Long-term budgeting: A cautionary tale from U.S. experience

by

Joseph White Ph.D.*

There is a growing tendency, among central budget-makers and commentators, to argue budgets should be made for the long-term, rather than the traditional annual budget. This tendency is especially strong in the United States, where it has become virtually a conventional wisdom. This article explains, first, why that approach fits very poorly with most of the goals of budgeting. It then evaluates U.S. experience with approximations of long-term budgeting of three types: i) medium-term limits on discretionary spending, ii) the Social Security programme, and the iii) Medicare programme. That experience illustrates the reasons why long-term budgeting would not be a positive reform. They include the fantastical nature of many long-term forecasts, strong incentives for both deception and self-deception about the effects of planned budget totals, and ignoring the basic task of budgeting, which is to reconcile preferences about policy details to preferences about budget totals in a way that considers each.

JEL Codes: H11, H41, H50, H68

Keywords: Accountability, Efficiency, Forecasting, Health, Long-term, Pensions, Representation, United States

* Joe White is Luxenberg Family Professor of Public Policy at Case Western Reserve University. He earned his A.B. from the University of Chicago and his Ph.D. in Political Science from the University of California, Berkeley. Before joining the Case Western faculty in 2000, he was an Associate Professor at Tulane University and first Research Associate and then Senior Fellow with the Governmental Studies programme of the Brookings Institution. His research focuses on both budgeting and health care policies and politics, with a focus on the United States but from a comparative perspective. In 2014 he received the Aaron Wildavsky award for lifetime achievement in budget scholarship from the Association for Budgeting and Financial Management. He has benefited over the years from the wisdom of many other scholars, none of whom bear any responsibility for any controversial assertions herein. But Dr. White especially appreciates the encouragement and support from Paul Posner for the work that became this article, and dedicates it to Dr. Posner’s memory.
1. Introduction

What is the appropriate time frame for a government budget? Budget processes normally have stated plans for the collection and allocation of resources over a period of one or occasionally two years (Caiden, 1982; Wildavsky, 1978). Recently, however, many budget-makers have promoted much longer-term perspectives.1

A long-term perspective could mean considering projections of possible long-term trends as part of the information used to make annual budgets. The OECD network of Senior Budget Officials endorsed a version of this view in OECD’s 2015 Recommendation of the Council on Budgetary Governance (OECD, 2015a; White, 2015a). The ninth of ten principles is to “Identify, assess and manage prudently longer-term sustainability and other fiscal risks.” This includes “publishing a report on long-term sustainability of the public finances, regularly enough to make an effective contribution to public and political discussion on this subject, with the presentation and consideration of its policy messages – both near-term and longer-term – in the budgetary context” (OECD, 2015a: 10).

In the United States, long-term projections have been common since they were created by GAO and taken up by CBO in the 1990s (GAO, 1992; CBO, 1997a). They were and are used to heighten policy-makers’ concerns about deficit control. When budget conditions looked particularly good, long-term forecasts were used to call for restraint based on possible deficits decades in the future (GAO, 2001).

Many of the leading figures in debate over the U.S. federal budget process, however, recommend going further. They maintain that instead of being judged by the prospective deficit in the next year, the federal budget should be judged by the deficits or debt it is estimated to produce decades into the future. In short, allocation decisions should be driven by long-term goals for totals, and policy details designed and judged based on their estimated effects on budget goals 20, 30, or more years into the future.

Thus GAO has proclaimed that, “Long-Term Focus is Critical” (GAO, 2004) and that any delay in addressing those concerns would be “destabilising” and unwise (GAO, 2011: 2). It has called for “enforcement mechanisms” to encourage better deficit control for longer periods of time (GAO, 2011: 3). In 2008 a group of budget commentators including three former Directors of the Congressional Budget Office proposed thirty-year budget caps for Social Security, Medicare, and Medicaid, enforced by “automatic adjustments in benefits, premiums, provider payments, or other revenues” in order to save the nation’s “fiscal future” (Antos et al., 2008: 2). Eleven members of the National Commission on Fiscal Responsibility and Reform (“Bowles-Simpson Commission”) in 2010 endorsed a plan that defined the goal of budgeting as to, among other things, reduce federal debt as a share of GDP in 2035 to 40% (National Commission, 2010). They clearly believed that budgets should be judged by whether they would (purportedly) hit that target.

In March of 2011, ten former chairs of the President's Council of Economic Advisers declared that, “repeated battles over the 2011 budget are taking attention from a more dire problem – the long-run budget deficit.” They called this deficit – which was a projection
decades into the future – “a severe threat that calls for serious and prompt attention.” They asserted that, even though “the actual deficit is likely to shrink over the next few years,” immediate action was needed because, “at some point, bond markets are likely to turn on the United States – leading to a crisis that could dwarf 2008” (10 Ex-Chairs, 2011). That was a remarkable claim considering that 2008 was the worst financial crisis in nearly eight decades. The battles over the 2011 budget to which the former CEA chairs referred included the possibility that the federal government would be unable to raise its debt ceiling, and therefore have to default on federal securities in August – an event which might be expected to have a more immediate and dire effect on the bond market. Nevertheless these economists were joined by the Committee for a Responsible Federal Budget in arguing that the long-term prospects were a similar crisis. In a June, 2011 statement, the CRFB declared that “we need to see” both “a debt ceiling increase as soon as possible” and “a companion debt deal that (a) Calls for at least USD 4 to USD 5 trillion in total deficit reduction” including “steps toward reforming unsustainable entitlement programmes.” “Failing to raise the debt ceiling,” CRFB declared, “would likely lead to a fiscal and financial crisis, but failing to address the debt would eventually do the same.” The statement acknowledged that entitlement “reforms” could not be drafted immediately, but called for a “credible process” including some set of binding automatic consequences if those reforms were not adopted (Committee for a Responsible Federal Budget, 2011a: 1, 3, 4).

Former CBO and OMB Director Alice Rivlin and former Senate Budget Committee Chair Pete Domenici (2015: 13) urged reforming the federal budget process so that Congress would “enact explicit long-term budgets for Medicare, Medicaid, and Social Security as well as other mandatory programmes” with “limits on automatic spending growth” that would be enforced by “re-establishing and simplifying pay-as-you-go rules for these mandatory spending programmes.” The National Budgeting Roundtable has now generated two analyses calling for enforceable long-term totals for either all or most of the federal budget. Phil Joyce (2016: 4) calls for establishing a statutory fiscal rule to determine long-term totals. Stuart Butler (2016: 3) called for “a 25-year budget for long-term mandatory programmes, together with a funding plan.” Each proposal says that Congress could make changes to the long-term plan. Each also, however, provides for automatic procedures, to make programme or revenue details fit the targets, as a default. In Butler’s (2016: 4) testimony, “if spending or revenues for these programmes exceeded or fell short of the corridor established in the original statute, automatic provisions would be triggered to maintain the original long-term budget.” In Joyce’s version (2016: 5), deviations from the agreed budget path would be countered with sequestration applied to “all spending and revenue changes, including tax expenditures.”

