto the panel. The presence of a supporting dissenter reduced the frequency with which the subject conformed with the researcher’s collaborators to about only 9 percent of the time.

In related studies, the role of the early-speaking majority is assumed by a confident, credentialed “chair” who speaks first and directs the group toward a specified decision. Such a voice has disproportionate influence on the group’s decision, even if it is based on error, because those with information that could derail the leader’s recommendation do not want to slow the group’s progress and take the time of busy people or to appear as one who is not a team player. Thus, the early speakers in a group decision process often create a “cascade” of support for a particular proposal that will prevent the group from hearing information held by self-silenced members.

2b. Reliance on Shared Information. One purpose of using groups to make decisions is to increase the pool of potentially valuable information used in the decision by combining the information held by individual members of the group. As noted in the study on estimating the distance traveled by a point of light, an overwhelming majority (“cascade”) can prevent some unique and relevant information from being heard. In fact, for many group decisions the most important information is shared, commonly-held information, i.e., “what everybody knows” (Strass and Titus, 1985; Gigone and Hastie, 1993). This result can be prompted if, for example, a group leader says: “We’ve addressed this question before. I don’t know of any reason for changing what made sense then.” In fact, for ease of agreement, nothing is better than reliance on commonly-held System 1 information. By contrast, processing unique information possessed by one or a few members would require the group to resort to an effortful System 2 analysis of the new information. Thus, the use of shared knowledge in decision making defeats one of the purposes for which the decision was referred to a group.

3. Overoptimism. The self-silencing behavior of group members with relevant information may also help explain what Kahneman and Tversky call the planning fallacy, or the tendency of group plans to be overly optimistic. They attribute this error to the cognitive ease of imaging the success of a plan instead of visualizing the myriad adverse
events that could prevent its realization. Most group plans are much closer to best-case scenarios than realistic expectations.¹⁹

**Proposals for Behavioral Changes in the Federal Budget Process**

*A Practitioner’s Guide to Nudging* (Ly, et al., 2013) suggests that the “first step in the process of designing an effective nudging strategy is to audit the decision-making process.” See Appendix 2 for such an audit for three cases.

Having identified six human behaviors (three affecting individuals and three affecting groups) that seem likely to retard the realization of more disciplined, sustainable, and efficient fiscal and budgetary policies, we now propose specific changes in the current process that might nudge decision makers in the direction of their intent for fiscal balance. We list these proposals by specific error-inducing condition: present bias, limited cognition, and so on.

We use this organizing device reluctantly because it may suggest that behavioral science has identified specific offsets or “cures” for each of a long list of departures from rational choice. No such general one-to-one correspondence exists between decision errors and specific features of choice architecture. Human behavior is the result of interactions among multiple parts of the brain and is affected by the entire choice environment. An adjustment in context or decision frame may affect the propensity to commit a variety of decision errors.

**Strategies for Improving Individual Decisions**

Kahneman’s framework suggests that to improve budget decisions it is necessary to address both System 1 and System 2 errors. For System 1 errors, the general rule is “to recognize that you are in a cognitive minefield, slow down, and ask for reinforcements from System 2” (p. 417). For example, if I had been able and willing to delay a decision on the $60K car purchase, System 2 might have told me that the present value of the extended payment option was no better than the cash price. Weakening the linguistic and cognitive separation between present and the future could also increase our connection to our future selves. It also is sometimes possible to use the influence of others to improve our decisions.

¹⁹ For evidence of optimistic bias in the Congressional Budget Resolution, see Bhatti and Phaup (2015), forthcoming in *Public Budgeting & Finance*. 
1. Present Bias

Using Goals and Feedback. One of the most reliable means of improving performance with respect to any objective is to establish meaningful quantitative goals and to solicit frequent feedback from credible observers (Thaler and Sunstein, 2008 pp. 99-100). As a means of offsetting present bias in public budgeting, one such change would be for the Budget Committees to establish 5-, 10- and 15-year goals for debt/GDP and other measures of fiscal balance in the budget resolution.

CBO could monitor legislative performance with respect to those fiscal goals and give public feedback. Performance assessments should note actions taken that are consistent and inconsistent with the goals as well as the absence of sufficient affirmative steps to assure their realization. The reports should be more visible and salient than current CBO scorekeeping reports prepared for use by the Budget Committees and should include tabulations like those in CBO’s March 2015 report, Legislation Enacted in the 113th Congress That Will Affect Mandatory Spending or Revenues.²⁰

Adopt More Salient Descriptions of Deficits and Debt. The U.S. government is regarded as one of the best credit risks in world. As a consequence, its debt securities trade in liquid markets at interest rates that are usually regarded as close to “risk-free.” This easy access to credit encourages the System 1 impulse to enjoy now and pay later. The easy option to borrow seems at times to be viewed as a permanent alternative to taxation. Yet, our System 2 cognition knows that every dollar that government spends must be paid for by someone eventually. Borrowing is a way of postponing taxation – an exchange of debt now for taxes later. This fact could be made more salient by referring to debt as “outstanding claims on future tax revenues” and deficits as “current period use of future taxes.”

Draw Future Total Costs of Choices Closer to the Present. The current budget process mostly uses intermediate projections of budget deficits and ratios of outstanding debt to national income for specific years to inform policymakers and the public of the future consequences of current policies. However, at least for System 1 cognition, those projections suggest the existence of a potential future problem, rather than a problem today. The harm from this depiction is that it is too easy to ignore options for dealing with excess borrowing of future resources; if no action is taken to rebalance consumption today with income until the bills come due in the future, then those bills will have to be paid from immediate reductions in consumption in the future.