Advocates for such proposals may argue they are not really long-term “budgets” because they do not intend for the “enforcement” provisions to go into effect. In this view, setting caps is not really budgeting; a budget must include decisions about budget details. The principle is correct: deciding about totals without considering what that means for details is an irresponsible and often dishonest way to budget. Yet enforcement provisions are details. They are the policy unless different laws are passed instead – just like any other budget law. As a default, the enforcement has all the power of the status quo in a political system in which legislating can be exceedingly difficult. Moreover, if Congress and the president could agree to waive the details of the enforcement, they also could agree to waive the totals – that is, to have higher deficits than planned. Therefore, proposals which combine long-term decisions about totals with enforcement procedures either should be
viewed as long-term allocation decisions to reconcile details and totals – that is, budgets – or should not be viewed as providing serious caps. Advocates who propose “enforcement” that they do not mean to occur are misleading the public, themselves, or both.

In these cases and many others, budget experts in the United States appear to believe that decisions should be based on forecasts of results far in the future. A long-term baseline would be estimated and then policies chosen because they are projected to reach some more desirable target. In this view, policies that do not move in a major way from the projected baseline to the target are irresponsible. So responsible budgeting must mean budgeting to meet long-term targets.

In international practice, long-term analyses have been encouraged but, “not used for allocation” (Schick, 2009: 17). Nevertheless, versions of the American budget community's belief in long-term budgeting, as opposed to long-term projections, appear to be attracting interest within the international budgeting community. In these pages Marc Robinson recently wrote that the “long-term fiscal projections and analysis which are now produced by most OECD governments (and by the European Union)... need to be more closely linked to budgeting” (Robinson, 2016: 45). Allen Schick has suggested that the future of budgeting might include both transforming Medium Term Frameworks into more binding budgets, and adapting the current “analytical tools” of long-term analysis into fiscal rules that “require or prohibit certain actions” (Schick, 2016: 28-31).

In this paper I will argue that American experience provides ample evidence about why the American emphasis on long-term budgeting should be rejected in other countries and scaled back in the United States. In the name of alleged “fiscal responsibility” it threatens many of the key purposes of budget processes.

The problems begin with the fact that long-term forecasts call for decisions based on information that is distinctly less credible than the information used for traditional budgeting – flawed as the latter can be. This is especially problematic for health care, a large part of any modern state's expenditure. Moreover, many of the standard purposes of budgeting – including support for representative government, accountability, transparency, macroeconomic management, and efficiency – are threatened by how long-term tactics work. Long-term approaches also favour particular policy views over others, rather than providing a neutral framework for informed and representative choice. Even the view that it would favour – one that prioritises “economy” in the form of lower spending over all other values – might be favoured mainly in a fraudulent way, with fake decisions.

The first section of this paper looks at these goals and purposes of budgeting, and considers how long-term budgeting might be expected to fit with those purposes. The next three sections evaluate U.S. national government experience with approximations to long-term budgeting. These include five- and ten-year ostensibly binding totals for discretionary spending (spending by bureaus that receive annual allocations for some purpose, as opposed to more open-ended mandatory or entitlement programmes); decision-making for the Old Age, Survivors and Disability Insurance (OASDI) pension programmes, known in the U.S. as “Social Security”; and budgeting for the Medicare programme that provides health insurance for the elderly and disabled. The final section of this article will summarise conclusions about the limited ways in which long-term budgeting might be useful, and the reasons why it is not a direction that budget reformers should follow.
2. Annual Budgeting, Long-Term Budgeting, and the Goals of Budgeting

“Budgeting is supposed to contribute to continuity (for planning), to change (for policy evaluation), to flexibility (for the economy), and to provide rigidity (for limiting spending)… Obviously, no process can simultaneously provide continuity and change, rigidity and flexibility. And no one should be surprised that those who concentrate on one purpose or the other should find budgeting unsatisfactory.” (Wildavsky, 1978: 501)

2.1. Goals for budget processes

As Aaron Wildavsky reminded us, the goals of any budget process are multiple and potentially conflicting. These themes go beyond what Robinson (2016: 30) calls “the main objectives of budgeting.” He emphasises “aggregate fiscal discipline”, “allocative efficiency”, and “effectiveness and efficiency of expenditure” or, in the more recent jargon, “value for money”. The OECD Recommendations and other literature (Axelrod, 1988; Lewis and Hildreth, 2013; Rubin, 1997) provide a much broader view of budgetary values that should be familiar to most readers.

2.1.1. Budgeting and representation

One well-known goal is to enhance accountability: of the government to the citizens and the executive to the legislature. Publishing plans and promises and then auditing performance provide ways for the citizens and their representatives to direct and/or oversee the government. Accountability thus is served by the norm of transparency: decisions are made in some visible way, with accurate reporting of how revenues are gathered and money spent. Both are major themes of the OECD Recommendations.6

Budget processes also are intended to help the legislature or political executive control the executive agencies, including to ensure that they spend in the way intended and no more than the amount intended. In Wildavsky’s (1978: 501) words, “Control over public money and accountability to public authority were among the earliest purposes of budgeting.”

Representation in the form of responsiveness to political leadership is also the principle behind OECD’s (2015a: 6) recommendation that processes “closely align budgets with the medium-term strategic priorities of government.” It addresses a concern that budgets tend to be frozen into set patterns, so not responsive enough to election of new governments. This function can be overstated as a presumption that each year’s budget should “set priorities.” In practice there are good reasons why budgeting is usually incremental, with modest changes in the priority of most policies (Schick, 2009; White, 1994). Shift points (Dempster and Wildavsky, 1979) are punctuations in normally stable equilibria (Baumgartner and Jones, 1993). Nevertheless, a budget process that could not accommodate changed preferences because of events (such as 9/11 or the collapse of the Soviet Union) or elections that change the governing coalition would violate the basic goals of representative government.7

The budgetary literature tends not to emphasise as much another aspect of representation, but that may be because it is taken for granted. A democratically accountable budget process is one that will not systematically distort budgeting to favour any specific group or interest in society at the expense of others. Budget processes should be procedurally fair – or, at least as fair as the rest of the political system. As part of this fairness, budget theorists have argued that budgetary staff must provide neutral competence in budgeting.
2.1.2. Economy and efficiency

Although frequently conflated, economy and efficiency are not the same (Simon, 1997; Wildavsky, 1966). Economy means, simply, limiting spending. It is the dominant goal for many participants in budget processes, especially participants who represent wealthier voters who are likely to pay more of the taxes that finance spending. Budget process reforms frequently are adopted as part of an economy campaign. But efficiency is a more fundamental goal, in part because economy can be, from a value-for-money perspective, wrong-headed. Economy could be served by eliminating highly useful programmes, or cutting programme spending in a way that reduces their output even more, thereby giving citizens a worse deal than they received before. A focus on efficiency requires no ideology, and making efficient decisions is, in principle, the most neutral of competences.8

Efficiency itself has multiple meanings. The most significant distinction for budgeting is between seeking efficiency in the operations of individual programmes (as in traditional budgeting, which Wildavsky called “pure efficiency”) or to make the most “efficient” choice about the package of programmes financed (in Robinson’s terms, “allocative efficiency”). “Performance Budgeting” in its many forms promises to create efficient choices among programmes. In practice, pure efficiency at the programme level is much, much easier to pursue than any version of efficiency that includes the choices among programmes (White, 2012a).9 The pursuit of efficiency at the programme level, however, is one of the core tasks of traditional budgeting – checking on agency proposals by “scrubbing the estimates.” If efficiency can be increased by improving agency management, enabling that improvement is also a goal of budget processes. Making management more difficult would be a failure.10

2.1.3. Budget totals, the macroeconomy, and “balance.”

The most visible set of goals involves totals for revenue, spending, and their difference (the “balance”). The relative priorities of spending, revenue, or balance values varies among policy-makers. In general, limiting deficits is a more consensual norm than limiting either spending or revenues. This need not mean a “balanced budget.” Instead, balance as a goal means choosing a difference between taxes and spending that is viewed as reasonable within that political unit (Rubin, 1997), such as the Maastricht standards for the European Union. The OECD (2015a: 6) calls for “clear, credible and predictable limits for fiscal policy,” rather than the narrower “balanced budget.”

Some participants in politics view balance from a distrust of debt by analogy to household debt (White, 1998a) or of the freedom given politicians if they are able to borrow (Savage, 1988). Most policy-analytic arguments about the proper budget totals, however, at least claim to be based on beliefs about economic or fiscal policy. Different perspectives then have very different implications for budget totals.

The demand management approach, derived from Keynesian economics, says that the right balance is the level of deficit or surplus that will have the proper effect on aggregate demand – stimulus to increase employment, or contraction to slow inflation. It therefore focuses on the balance and assumes that should change with economic conditions. An alternative approach, particularly prominent among American economists, asserts that budgets should be used to increase national savings, so de-emphasises response to macroeconomic fluctuations (White and Wildavsky, 1991) and promotes budget surpluses (White, 2003). The savings argument, unlike demand management, sets no standard for totals in any given year: more savings is always better.
A third view at least purports to establish financial market confidence as the core standard for the use of the budget to shape the economy. It claims the markets will punish the nation for whatever budget totals the advocates for this approach do not like. The argument could in principle be used for any policy: the deficit-increasing mix of Reagan administration initiatives in 1981, the deficit-reducing initiatives of the Clinton administration in 1993, or the Trump administration tax cuts. Nevertheless, the usual claim is that deficits must be lowered because, otherwise, high real interest rates will torpedo the economy (White and Wildavsky, 1991; Woodward, 1994). Since participants in financial markets are assumed to be forward-looking, expected future deficits could be as damaging as current ones. The argument persists in spite of substantial contrary evidence. Last but in the U.S. at present not least come arguments that the government is too big, and that size itself – both taxes and spending – harms the economy. A telling example was the George W. Bush administration’s argument that taxes should not be raised to pay for new “homeland security” spending “because of the economic distortions introduced by the tax system” (Office of Homeland Security, 2002: 65). This view is the core fiscal policy of the current Republican party, in which claims about the need to cut entitlements due to their long-term budgetary effects are common but a massive tax cut which increased projected deficits by USD 1.8 trillion over ten years is just fine (CBO, 2018; Stein, 2017). The most obvious budget process reform based on this view is proposals to set fixed limits on spending or taxes as a share of the economy.

Readers must have their own opinions of the merits of these views. Each is based on ideologies or distributional preferences. Therefore standards for evaluating budget totals are one of the most important political inputs to budgeting, and designing the budget process to favour a particular economic approach must be controversial at best and economically damaging at worst.

2.1.4. Honesty and conflict management

Two other standards are crucial for evaluating any budget process, even though they may be taken for granted and not stated as explicit goals. The first is that budgets should be honest and accurate. Budgets built from fake estimates or on deceiving the public about what is included duck accountability, are not transparent, ignore or fake efficiency, do not reveal the actual economic policy, and misrepresent priorities. The OECD therefore made providing for the “integrity and quality” of budget decisions its tenth recommendation, mentioning measures such as independent audit and forecasting institutions. Annual budget processes are repeated games in which the victims of lies can respond the next year, and that might encourage honesty. Nevertheless, there can still be rewards for deceptions that are clever enough to not get caught, sometimes the deceiver has allies who can protect it, and sometimes honesty is in hardly anyone’s immediate interest.

As the last caveat suggests, in practice, the goal of honesty may be violated because of another goal: budgeting must enable decisions without overwhelming the system with conflict. Making a budget might sound like an optimisation problem on many variables, but that would ignore the fact that different people would optimise, or satisfice, in different ways. The potential for conflict is maximised by peculiarities of budgeting as a social choice problem. The core conflict within budgeting involves mismatches between preferences about details and preferences about totals. Preferences about details for spending and revenue tend strongly to sum to deficits that are larger than preferences for deficit totals.

The traditional political theory of budgeting (Wildavsky, 1964) frames this conflict as between budgetary “claimants” and “guardians.” In this view selfish “special interests” in
programmes or avoiding taxes oppose and sometimes block pursuit of the “general interest” in lower deficits. Budget-makers see themselves as the guardians of the general or public interest in more controllable, sustainable, responsible (choose your adjective) totals.

A less common theory, however, may be more accurate, especially under current conditions. The guardian/claimant distinction, and the guardian role, may apply better in situations where the totals are generally acceptable, so the challenge only is to prevent deviation, then when totals are already unpopular. In that situation, an inconsistent social choice problem, identified most clearly by Anthony Downs (1960), appears more fundamental.