An alternative means of reframing fiscal imbalance as a present rather future problem is to give greater emphasis to the “fiscal gap” while assigning less importance to

²⁰www.cbo.gov/publication/50052
annual budget deficits. Fiscal gap uses the difference between the present value (stocks) of future expected revenues and future expected outlays under current law to measure multi-year, longer-term fiscal imbalance. Fiscal gaps are now routinely calculated for 25-year, 75-year, and infinite horizons by relevant federal agencies and are required information under Generally Accepted Accounting Principles (GAAP), or standards issued by the Federal Accounting Standards Advisory Board.

A major advantage of routine use of fiscal gap is that it includes a widely-used measure of the long-term budget constraint – the present value of expected revenues under current law.\(^2\) The disadvantage of all present-value calculations is that judgment is required in projecting key variables such as the growth of national income. This leaves space for opportunistic “manipulation” of the numbers by political authorities, much as “rosy scenarios” are used to strategically reduce the projected deficit in specific years for a proposed budget in the current process. One way of limiting such abuse is to enable a number of capable authorities to prepare such measures independently and to require a reconciliation of differences. While errors in long-term projections are likely significant, projections can be routinely re-estimated and corrected over time, as policy is adjusted gradually.

**Draw Future Costs of Individual Programs with Deferred Payment Closer to the Present.** Not only can use of present value estimates improve the salience of long-term fiscal imbalance in aggregate, it may also be a more effective budgeting tool to bring the future costs of individual deferred payment programs closer to the present. In particular, current budgetary accounting for many social and financial insurance programs separates in time the recognition of cash inflows, which usually occur first, from the program’s corresponding cash outflows. Humans may be misled by the earlier inflows of cash, which can be misinterpreted as an increase in budgetary resources. This may incline decisionmakers to spend monies already committed to another purpose. The CLASS Act, discussed in more detail in Appendix 2, is one example of a program explicitly intended to be operated on an actuarially sound basis. But the CBO cost estimate of the bill (the Affordable Care Act) that included the new insurance program credited the whole bill with billions of dollars of deficit reduction over the first two decades of operation of the CLASS program. Social Security, Medicare, defined-benefit pension plans, and deposit and pension insurance are afflicted with this same potential for distorting the cost of current-period budget decisions.

The technical source of difficulty is that these programs are accounted for using on-budget trust or revolving funds that are authorized to spend fund balances for earmarked purposes without further legislative action. Thus, in most periods, when cash inflows exceed cash outflows, the program appears to be a net source of budgetary

\(^2\) The use of this long-term measure of fiscal imbalance also affords policy makers a feasible means of restoring fiscal balance while avoiding a behavioral barrier to reductions in scheduled, mandatory benefits, loss aversion, by restricting reductions to new entrants in the deferred payment program.
resources that may be used to finance other public activity. Of course, for programs intended to be self-financing or “actuarially sound” using only earmarked taxes or premiums plus interest on those balances, the net present value of inflows and outflows is expected to be zero.

An important exception to this practice is the current budgetary treatment of federal loan guarantees. Before 1992, loan guarantee programs were credited with the inflows of fees received when loans were extended, without any recognition of the cost to be borne for claims for default later. Since enactment of the Federal Credit Reform Act of 1990, the net expected cost of guarantees has been shown as budget outlays when guaranteed loans are extended. This prevents the budget from showing large initial gains in budget resources followed by large, unexpected losses. And, it does so without diminishing the ability of the government to account for all cash received and disbursed. A similar change in the budgetary treatment of other programs with deferred payments that are financed from current earmarked taxes and fees would make cost estimates (and baselines) related to such programs more salient depictions of their contribution to fiscal imbalance.

A related distortion arises under current practice for programs that are intended to be fully financed with earmarked taxes but whose funding is inadequate to pay benefits in the budget projection period. These include the Highway Trust Fund (HTF), Social Security disability insurance (DI) trust fund, Medicare hospital insurance (Part A) trust fund, and the Social Security Old Age and Survivors Insurance trust fund.

The issue stems from the rules governing the construction of the baseline (section 257 of BBEDCA). These rules require CBO to assume that the benefits calculated according to formula for each program under current law (or, for the HTF, as enacted for this year’s highway program) continue to be paid (1) even if the trust fund does not have sufficient resources to support those payments and (2) even though the law governing each program requires the agency administering the program to reduce payments to beneficiaries so that total payments do not exceed the resources available in the trust fund.

For example, the HTF collects roughly $30 billion per year in gas taxes but the baseline rules require CBO to assume that it will spend $40 billion per year, based on the most recent enacted spending authority. With no accumulated surpluses or balances, the HTF is expected to be insolvent in the summer of 2015. In that situation, the law requires the Department of Transportation to delay sending checks to states for the federal government’s share of highway projects already incurred until additional gas tax revenues come in. Absent any change in law to remedy the situation, outlays will be lower, yet the baseline for the year assumes that spending will be higher and “on time.”
For another example, CBO estimates that by early 2017, the DI trust fund “will be exhausted.”\(^{22}\) According to law, the Social Security Administration will be required to reduce every DI benefit check from the formula level so that total spending does not exceed the income available to the fund that year. Nonetheless, the baseline rules require CBO to assume that benefits are not reduced.

Because the cost of legislative action is scored against the baseline, the effect of this rule is to permit Congress to change current law and to spend more than is permitted under current law, but for that action to have a scored cost of zero. An obvious solution would be to change the baseline rule.