In Downs’ words, the package of government programmes always includes “at least one act which any given voter opposes” – in fact it surely includes many more. Because different people or groups want to cut different programmes, however, significant packages of spending cuts or tax increases that would be accepted by majorities are much more difficult to identify. Hence consistent individual preferences (each person could balance the budget) sum to inconsistent social preferences (no majority for any specific way to do so).15

From the guardians vs. claimants perspective, totals can be reconciled with details by simply declaring the totals and forcing the details to fit. Advocates for details are seen as selfish special interests so should be subordinated to the disinterested guardianship about totals. The OECD, accordingly, strongly endorses “top-down budgetary management” (2015a: 6). But this does nothing to ensure that there is any set of details that can fit the total and win majority support. Opponents of any measure do not have to feel like hypocrites because they do support alternatives that would fit the totals. So the total for the top does not produce moral suasion for compromise. A set of totals that looks reasonable without considering the details might be viewed as damaging by majorities once the details are filled in.16 So in normal budgeting consideration of all the consequences leads to adjusting preferences about details and totals to each other in an iterative process – rather than simply setting a total.17

2.2. Annual vs. long-term budgeting, in principle

Although the traditional budget process or schedule in both the United States and other countries has been annual (Wildavsky, 1978), that has never meant that all decisions were made for only the following year (Caiden, 1982). Just as families will make long-term commitments such as mortgages, buying cars, and committing to pay a child’s college tuition, governments engage in some activities for which they make commitments for multiple years: such as buying an aircraft carrier (they take a while to build), or committing to pay pensions (such as government employee pensions, or Social Security), or promising to pay back loans (much like a home mortgage). Yet almost all governments have had annual budget processes in the sense that each year they reviewed and tried to project what the spending and revenue totals would be if they kept doing what they did before, and considered whether to make changes in budget details in order to achieve different totals.

2.2.1. Annual budgets and the purposes of budgeting

Annual budgets fit with representative government for a series of reasons. One is that they serve the goal of controlling agencies, by setting and enforcing targets for their spending. Annual decisions also involve a time frame that voters and legislators can understand and assess, comparable to the standards used to evaluate performance in any other organisation (imagine if investors were expected to rely on five-year rather than annual reports!).
In annual budget processes, Central Budget Agencies and the legislature can review agency plans to search for ex ante efficiencies. Annual decision-making fits best with this pursuit of efficiency, because many of the factors in agency production of goods and services, such as prices of fuel or wages, change frequently. Demand factors for some programmes change with events in the economy. Plans for longer than a year can easily be superseded by events.

The theory of economic management which emphasises aggregate demand also fits best with annual budget decisions. Changes in the extent to which government spending and taxes add or subtract money from the private economy will influence short-term inflation, unemployment, and interest rates. Annual budget decisions are not as effective a tool for fiscal policy as was assumed when Keynesian demand management was first implemented by governments. Legislative responses to economic change are slow and often poorly timed (e.g. the spending occurs when the slump has ended). Therefore the best response to economic cycles is “automatic stabilisers”: programmes such as the graduated income tax and unemployment benefits that are countercyclical. Nevertheless, the annual budget serves fiscal policy goals by forcing a review of trends and a statement of the government’s approach. Legislated fiscal changes also are highly justified in extreme circumstances such as the shock created by the global financial crisis.

As discussed above, the repeated interactions within annual budget processes on balance also can encourage relative honesty. For these and other reasons, the traditional annual budget process has persisted in spite of many criticisms. We turn now to those criticisms, and to whether long-term budgeting is likely to be an improvement.

2.2.2. The flawed claims for long-term budgeting

Advocates for long-term approaches may claim that they make government more transparent by highlighting consequences that would otherwise be ignored. Information can only improve transparency, however, if it is true. The United States represents the international state of the art for long-term forecasts, both in terms of the sophistication of methods and having multiple, competing sources. Nevertheless these forecasts are highly unreliable guides to budget totals far in the future. As Rudolf Penner has written, if the purpose is simply to call attention to long-term issues, “the huge inaccuracies do little harm.” But one should be very cautious about assuming any budget forecast is good information because, as he wrote, “it has been shown that forecasts become rapidly less reliable as the forecast period is extended” (Penner, 2001: 14, 12).

A further argument asserts that long-term budgeting improves representation by bringing in the concerns of future generations. In Eugene Steuerle’s version of the argument, both entitlements to pensions and health care and Republican tax cuts, by shaping future spending and revenues, mean that “dead men” rule. “Yesterday’s policymakers have robbed their successors of their fiscal freedom” (Steuerle, 2014: 11). Any programme, such as pensions, that makes long-term promises is suspect because it is “set in law and, in the real world of policymaking, changing the law to break past promises to voters is easier said than done” (Steuerle, 2014: 9). He therefore calls for immediate long-term measures to reduce entitlements for the elderly.

Yet the power of the status quo applies to any status quo: for example, privatizing Social Security would make it very hard to restore the current system. Any automatic Medicare spending cuts would have the same advantage whether or not they look like good policy
when the time comes. If enforcement procedures are protected by supermajority provisions, as is commonly proposed, they would have a further advantage. Either way, decisions made in 2018 will influence results in 2048, so why is trying to force cuts now more democratic than trying to enforce promises about benefits? The long-term approach is particularly suspect when its advocates, as in the Bowles-Simpson plan, set rules about the “size of government.” If one wants to avoid foreclosing choices for future voters, legislating automatic mechanisms that would be backed by the power of the status quo and target pain on particular future voters does not seem an obvious approach.

In practice, some public policies, especially pensions, make very little sense if they cannot include long-term and fairly binding commitments. In the United States, the big federal entitlement programmes also have been extremely popular for a long time. The political pattern that campaigners to reduce entitlements want to “fix” reveals the point: proposals to cut benefits (as opposed to payments to medical providers) very rarely pass. Over fifty years of experience with Medicare and eighty with Social Security show that support for those programmes has persisted across “generations.” The whole framing in terms of generations should be challenged, because aging is part of the life cycle. Any day’s taxpayers are the future’s beneficiaries, and making their health and income when they become elderly less secure may not seem like helping them.

Budget analysts and policy-makers can reasonably disagree about what form of “binding the future” is better policy (Rubin, 2009; White, 2009a). Yet claims that capping entitlements would make budgeting more representative in principle are highly questionable. They may be attractive from some public finance perspectives – Steuerle’s “fiscal freedom” may look similar to the IMF’s and OECD’s advocacy for “fiscal space” (Heller, 2005; Marcel, 2015; Schick, 2009; Vallmalle et al., 2014). Social insurance advocates’ arguments that programmes can and should have dedicated funding (showing the public is willing to pay for them) is opposed by public finance specialists’ dislike for dedicated revenues. Nobody should imagine, however, that these disciplinary perspectives are representative.

Advocates for longer time frames frequently argue that this will increase honesty and accountability by reducing budgetary gamesmanship (GAO, 2011: 3). U.S. experience shows that they change the repertoire but not prevalence of budgetary evasions and deceptions. The 2001 tax cuts, for example, were kept within ten-year budget control rules through having them all expire 9 months early – creating a massive and automatic tax increase that reasonable people (and the law’s sponsors) could expect would not be allowed to occur (Horney and Kogan, 2007). A ten year budget horizon has allowed budget-makers to increase the deficit in the near future but promise to “offset” those increases in the ninth or tenth year – a process that can be repeated year after year. My account of specific policy areas below will show many examples of the ways long-term “plans” get gamed.