**Consideration of Costs and Benefits by Proposers.** Views and Estimates reports are current used by the other committees to inform the Budget Committees of plans for the coming budget year and for possible inclusion or exclusion of associated costs or savings in the Budget Resolution. The content of these reports could be expanded to require the inclusion of (1) a specific recommendation for funding any proposed new activity, including possible reductions in existing programs, earmarked fees or taxes, or an increase in the income tax and (2) expected program benefits as estimated by an identified source. These additional requirements would be difficult to meet using System 1 and could prompt early and systematic consideration of both the costs and benefits of the proposed changes in policy. The revised Views and Estimates report should be easily accessible on the Budget Committee website. Public comment and alternative views on the policy proposals and estimates should be encouraged and facilitated.

**Slow Down System 1 Decisions.** The rules of both the House and Senate could be amended to require a delay of one legislative day or more to waive a point of order raised against legislation that would violate the budget resolution. As for the budget resolution itself, there should be a new point of order against a budget resolution that increases spending or reduces revenues relative to the baseline. Congress has already taken a small step in this direction by establishing in the Budget Control Act of 2011 (BCA) a supermajority (in the Senate) point of order against any budget resolution that includes discretionary spending levels that exceed the statutory caps set out in the BCA through 2021. But as the case studies of the Medicare prescription drug bill and the highway program illustrate (see Appendix 2), it is currently far too easy to have the budget resolution assume higher levels of mandatory spending or lower levels of revenue than current law. Detailed records of votes cast for and against waivers of budget points of order could be posted on Budget Committee websites.

---

2. Limited Information and Processing Capability

Reduce the Strain on System 2 – Simplify. The process changes suggested for reducing present bias in budget decisions could also ease the burden of deliberative thinking. For example, avoiding the false impression that an inflow of new fees or taxes already earmarked for a specific purpose (e.g., gas taxes or insurance premiums) provides additional resources available for other unrelated purposes, could simplify the decision process for policymakers. They would not, for example, be compelled to understand the detailed explanations of results in CBO cost estimates. Instead, CBO could use its best judgment to show the most relevant numbers on the bottom line that would summarize the relevant effects of the bill without distortion.

The use of the “language” of cost benefit analysis could also make it easier for policymakers to switch from System 1 to System 2. Sunstein reports good results, for example, in asking proponents of policy change to enumerate and offer estimates of the costs and benefits of that proposal, during his stint as Director of the Office of Information and Regulatory Affairs in the Office of Management and Budget.

For some decision makers, present values of strings of annual cash flows are easier to keep in mind than the underlying annual values. For others, dollars per year per household are easier to work with than billions, as illustrated by their effective current use (Eugene Steuerle, and the Committee for a Responsible Federal Budget, CRFB).

Increase the use of defaults where decision are difficult because of the complexity of alternatives, if there is a clear advantage to one option relative to no action. Two possible candidates are: automatic continuing resolutions as an alternative to government shutdown and the CRFB recommendation for increases in the debt ceiling consistent with the budget resolution but via reconciliation for both the House and the Senate (building on the “Gephardt rule”).

3. Inclination to Herd

Herding is a deeply ingrained proclivity of humans that is difficult to restrain. Since herding is quite literally a spatial phenomenon, one could start there. Herding by party could be diminished by eliminating the current practice of physically separating the parties in the chambers as well as in hearing rooms. (Some members have already started to experiment with this approach on their own initiative, albeit but once a year, by pairing with a member from the other party to sit together at the president’s State of the Union Speech). To further increase connections across groups, competitions for charity (following the examples of the Congressional basketball and hockey games) could be

24 http://crfb.org/blogs/debt-ceiling-returnsand-10-ways-reform-it
social occasions pitting Members against Staff, committee against committee, House against Senate, or Legislative against Executive. The Budget Committees could be an ideal place to experiment with innovation. The Budget Committees are different: they are “new,” and they do not have jurisdiction over “substantive” or program legislation. They could justify some “out of the box” experimentation with respect to seating.

**Strategies for Improving Group Decisions**

Extreme polarization, the loss of vital relevant information and the associated reliance on “shared” common knowledge, and excess optimism are all barriers to informed decisions and judgments by public authorities. Modifications to the group decision process can improve group decisions if some of those pitfalls can be avoided. Process changes that (1) permit individuals to seek and maintain membership in more inclusive tribes, (2) promote the expression of individual preferences and observations that make genuine dissent an obligation, (3) trigger System 2 into action, and (4) restrain over optimism – all have the potential to improve group decisions. The general strategy for avoiding poor group behavior is to build into the decision process a means of eliciting information from all members of the group before committing to a limited range of alternatives.

1. Polarization

Political science research has confirmed that polarization in Congress has increased, though agreement regarding cause is not as complete. Nevertheless, tools are available to reduce polarization if they can be brought into the deliberative process.

For example, many members are admirers of private sector practices, suggesting that government has much to learn from the business example. One widely observed practice in those sectors is the use of retreats – taking time for more shared reflective thinking and relaxed dialogue. These events often include outside facilitators whose expertise enable them to create a freer than those that occur in the regular workplace.

Congress has its own model of this. While current Congressional retreats primarily for member and are separated by party, one program seeks to promote good working relations among the entire congressional staff. The Stennis Fellows Program (of the Congressionally-created John C. Stennis Center for Public Service Leadership) is a bicameral, “bipartisan [program] for senior-level staff of the United States Congress...[that] brings together [about 25] chiefs of staff, committee staff directors, legislative directors, and others [each Congress] to explore ways to improve the effectiveness of those who work on Capitol Hill.”25 In place since 1993, the program’s class of fellows for each Congress sets its own research theme and working agenda. For

---

25 http://www.stennis.gov/stennis-fellows
example, the title selected by the fellows in the 113th Congress was “Meeting the Challenge: Bridging Boundaries for the Common Good.”

In the regular course of business on Capitol Hill, Republican staff usually meets with Republican staff, and Democratic staff with Democratic staff. And within those separate groups, leadership staff speaks first and sets the agenda. In the Stennis program, participants are from both parties, and none has preeminence. To reinforce and facilitate the idea that staff should not bring the usual ways of the workplace to the fellows’ sessions and retreats, the Stennis program uses outside discussion leaders to make sure that all thoughts, ideas, and information have an opportunity to be presented and considered. The retreats, conducted “off-site” typically during recess periods, are more conducive to a System 2, slow-down approach rather than an impulsive, “let’s get it over with” System 1 approach. Senior staff—who, in workplace situations, would perhaps ordinarily fight tooth and nail to represent their bosses’ views without considering those of others—when they are in the Stennis Fellows environment are often hard to identify as Republican or Democrat.

A related proposal is simply staying in the same work or social space together until participants figure out how to reach agreement (aka Kissinger strategy). An example of this approach is the negotiation at Andrews Air Force Base in the summer of 1990. While often maligned, the decision to bring representatives from the Administration of President George H.W. Bush together for days with Democratic and Republican members of the House and Senate (and their staff) eventually resulted in a deficit reduction agreement and the Budget Enforcement Act of 1990.

Members of Congress also already practice, in small scale, an aspect of this approach of staying in the same space by going on (also often-maligned) congressional delegations (codels) overseas. For example, when then-Senators Conrad and Gregg (chairman and ranking member, respectively, of the Senate Budget Committee) traveled with a handful of other senators to South America26 before the beginning of the 110th Congress, it was the start of discussions that bore fruit in their introduction nine months later of the Bipartisan Task Force for Responsible Fiscal Action Act 2007. While the provisions in that bill were never enacted, its core was the basic model for the National Commission on Fiscal Responsibility and Reform (Bowles-Simpson Commission) created by presidential order in 2010. Though its list of policy options was not enacted into law, the Commission, whose work spread over nine months, represented a more serious and more bipartisan effort at limiting the deleterious effects of polarization on fiscal decision making than was the Joint Committee on Deficit Reduction (or, Supercommittee) created by the Budget Control Act.

2a. Loss of Diversity of Observation and Dissent and 2b. Reliance on Shared Information

Group decision making often fails to access all the information possessed by all its members. “First speakers” (such as the chairman of a committee) have disproportionate influence. Later speakers self-censor their knowledge of contrary evidence to avoid delaying group progress and to demonstrate a “team player” orientation. Illustrating our caution that no “one-to-one correspondence exists between decision errors and specific features of choice architecture,” one proposal that would offer an improvement to the loss of diversity problem is the one described above on polarization. The Stennis Fellows program creates a deliberative environment that consciously seeks to guard against the disproportionate influence of first speakers. And certainly the private sector has ways of making sure that creative ideas and proposals from junior members of team are not systematically ignored or not taken seriously. Members could avail themselves of off-the-shelf tools to consider information beyond that they are predisposed to value.

For example, in a variation on the Delphi method, Budget Committee members could be polled anonymously at the beginning of each budget season and asked to select, say, the 5- and 10-year targets (for the ratio of debt to GDP and for the reduction in the fiscal gap) that they would support for inclusion in the budget resolution. As part of the exercise, the survey could include questions that require System 2 thinking, i.e., how would you justify your choice of 5- and 10-year targets for debt/GDP or reduction in fiscal gap? Members of “the herds” could remain individuals (via the anonymity of the survey) and express their personal views of what is desirable and feasible. Staff would tabulate and distribute survey results to all committee members. The chairman would then have a closed committee session to discuss the results and reasons given, and then repeat the anonymous survey. The results will likely converge somewhat, and some of the reasons will be novel and persuasive. The chairman might then use the mean/median of those second-round results in his mark. As a consequence, the chairman’s mark might be more realistic and win more support than his own initial starting position would have been.

The advantage is that all members share their unique information, then talk to process it, and then report their revised, more informed, but still anonymous views. Disclose, talk, disclose. There would be less likelihood of cascades; all information would be on the table. This would represent a stark contrast from the approach taken on the 2016 budget resolution by the chairman of the Senate Budget Committee. When the chairman reported his mark out of committee, as amended by the committee markup, there was, for the first time ever, no committee print or report, as required, to translate the committee’s resolution into more salient terms such as committee allocations or resulting debt-to-GDP ratios.
3. Overoptimism

Phaup and Bhatti (forthcoming in Public Budgeting & Finance, 2015) have outlined a detailed proposal for correcting the “planning fallacy” or overoptimism in budgeting. They propose that CBO calculate and publish the difference between the deficits in the budget resolution and the actual annual deficit reported by the Treasury. (They find that for the past 40 years the annual average deficit assumed in the budget resolution has been too low by about $80 billion.) With this proposal, the average expected error could be routinely included in the budget as an outlay allowance that would increase the deficit estimates in the resolution above the levels they otherwise would be without the error correction. As CBO’s estimate of the annual expected error changes, the annual allowance would be adjusted. One advantage of this process is that the correction for persistent deficit estimation errors would provide policymakers with a convenient reminder of the inescapable uncertainty of the fiscal consequences of policy decisions.