Under the right circumstances, keeping score of budget results for more than one year serves the goal of economy by making it easier to pass deficit-reducing legislation. Once the U.S. deficit became embarrassingly large in the 1980s, it was easier to get credit for making a dent in it by spreading action over a number of years. The 1990 deficit reduction package, for example, was described as a USD 482 billion five-year package – which sounds much better than the USD 33 billion deficit reduction in the first year. Unfortunately, long-term budgeting can also lead to claiming credit for big savings that may not occur because the details have not been specified.
Long-term budgeting also has logical implications for other budgetary values. From the standpoint of efficiency, its advocates may assume that, by simply capping the spending for some agency and telling its managers to live within the allocation, they will “incent” the managers to operate in the most efficient way. But that simply assumes away the work of discovering and validating efficiencies. Normally, strict budget constraints encourage short-term coping such as starving investment. The only exception to this concern would occur if short-term investments are needed to achieve long-term savings. Then a long-term focus may counter budgetary shortsightedness. Unfortunately, arguments about long-term savings are often wrong, and sometimes intentionally deceptive, even if attractive; so long-term budgeting could increase short-term costs without helping in the long run. This problem will be discussed further in the section on health care costs.

The campaign for long-term budgeting is associated with beliefs that the primary goal for budget totals should be either to reduce deficits or to limit both spending and taxes, as in the Bowles-Simpson plan. It therefore is biased against the demand management view of economic policy. In principle, demand management can operate from any baseline balance, e.g. as variations around some norm associated with savings targets. In practice, any provisions to leave room for such responses to economic distress are suspected of being a way to bust the necessary fiscal constraints. The Peterson-Pew Commission, although a fervent supporter of long-term “fiscal rules” (2011a, c), explained the problem nicely (2011b: 19):

There are questions about how to formulate this long-term rule in a way that is both transparent and sufficiently flexible. The former is necessary... to ensure public pressure is brought to bear when the rule is violated. The latter is necessary to accommodate inevitable short-term shocks that will require deficits and thereby sustain support for the rule over time. However, experience with cyclically adjusted balanced budget rules is insufficient to judge whether they are sustainable both economically and politically.

In general, by raising the level of concern about totals, and the probability that totals could be deemed unacceptable, long-term budgeting should increase budgetary conflict and make settling it more difficult. This is a rewording and reframing of the perspective of budget hawks who believe that politicians are avoiding the tough (read: conflictual and difficult) choices and so they need to be forced to confront them through either information or procedures that create greater pressure to cut spending (or less commonly, raise taxes). More challenging goals for the totals should make it more difficult to reconcile preferences about details and totals.

2.3. Summary

Compared to traditional annual budgeting, efforts to seriously budget for the long-term might make budgeting less honest and accountable, rather than more. They can make pursuit of efficiency more difficult, require that decisions be made based on less reliable information, and increase conflict. The possible effects discussed above are summarised in Appendix 1. The sections of this article which follow provide more evidence for those descriptions.

The discussion to this point, however, has excluded perhaps the most basic question: whether long-term budgeting is even possible. The following sections of this article address a series of difficulties that are illustrated by U.S. experience. The arguments made in the following sections apply to other nations as well, but in a few cases are especially relevant to the United States.
We will see that multiyear caps on discretionary spending separate decisions about
details from decisions about totals in ways that are hard to sustain because of the
consequences for details. The U.S. Social Security programme is the best case for long-term
budget planning, but it is very difficult to forge agreement about how to do that. Both
underlying ideological disagreements about the programme and disagreements about
scorekeeping within the budget community create problems. In the case of U.S. Medicare,
there is simply no good way to make long-term budget allocations. Health care spending
control is a short-term problem; the best way to control spending in the long-term is to
control spending each year. Decisions made in one year may be superseded long before the
future years in which they were supposed to shape budget results. In short, ostensible
long-term decisions may be “unsustainable,” in part because they do not satisfy many of
the goals of budgeting.

3. Caps, Enforcement, and Discretionary Spending

Many U.S. proposals to reduce deficits over the long run through some sort of
enforcement procedure target enforcement on entitlement programmes. Some
recommend backup tax increases. To date, however, legislation that actually passed has
targeted enforcement on discretionary spending: bureau programmes that are annually
appropriated, such as the Army, Environmental Protection Agency, and National Institutes
of Health. Congress and the president have agreed on five laws with such provisions: the
Gramm-Rudman-Hollings Act of 1985 (GRH); the Omnibus Budget Reconciliation Act of
1990 (OBRA-1990); and in particular the Budget Enforcement Act, or BEA, that was Title XIII;
the Deficit Reduction Act of 1993 (also known as OBRA-1993); the Balanced Budget Act of
1997 (BBA-1997); and the Budget Control Act of 2011 (BCA).

The idea of multiyear caps developed in part from increased time frames for the
Congressional Budget Resolution. This Resolution, a set of targets and constraints first
created in the Congressional Budget Act of 1974, at first only provided figures for the
following budget year: so the resolution adopted in 1978 set figures for fiscal year 1979.
Beginning in 1980, in part to limit some games and in part to make deficit reductions seem
larger, the period was expanded to three years. So the resolution adopted in 1983, for
example, went through FY 1986. The time horizon for both budget resolutions and real
budget decisions was expanded to five years in 1990; the deficit reduction package
negotiated by President Bush with Congress covered a five-year period and the first step was
to enact a (previously long-delayed) budget resolution outlining the plan. For a few years
around the turn of the millennium the time frame for budget resolutions was increased to
ten years; it has since varied between ten and five, if Congress can agree to pass a resolution
at all (Heniff, 2015). But both the expansion to five-year resolutions and the development of
discretionary spending caps were in part reactions to Gramm-Rudman-Hollings.

3.1. Gramm-Rudman-Hollings

In 1985, a coalition of senators took legislation to raise the debt ceiling hostage,
threatening default on existing debt unless Congress passed legislation to force a balanced
budget. Since there was no majority for any particular way to do that, they demanded a
process with phased deficit reductions (ultimately five increments of USD 36 billion each,
to balance the budget at the end of five years) that would be enforced by automatic cuts
(“sequesters”) to discretionary spending programmes. We see here the origin of multi-year
deficit reduction through “enforcement” provisions. When House Democratic leadership
decided that desire to “do something” about the deficit meant they would lose conservative southerners so had to adopt some version of GRH, they worked to make it as noxious as possible to Republicans. Nobody claimed that the discretionary cuts were desirable, and almost everyone involved rationalized them as so terrible that, rather than letting them go into effect, the other side would have to give in and do something else.

The political theory of GRH was flawed because all sides had already showed that they cared more about entitlements and tax policy than about discretionary programmes, by excluding the first two from the sequester in the first place. It was easier to threaten cuts in discretionary programmes in general because they did not have to be specified in advance, so the voters would not know what to protest against. But in order to make sequesters as threatening as possible, the law was written in a way that made it as arbitrary and irrational as possible.

The result should have been predictable. Hardly anyone was willing to give up their priorities in order to avoid sequestration, because they thought the sequester terms were so obviously horrible that the other side would have to give in. Since the threat did not create agreement, when the crunch hit each year the only response on which majorities could agree was to combine a small amount of real deficit reduction with a larger amount of flim-flam. In the three years prior to passing GRH, Congress and the President passed a major deficit reduction package in 1982, a modest but real one in 1984, and a Social Security refinancing package in 1983 that included significant policy changes (especially in Medicare) that would reduce the unified budget deficit. Between GRH and OBRA-1990, they did much less.

The GRH experience led a sufficient number of centrist policy-makers to draw some useful conclusions. The first was that holding discretionary spending hostage to force action on entitlements and taxes didn’t work. Second, if policy-makers want to cut entitlement spending or raise taxes they should do that directly. Third, attaching enforcement to specific deficit targets is a bad idea because outside events can require much larger cuts than expected in order to meet the targets. Fourth, targets for total discretionary spending should bear some plausible relationship to acceptable details.

3.2. The Budget Enforcement Act and 1990s extensions.

The Omnibus Budget Reconciliation Act of 1990, and especially its Budget Enforcement Act (BEA) section, reflected all these conclusions. The BEA created discretionary spending caps as part of a broader deficit-reduction package that also included, among other provisions, Medicare savings and tax increases. The caps were defined as specific numbers for each of the following five years, but the law allowed for automatic adjustments in response to some economic factors (CBO, 1990: 5). BEA kept a sequester process, but sequesters would only be triggered if Congress acted to increase the deficit beyond what OBRA-1990 would have produced under the actual economic circumstances. So the sequester was to protect the law’s provisions, not to force future legislation. A series of points-of-order were designed to further constrain Congress to offset any new spending or revenue reductions. Both the points of order and sequester process were called “PAYGO” (pay-as-you-go) rules.

This basic approach was continued in the 1993 Deficit Reduction Act and the 1997 Balanced Budget Act. BBA-97, however, differed from the first two laws in two significant ways. Because it was passed in conjunction with the Tax Relief Act of 1997, it made no
contribution to the balanced budget that emerged (to general amazement) by the end of 1997. In addition, while the discretionary spending caps in the 1990 and 1993 laws were pretty much enforced, the 1997 caps were spectacularly overshot (CBO, 2003; GAO, 2002). The reasons for this difference in enforcement are instructive.

For the first few years, the details matched the totals because of a fortunate budgetary event, the collapse of the Soviet Union, and in 1990 a deal about the details between congressional Democrats, especially Senate Appropriations Committee Chair Robert C. Byrd (D-WV), and President Bush’s OMB Director, Richard Darman. The former allowed defense outlays to be cut by about USD 30 billion between FY 1990 and FY 1998, declining from 5.1% of GDP to 3.0% of GDP. Meanwhile, the 1990 caps allowed an increase in nondefense discretionary spending from 3.4% of GDP in FY 1990 to 3.6% in FY 1992.

The BEA targets for fiscal years 1994-1995 were also enforced, and actually lowered in the 1993 legislation, because of a powerful anti-spending mood in Congress after Ross Perot’s 1992 presidential campaign. Conservative Democrats demanded a spending “freeze” early in 1993, and helped kill President Clinton’s proposal for a short-term budgetary stimulus. Supplemental appropriations that under the BEA would normally have been allowed as an emergency were largely offset (Hager, 1993a, b, c; Towell, 1993). The Republican victories in the 1994 election put control of Congress, and especially the House, in the hands of a majority that made cutting domestic spending its highest priority. Although they could not enact their “revolution” (Drew, 1997; Joyce and Meyers, 2001), some discretionary spending cuts should have been expected even if there had been no caps.

To summarise, the caps for the first three years were based on expectations of defense constraint that was enabled by events and included a temporary increase in nondefense discretionary spending. After 1993 the caps were enforced partly by the reduced need for defense spending and partly by a strong anti-spending mood, driven by the 1992 and 1994 elections. Neither set of caps forced spending restraint that was not likely to have happened anyway.

The caps enacted in 1997 then were exceeded by large margins for two basic reasons. First, they were not needed to balance the budget – as CBO (2003: 114) later noted, “the surplus eliminated the essential purpose of the [caps] – to combat and control deficits.” Second, after years of constraint legislators in both parties wanted to spend more on some specific programmes. Republicans, for example, wanted to spend more for transportation projects and biomedical research. Republican leaders thought they had lost the veto battles with President Clinton in 1995 and were worried about maintaining their majorities. Republican leaders of the appropriations committees didn’t think they could pass bills with the required constraint on the details.

Therefore policy-makers avoided the caps through procedural moves that were available before but now had majority support. These included advance appropriations, payment delays, and defining more spending as “emergencies” so not subject to the cap. After the surpluses appeared, emergency designations grew to over USD 30 billion per year from FY 1998 on. Beginning in 2000 those measures were not large enough, so the caps were simply raised within some convenient piece of legislation. Sequesters were prevented by legislatively eliminating balances from the PAYGO scorecards (CBO, 2003, 116; also GAO, 2002, 11-12). By fiscal year 2001, outlays exceeded the caps by 15%, nearly USD 85 billion. In short, neither the immediate budget situation nor actual preferences about details supported enforcement of the caps adopted in 1997, so they were violated.
3.3. Return of the caps: 2011 to the present

The story of the current round of caps is incomplete, but fits with scepticism about both the honesty and reliability of the approach.

The 2011 Budget Control Act combined Gramm-Rudman with the cap approach from the 1990s, only for ten years instead of five. It first created a set of discretionary spending caps for fiscal years 2012-2021. Relative to CBO’s baseline, this involved cutting budget authority by 4% in FY 2012, rising to nearly 9% by FY 2021. In this sense it was similar to the 1990, 1993, or 1997 plans. But it also called for a Joint Committee process that was supposed to yield another USD 1.2 trillion of deficit reduction over the same ten years. If the Joint Committee failed, a backup sequester, as with GRH, would go into effect. The size of this sequester, however, was defined in advance: it would cut spending by a further USD 984 billion from FY 2013 through FY 2021. Half of the savings would come from defense and half from domestic spending. A portion (just below 35%) of the domestic sequester would be applied to entitlements (mainly Medicare payments to providers), but about 83% of the total sequester would come from discretionary spending.