Closing Comment

In this paper we have attempted to summarize some key findings of behavioral research and used those results to suggest possible changes in the federal budget process. There is strong evidence that behaviorally inspired changes in choice architecture have been effective in improving decisions in other settings. No one can be certain that this approach will bring similar improvements to federal budgeting. Nonetheless, these seem to be worthy of further discussion and thought and possibly an experimental trial on grounds that they are low-risk, have low costs, would be easy to reverse, and are fully consistent with democratic governance.

Statutory PAYGO

Initially enacted in 1990, statutory PAYGO was in effect for legislation enacted from 1991 to 2002. Early experience with this enforcement mechanism was encouraging. Notwithstanding the technical way the law was implemented, many Members who wanted to propose a bill or amendment to spend more on mandatory programs or cut taxes were under the impression that they had to include other provisions in their measure that offset the deficit increase, for a net deficit effect of zero. Although statutory PAYGO was enforced on a scorecard basis at the end of each year, rather than on a bill-by-bill basis, a temporary, unwritten norm of “pay as you go” did prevail. Nonetheless, the first PAYGO mechanism was abandoned after Congress and the President agreed to “turn off” the sequester that was supposed to have enforced the PAYGO law after the 2001 tax cut was placed on the scorecard (or as much of it as would fit in the first five years, which was more than $500 billion).27

After that law expired, many Members called for its return. After a seven-year hiatus, statutory PAYGO was re-enacted, but in a different form. While the name of the rule remained the same, its operation did not. The strict year-by-year discipline of the first incarnation of the pay-as-you-go scorecard was omitted in the second version. And certain legislation that was expected to be enacted (which would increase the deficit significantly, such as extension of certain expiring tax cuts and the “doc fix”) after the Statutory Pay-As-You-Go Act was enacted in February 2010, was peremptorily exempted from the discipline of that law.

One big difference between the 1990 and the 2010 version is that the first pay-as-you-go scorecard (the tally of violations and offsets) started with a zero balance. The point was that the Omnibus Budget Reconciliation Act of 1990 (which enacted PAYGO by including the BEA) included many provisions to reduce spending and increase taxes, thereby reducing the deficit by an amount approaching about $250 billion over five years. That deficit reduction was not placed on the PAYGO scorecard; rather, the PAYGO scorecard was set to zero so that subsequently enacted laws would not erode the deficit reduction achieved by OBRA 1990. Subsequently enacted legislation had to be paid for, or else OMB would have to execute a sequester. In contrast, the pay-as-you-go scorecard that opened up for business in the new version in 2010 was effectively endowed with a “credit.” One of the first laws enacted after the Statutory Pay-As-You-Go Act of 2010 was the Affordable Care Act (ACA), estimated to produce a large amount of deficit

27 http://www.whitehouse.gov/sites/default/files/omb/legislative/02janseq.pdf
reduction. But $74 billion of the deficit reduction in the ACA was credited to the first five years of the new pay-as-you-go scorecard. Legislation enacted subsequently that increased the deficit has simply drawn down that credit.

So the threat of sequestration that is supposed to enforce the 2010 version of statutory pay-as-you-go has thus far not been a serious threat because the PAYGO scorecard started out with a large credit instead of starting out at zero like the PAYGO scorecard in 1990. If the ACA had been left off the 2010 Statutory PAYGO scorecard (like the OBRA 1990 deficit reduction was left of the first 1990 PAYGO scorecard), the ACA savings would have been preserved for deficit reduction. Subsequent legislation that increased the deficit would have required offsets. Or, such legislation would never have been enacted. Or, if enacted, there would have been a sequester to fulfill the intent of the resuscitated pay-as-you-go law.

Another big difference between PAYGO of the 1990s and the current incarnation is that the first PAYGO scorecard was enforced on a year-by-year basis. If Congress increased spending or reduced taxes next year, it had to eventually, if not immediately, also enact a commensurate offset of the same amount for next year, or else the President would have to order a sequester to produce the required amount of deficit reduction. PAYGO was specifically enforced on a year-by-year basis. Congress was not able to increase spending or cut taxes next year by enacting a spending cut or revenue increase, whether real or phony, that would not take effect until 10 years from now.

In contrast, PAYGO as it currently operates has neither a first-year test nor a year-by-year enforcement. Rather, it “smoothes” the scored year-by-year effects of enacted legislation over the two scorecards on the scorecard (one for five years and one for 10 years). In this way, Congress can easily increase the deficit now as long as legislation is eventually enacted that is scored as decreasing the deficit at some point well into the future, even if that future deficit reduction results simply from a timing shift of revenues from even farther in the future (say, outside the scoring window of 10 years, in years 11 and 12) into the end of the window of the PAYGO scorecard. Given that one of the most problematic decision errors of Congress is to prefer consumption now and pay for it far later, the very first step to correct for this temptation is to (1) reset the current PAYGO scorecard to zero (i.e., erase the credits) and (2) return the operation of PAYGO to its form in the 1990s when it was enforced on a year-by-year basis and spending increases/tax cuts today had to be offset by spending cuts/tax increases today, or else spending cuts via sequestration.

   http://www.whitehouse.gov/sites/default/files/omb/assets/paygo/2013reportcomplete.pdf
The Budget Control Act of 2011 (BCA)

As the BCA is the most recent of significant laws affecting budget process, it perhaps needs less recapitulating than those that came before, but the thumbnail sketch of the components with the longest-lasting effects is as follows:

- The law set statutory caps on discretionary spending (similar to the BEA caps for 1991-2002) for 2012-2021 that were $825 billion lower than baseline discretionary spending (i.e., 2011 enacted appropriations adjusted for inflation).