It should have been no surprise that a political system that could not agree on different deficit reductions in July of 2011 could not agree on them a few months later. Therefore the Joint Committee process failed, and the sequester occurred in 2013. The law then automatically lowered the caps for subsequent years. The overall effect of the BCA, therefore was to create budget authority limits that were about 14% lower than the CBO baseline for FY 2013, and 15% lower for FY 2021. In policy terms, these cuts were much more drastic than they sound, because the CBO baseline only assumed that spending would increase at the inflation rate from year to year. It thereby did not allow for economic and population growth, and this effect would compound over time. As a result, the baseline already assumed spending would decline from 9.1% of GDP in FY 2011 to 6.7% in FY 2021 (CBO, 2011c: 15). The largest effect of the BCA, then, was to turn the baseline from an estimate into the law; the sequester and other cuts then reduced legally allowed spending in FY 2021 further, to about 5.7% of GDP. Over the long run the expected cuts, therefore, were much more severe than in the 1990s, especially considering that the U.S. military is much more active now than then.

What have been the effects of the BCA, and to what extent has it been enforced? Actual spending figures are not straightforward evidence for two reasons. First, for the first couple of years there were still extra domestic outlays from earlier anti-recession appropriations. Second, substantial defense spending was excluded from the targets as “Overseas Contingency Operations” that in principle would rise or fall with the course of the “war on terror.” Its trend, however, was subject to contradictory impulses: the Obama Administration’s efforts to decrease those operations, and the temptation to reclassify regular defense spending as “contingency” so as to protect it (CBO, 2016a; Davidson and Booking, 2015; Tucker, 2015). On balance, certainly by 2015 and 2016, the OCO designation was being used to exceed the caps (CBO, 2016a).

OCO manoeuvers aside, by my calculations appropriations for Fiscal Year 2016 were 10.1% below the original CBO 2011 baseline, compared to the 14.3% that would have been expected from the BCA’s enforcement. This was largely because of laws moderating the caps, ostensibly in the short term. The 2013 sequester was delayed and modestly reduced as part of the American Tax Relief Act (ATRA), otherwise known as the end-of-2012 “fiscal cliff” deal. The caps were modestly raised again in both the Bipartisan Budget Act of 2013.
and the Bipartisan Budget Act of 2015. In each case Congress and the President raised the caps for two years and then claimed that spending would return to the previous capped levels afterwards.36

As part of BBA-2015, an extra USD 32 billion was provided for OCO spending, equally divided between defense and nondefense categories (CQ, 2015a). It seems fair to project that OCO appropriations are not likely to decline as much as OCO activities, thereby providing a bit more wriggle room for defense and some international affairs spending under (or out from under) the caps (Reich, 2016).

At the beginning of 2018, President Trump and Congress passed a short-term continuing resolution, providing temporary appropriations for the federal government that included substantial further legislation including the latest “Bipartisan Budget Act.” BBA-18 raised the caps dramatically, by about USD 60 billion a year for two years for domestic discretionary spending and USD 80 billion a year for defense spending (Parlapiano, 2018). This time Congress and the President did not pretend to offset much of the increases. Projected domestic discretionary spending for Fiscal Year 2019 remains below the inflation-adjusted pre-BCA figures (Reich, 2018a), and the law still requires that it plummet after two years. There is some reason also to expect Republicans to try to spend less than the domestic caps – as is proposed in President Trump’s budget (Davis, 2018) and seen in a variety of technical manoeuvres (Reich, 2018b). So the most recent events both show the weakness of caps based on ignoring details, and that in any given year spending will be driven by the political dynamics of that year.

When offsets were provided in laws that raised the caps, they heavily relied on timing games. ATRA’s offsets were mainly back-loaded or, in the long run, cost more than they saved (Driessen and Labonte, 2015; PGPF, 2013a).37 The offsets for BBA-2013 also were heavily back-loaded (PGPF, 2013b). The offsets for BBA-2015 included some more plausible payment reductions for Medicare, some one-term benefits from selling government assets, and again were heavily back-loaded (Driessen and Labonte, 2015; PGPF, 2015; “Winners and Losers” 2015).38

3.4. Lessons of a sort

U.S. experience with 5 to 10-year caps shows, first, why they are attractive short-term political manoeuvres. Agreement is possible because a) the public cannot see who will be hurt – unlike, for example, cuts in Social Security, for which any plan must include specific changes in benefit rules; and b) the politicians cannot see who will be hurt, either – because opposing negotiators guess differently about what will be cut in a few years – for example, defense vs. domestic spending. Following the logic explained by Downs, each voter may assume that other peoples’ programmes will be cut. The problems arise later, when policymakers have to fill in the details, and can be blamed for them.

Except in 1990, the totals do not appear to have been based on any coherent understanding of the details or even a plausible rule of thumb. It is especially hard to project long-term figures when over half of the spending is for a category, the military, that is heavily influenced by world events. The “government is too big,” “increase national savings” and “satisfy the financial markets” approaches all are used to argue for lower spending, but do not set any specific bounds. Moreover, the totals are subject to perceptual distortions – such as confusing nominal dollars with spending power. Such errors have limited effects when budgeting is for one year, but their distortions compound over longer
terms. In 2011 totals were defined as targets for savings from a baseline, and since the inflation-adjusted baseline was itself, in the long run, quite strict, the long-term result could be much more stringent than the public would understand.39

Experience also supports serious concerns about how effectively caps constrain spending. They certainly create procedural difficulties. Yet when Congress and the President wanted to, they have easily exceeded the caps, with no political consequences. When legislation has offset cap increases, it has usually claimed savings years in the future, so that the long-term focus enabled higher deficits in the short term. In general, setting caps without explaining the details involved, and then offsetting some of the effects with what Robert L. Bixby and others call “gimmicks” (CQ, 2015b), does not seem to serve honesty, transparency, and accountability. Whether caps have been met or exceeded has depended largely on the political conditions at the time: public moods about deficits, elite pressures, and the partisan division within the government. Those are also the factors that normally determine annual budget action. So, while long-term caps may have had some impact on spending and deficits, much of the results in the cases reviewed here would have happened anyway. Caps have been input into later decisions but not actual long-term commitments.

Last but not least, at some points discretionary spending caps have been associated with spectacularly broken budget processes. Nevertheless, it would be unfair to conclude that the caps in the United States have caused process collapse. It makes more sense to conclude that both the caps and the collapses have been produced by the combination of intense budgetary pressure and disagreement about how to respond to that pressure.

4. “Social Security” in the United States40

In the U.S., the term “Social Security” refers to a social insurance package of age-based and disability-based pensions: Old Age, Survivors and Disability Insurance (OASDI). The old-age portion, OASI, was enacted in 1935 and first paid benefits in 1940. Disability Insurance, DI, was enacted in 1956, with first benefits paid in 1957. In 2016 OASDI accounted for 5% of GDP and about 24% of federal government spending (CBO, 2016a: 64).