- The law created a Joint Committee on Deficit Reduction (aka Supercommittee) that was to attempt to vote out a deficit reduction package equivalent to at least $1.2 trillion (including interest savings) over 2012-2021.

- If the Joint Committee failed (and it did), the BCA directed OMB to carry out a “back up” plan over 2013-2021, whereby OMB would have to reduce spending by nearly $1 trillion, or $109 billion per year. OMB is required to achieve this by reducing the initial BCA discretionary caps even further (which would accomplish about 83 percent of the required annual deficit reduction) and by sequestering non-exempt mandatory accounts each year (which would accomplish the balance of the required annual deficit reduction).

- For the first year (2013) of this “back-up” mechanism, Congress mitigated the effects by delaying the implementation by two months. For 2014 and 2015, Congress enacted the Bipartisan Budget Act of 2013 (aka Ryan-Murray) that increased somewhat the discretionary caps from the levels that OMB was required to enforce for those years.

- For 2016 through 2021, the BCA remains in place to be carried out as originally enacted. In addition, the sequestration of mandatory spending is now slated to continue at about $19 billion per year for 2022-2024.

A Practitioner’s Guide to Nudging (Ly et al., 2013) suggests that the “first step in the process of designing an effective nudging strategy is to audit the decision-making process.” The best way to do that here might be through some case studies. Any such “auditor” of the Congressional decision-making process might naturally assume that the only way Congress makes decisions inconsistent with its fiscal intent is by enacting legislation that exceeds its budget plan, which is a frequent event under the current budget process. But often instead, Congress builds undermining decisions (i.e., decisions that produce fiscal results counter to the lower-deficit, lower-debt goal that Members of Congress profess they desire) into the budget plan – “for free.” Consider two examples of decisions made without having to recognize that the decision is counter to the goal of fiscal balance and responsibility.

Medicare Prescription Drug Benefit

Perhaps the poster child of a fiscally irresponsible budget decision made jointly by the President and the Congress is the Medicare Prescription Drug, Improvement, and Modernization Act (MMA), enacted in December 2003. A thumbnail sketch of the development of this law goes as follows:

- On February 3, 2003, President George W. Bush proposed in his 2004 budget to spend $400 billion over 2004-2013 on a new prescription drug benefit for seniors under Medicare (the next Presidential election was in 2004). A month later, CBO estimated that the President’s 2004 request, including his proposal to create this new prescription drug benefit, would double the deficit over the 2004-2013 period from $0.9 trillion to $1.8 trillion (2003). Nonetheless, the possibility of paying for this new debt-financed entitlement program was never seriously discussed by the Congress or the Administration.

- The Republican-controlled House and Senate decided to facilitate enactment of the new drug benefit by including the cost of it in the budget resolution. Because of its privileged nature from the 1974 Budget Act, a budget resolution cannot be filibustered in the Senate and therefore can pass both bodies by a simple majority. And there is no point of order against a budget resolution that includes plans to spend more or tax less in the future than current law would already do. So the conference report on the budget resolution for fiscal year 2004 included a $400 billion reserve fund. The reserve fund could be released by the chairmen of the House and Senate Budget Committees as an increased allocation over the baseline to the relevant committees of jurisdiction (the Finance Committee in the Senate and the Ways and Means Committee in the House) to accommodate legislation
creating this new entitlement, dependent on the condition that the legislation reform (i.e., “modernizes” or “strengthens and enhances”) Medicare.\footnote{SEC. 401. RESERVE FUND FOR MEDICARE MODERNIZATION AND PRESCRIPTION DRUGS.}

- There was little in the eventual MMA legislation produced by the committees of jurisdiction that “modernized” or “strengthened and enhanced” the Medicare program. But the conditions in the reserve fund language in the budget resolution were ignored, and the Budget Committee chairmen released the allocations from the reserve funds to the relevant committees, with the idea being there would be smoother sailing for the prescription drug legislation on the floor of each body of Congress (i.e., less chance of a point of order being raised that would require 60 votes in the Senate to waive and proceed with the bill; indeed, no point of order was raised against the MMA bill when the Senate considered its version, and the bill passed 76-21).

- Despite these well-laid plans, immediately after the Senate invoked cloture (by a vote of 70-29) against a filibuster of the conference report on the MMA, the Senate Minority Leader, Senator Daschle, raised two points of order under the Congressional Budget Act against the bill, but not because it spent more than the $400 allowed by the budget over 2004-2013 (it didn’t; in fact, the net cost estimate of the bill came in at $395 billion – under the amount allowed). Instead, Senator Daschle raised the points of order because the bill spent $4 billion more in 2004

\footnote{http://thomas.loc.gov/cgi-bin/cpquery/z?d=c108h&s chamber=h&j=hr071.108&b=1&i=3&d=cp108FXZy6&m=hr071.108&j=hr071.108&hd_count=50&i=3&n=3&sel=TOC_76877&}

32
alone than assumed in the budget resolution (and the same $4 billion more than was allocated to the Finance Committee for 2004). Senator Daschle attempted to use the points of order arising from spending an extra $4 billion more in 2004 to kill the conference report – not because it increased the deficit by too much, but rather because it did not increase spending and the deficit over 2004-2013 by an amount even larger than $400 billion, as he argued:

I don’t challenge the $400 billion. Frankly, I don’t think that it is adequate to provide a meaningful drug benefit. We have to do better than that.\(^{33}\)

In fact, Senator Daschle and others wanted to spend well more than $400 billion on the bill to “close the donut hole.” But, with one vote to spare (by a vote of 61-39), the Senate waived Senator Daschle’s point of order, the conference report survived, and the Senate proceeded to pass it by a final vote of 54-44.\(^{34}\)

- President Bush signed the MMA into law, which has been in operation since January 2006. There has never been an effort to pay for the costs of the program or otherwise prevent the MMA from being a pure add to future deficits and debt. Proponents of the bill instead have attempted to burnish their fiscal-discipline bona fides by observing they have been “correct” all along, as the actual cost of the bill has not turned out to be quite as much as initially estimated in 2003.\(^{35}\)

**Free Highway Spending**

Another example of Congressional fiscal irresponsibility is the use of the Congressional budget resolution to circumvent the fiscal limit that is supposed to be provided by the use of a trust fund to link spending on highways to amounts collected in taxes by users of federal transportation infrastructure. The federal Highway Trust Fund is credited with receipts (primarily gasoline taxes), which is supposed to ensure that such taxes are only


\(^{34}\) For completeness, it is important to note that even though no Budget Act points of order came into play during the House’s consideration of the legislation, passing the bill in the Republican-controlled House was not a cakewalk, even though it was written by Republicans for a Republican president to sign. When the bill was first considered in the House in June 2003, three Republicans had to change their initial “nay” vote before it could be passed 216-215. When the conference report came back to the House on November 22, 2003, in an unusual move, the vote was kept open from 3 AM until about 6 AM, until it passed 220-215.

\(^{35}\) “Over the 2006–2013 period covered by CBO’s original cost estimate, net federal spending for Part D was projected to be $550 billion; actual spending was $353 billion, or 36 percent less.” See page 5 of CBO, *Competition and the Cost of Medicare’s Prescription Drug Program*, July 2014 ([http://www.cbo.gov/sites/default/files/45552-PartD.pdf](http://www.cbo.gov/sites/default/files/45552-PartD.pdf)).
used to pay for highways and, symmetrically, that highway spending is limited by the amount of those receipts.

But since 1998, the Highway Trust Fund has been running out of money (CBO, 2014). This result is due to annual Congressional budget resolutions that continue to allocate more spending authority to the highway committees than can be paid for by the incoming revenue (primarily gas taxes) earmarked for the fund. The budget resolution provides the increased spending authority to the committee of jurisdiction “for free” – it neither plans for nor instructs any other committee to create additional revenue from sources related to highway usage (or any other source for that matter). And, as with the Medicare Modernization Act, the budget resolution that provides free money can be adopted by a simple majority.

Given this increased spending allocation, when the highway bill comes to the floor of each body with the level of spending assumed in the budget resolution, that highway bill does not face a point of order. Once that bill is enacted, the Department of Transportation then obligates the amounts as directed, and when it comes time to write the checks to states to liquidate those commitments that exceed the available Trust Fund balances (i.e., gasoline tax revenues), Congress has thus far decided it has little choice but to retroactively make good on those outlays in a timely manner by authorizing Treasury borrowing to pay for it, thereby increasing the federal debt. The bottom line is that Congress assumes excess highway spending authority in budget resolutions even though it knows the Trust Fund does not have the money to pay for it and that Congress ultimately will have to enact a subsequent law that increases Treasury borrowing to cover it, which means that future taxpayers (i.e., children and grandchildren) will have to pay for it instead.

The CLASS Act of 2010

Our third case study illustrates how the cash-basis presentation of some of the largest federal social insurance programs, such as Social Security and Medicare, encourages a “present bias” in budget decisions because it delays the recognition of current-period costs until they are beyond the responsibility of current budget decisionmakers. In those cases, cash-basis budgetary accounting violates the “concept (or more correctly, the functional necessity) of matching costs in the budget with the act of allocating scarce resources to alternative purposes. As a rule, cash-basis accounting for cost understates (1) the current-period claims on resources that are being made by payments that are deferred until later and (2) the resulting contribution of those claims to the deficit and the debt. In an a odd twist, because of the mistreatment of interest, this method of accounting also overstates the net long-term outlays of an actuarially sound program, where premiums are set in present values equal to the future value of claims.
All of these shortcomings of the currently-used cash accounting are spelled out clearly by CBO in its 2009 cost estimate\(^{36}\) (and explanatory letters to Senator Harkin\(^{37}\) and Congressman Miller\(^{38}\)) for the Community Living Assistance Services and Supports Act (CLASS Act), which was part (Title VIII) of the Affordable Care Act of 2010. This legislation attempted to create a voluntary long-term care insurance program for those who might enroll in case they are later unable to perform basic living functions for themselves, such as dressing, cooking and eating. Care was to be provided in the community, either in the home of the disabled person or in a residential facility. The insurance would have been provided through employers after a five-year vesting period. The program was explicitly intended to be self-supporting from premium income. Accordingly, the Secretary of the Department of Health and Human Services was directed to either (1) set premiums and adjust benefits as needed to assure the financial viability of the program, or else (2) never initiate the program if it was decided that the program would never be sustainable and self-financing.

The CBO cash-basis cost estimate faithfully reflected the provisions of the legislation:

- Premium income (offsetting receipts, i.e., negative outlays) was projected to flow into the government for the first five years. With no possibility of claims during this initial vesting period, the insurance program would have accumulated a balance of more than $30 billion, further increasing to more than $73 billion over years 6-10, even after allowance for expected claims and administrative costs.

- For the 2010-2019 period, the CLASS Act was projected to “reduce” the budget deficit by $72 billion (the excess of premiums over claims).

- Premiums were also projected to exceed claim payments and other costs in the next decade (2020-2029) and to similarly reduce the deficit by the difference, though by less than in the first decade.

- Only in the third decade did CBO estimate that insurance outlays for claims would exceed premium income and the program would begin to show a cost in net outlays and the deficit.

- Because the CLASS insurance fund balances were to be invested in Treasury securities, the CLASS Act would also reduce federal borrowing from the public by the same amount and result in savings in interest payments to the public that


\(^{38}\) [http://www.cbo.gov/publication/41853](http://www.cbo.gov/publication/41853)
would now be paid to the trust fund as interest on fund balances. Those payments would have been required by the trust fund to permit premium balances to grow into the future value of insurance claims.

- However, interest payments to the insurance fund by Treasury are intragovernmental transfers that have no effect on net federal outlays. Thus, over an entire cohort of insured employees, the actuarial payment of claims by the trust fund was scored as exceeding the collection of premiums by the amount of interest earned by the trust fund on its balances.

As a result, the CBO cost estimate of the CLASS Act under current cash-basis accounting indicated to members and the public:

- The CLASS Act was projected to provide $72 billion in extra income or budget “resources” over the first 10 years, and additional unspecified amounts in the next 10 years that were available to either offset new spending provided elsewhere in the ACA or else to allow the authors to claim deficit reduction.\(^{39}\)

- The interest savings from collecting premiums in advance of claims and reducing public borrowing and interest is not a legitimate offset to other federal spending.

- This cost estimate gave credit for and counted a false saving (actuarially fair premiums already obligated to pay future claims), while ignoring a genuine one – interest saved on debt not sold to the public because of the accumulation of premiums.

- Enacting this legislation would impose a cost on government in the future, but for the next 20 years, the CLASS Act would be a fiscal positive.

As there was no feasible accounting solution available under law to this misleading accounting, one could be sympathetic to CBO’s tied hands. As many expected, however, in October 2011, the Secretary of HHS determined that no feasible premium schedule could be established that would permit the insurance to be offered as a self-sustaining program because of the threat of adverse selection. That is, only the least healthy, most-likely highest claim employees would enroll, thus driving premium rates higher until only the worst risks were in the insurance pool. Consequently, as directed by the ACA, the Secretary halted efforts to establish the insurance program. (For good measure, in January 2013, Congress repealed the CLASS Act.)

\(^{39}\) Senator Conrad, Chairman of the Senate Budget Committee in 2010, prevented this sum from being included as a credit on the just-enacted statutory PAYGO scorecard. Hence, the funds could not be used to avoid a sequester. However, nothing offset the general perception and advertising that the ACA (including the “savings” from the CLASS Act) represented a “hard decision” to reduce the deficit.
Even though the net premium income for the first 20 years would eventually be needed to make good on insurance claims in the future, the cost estimate clearly indicated that the CLASS Act premium income contributed $72 billion to deficit reduction in the first 10 years of the ACA.\footnote{A slightly more complicated, but more relevant, case is the Old Age and Survivors Insurance (OASI) program of Social Security. Currently earmarked payroll taxes are about 1.7 percent of taxable payrolls short of the amount required to pay accruing scheduled benefits. To get a more relevant and salient measure of the cost of Social Security might require moving the OASI trust fund outside the federal corpus and requiring an annual government payment to the fund equal to the tax shortfall. In addition, payroll taxes might be scored as both revenues and outlays when received and paid to the trust fund. This treatment would make clear the amount of the periodic cost of Social Security that is within the reach of current budget decisions, the amount that is not, except by default on past promises, and the total public debt including existing obligations to make payments under current law.}

However, an easy solution is readily available: simply define the insurance fund to be outside the definition of the unified budget, just as the financing accounts for direct loans and loan guarantees are treated under credit reform and accrued interest on the public debt by long-standing rule.

If insurance programs were moved outside the federal corpus (or “off-budget”), premiums would not be scored as federal receipts, and equal value claims payments would not be scored as outlays. However, interest paid on Treasury debt would be scored as outlays, and debt held by the insurance program would be included in debt held by the public, just as such debt held by the Federal Reserve is treated.

As a result, collecting premiums in advance of claims payment for insurance funds would not be scored as deficit reduction, but interest payments on debt held by those funds would be scored as outlays as earned.

**What Does the Audit of the Decision-Making Process Tell Us?**

It is clear from just these three examples that there are no impediments that currently restrain a budget resolution or prevent misstatement of the budget effects of payments to an insurance fund. After Congress approves the deficit increases in a budget resolution (or spends insurance premiums on something other than the insured purpose), the follow-on legislation to implement those assumptions usually has an easier path to enactment.

The 40-year-old Congressional budget process, despite providing lawmakers with much of the information they need to realize their goals for budget policy, appears to fail to provide sufficient cues and nudges in their day-to-day decisions consistent with their own professed goals and preferences for the country. For one, it fails to make the future costs and risks of today’s decisions salient today.
REFERENCES


Ly, Kim, Nina Mazar, Min Zhao, and Dilip Soman (2013). A Practitioner’s Guide to Nudging, Rotman School of Management, University of Toronto.