The programmes in OASDI are funded mainly by a payroll tax, with contributions creating the entitlement to benefits. The contribution rate reached 6.2% of cash wages from both employee and employer (so 12.4% total) in 1990, having been raised many times before. The contributions apply up to a threshold, which in 2018 is USD 128 400. The trust funds also receive income from income tax on benefits and from interest credited on the funds’ balances. Therefore benefits are essentially paid from three streams of income: current payroll taxes; current taxes on benefits (which could also be seen as just a reduction in benefits); and the interest created by previous surpluses of taxes over benefits.

Considered in isolation, Social Security has always been budgeted for the long-term. The first Advisory Council on Social Security (1938), whose report set the stage for the important 1939 Amendments, declared that,

“V. The planning of the old-age insurance programme must take full account of the fact that, while disbursements for benefits are relatively small in the early years of the programme, far larger total disbursements are inevitable in the future. No benefits should be promised or implied which cannot be safely financed not only in the early years of the programme but when workers now young will be old.
“VI. Sound presentation of the government’s financial position requires full recognition of the obligations implied in the entire old age security programme and treasury reports should annually estimate the load of future benefits and the probable product of the associated tax programme.”

As one example of such looking ahead, the original 1935 Report of the Committee on Economic Security estimated, quite accurately, the proportion of aged persons in 1960 and 1975.41 From its beginning, the programme's financing plans included scheduled increases in the payroll tax to meet expected cost increases.

Contributions to both OASI and DI are paid into trust funds, governed by trustees who report each year on the funds' prospects for the next 75 years. In 1983, the combined funds were expected to have less money than would be required to pay full benefits by around 1 July. In response, Congress and the President agreed to the 1983 Social Security amendments. While getting through the immediate problem was the highest priority, policy-makers also sought to enact changes that would eliminate the long-term actuarial shortfall (difference between projected costs and revenues over 75 years) (Light, 1985; see also Goss, 2010; Kingson, 1984; White and Wildavsky, 1991).

The rationale for this forecasting is that people contribute to Social Security based on the promise of future benefits, so need to know that promise is secure. Benefits are related by formula to enrollees' contributions. So, unlike most programmes, there is a very specific promise and claim to that promise. The claim is not quite equivalent to a contract, because Congress is sovereign and (with the President’s support) can break its promises by changing the law. Moreover, the trust fund process establishes that beneficiaries only have a claim to the extent funds are available (see Goss 2010 for discussion). But major reduction in the promise was intended to be difficult: President Roosevelt famously explained that the payroll contributions were “there so as to give the contributors a legal, moral, and political right to collect their pensions…” 42

The 1938 Advisory Council’s statement clearly did not, however, view the programme entirely in isolation, separate from the rest of the federal budget. The quote above emphasises “the government’s financial position” as a whole, not simply the programme’s. 43 So one of the basic questions about Social Security is the relationship between Social Security's financing and the overall federal budget, which is not as straightforward as even many “experts” believe. The second is to what extent it is really practical to plan for the programme many years in the future.

DI has 75-year projections because it is attached to OASI’s financing, and was viewed as a form of retirement protection like OASI (Social Security Administration, 1986). As a result, “most analysis of the actuarial status of the Social Security programme is done on a theoretical basis where the two trust funds are considered on a combined basis” (Goss, 2010: 116). The two funds are treated as interchangeable, to the extent that under current law, the payroll contribution will have been divided six different ways between 1990 and 2019.44 Nevertheless, the logical time frame for DI budgeting is much shorter than for OASI.

4.1. The trust funds and Social Security within the federal budget.45

If spending is likely to grow over time, how can policy-makers plan to meet future costs of benefits? One approach would be to schedule, in the law, future contribution increases to match the spending increases, or spending cuts to match the contributions, or some of both to bring the two into balance. A second would be pre-funding: to build up
surpluses in the trust fund(s) that could be used to pay for some of the future costs – either through interest earnings on the trust funds or interest plus spending down the principal.

As enacted in 1935, Social Security included some moderate pre-funding, but that was pared back for many reasons in the 1939 amendments – one argument being that collecting taxes without paying benefits had helped put the economy back into recession in 1937 (for a further explanation that provides a good introduction to the programme, see Corson, 1940). But, since all covered workers (which then did not include the self-employed and agricultural workers) were paying in, and pay-outs were low for a long time because few workers had made many contributions, the OASI trust fund still far exceeded benefits into the 1950s. From about 1960 on, however, spending and tax revenue had become fairly similar, and as both grew, the trust fund balance shifted to about 100% of annual spending – a prudential margin to cover short-term bad news. At that point the programme could be described as basically pay-as-you go, with some tax increases still scheduled to cover projected spending increases, and a trust fund cushion against bad economic news.

During the 1970s a combination of policy decisions and unexpectedly bad economic news threw the programme out of financial balance. Financing reforms enacted in 1977 failed largely because of miserable economic performance over the following five years, and the actuarial projections then showed the trust fund balance falling to zero, so inability to pay full benefits, by July 1 of 1983. This led to the Social Security Amendments of 1983 and a policy change to more of a pre-funding approach (for accounts see Svahn and Ross, 1983 and Light, 1985).

The 1983 law accelerated scheduled payroll tax increases, but did not change the projected rate of 12.4% combined (for OASDI) in 1990. Nor did it schedule further tax increases in the future. It included one significant benefit cut in the future, an increase in the Normal Retirement Age from 65 to 67, which would be implemented in two stages over the years from 2000 to 2022. Its major other provisions were a mix of more immediate modest benefit cuts or tax increases, with the lynchpin being one measure, taxing benefits for higher-income beneficiaries, that each political side could count as a “win” (Democrats as a tax, Republicans as a cut). These measures were projected to build up surpluses to help finance costs through the baby boom generation’s retirement. In essence, the baby-boom “generation” would pay higher taxes than needed to pay for its predecessors’ retirement, in order to help finance its own. Current estimates say it didn’t quite work: the trust fund balance is expected to hit zero in 2034, when the oldest baby boomers will be 88 and the youngest will be 70. Still, starting at near-zero and ending up with a positive balance for the next fifty years is pretty good, though not quite as good as it looks.

In 2015 the Social Security Trustees projected that costs in 2089 would be 17.97% of taxable payroll, and income 13.32%, so the difference 4.65% of taxable payroll. By trust fund logic, however, the actuarial deficit – the change in 2015 that would finance the programme in 2089 – was 2.68% of taxable payroll. The reason is that the immediate change would create surpluses that would build up over time, and then could finance later spending by being used up (Board of Trustees OASDI, 2015: 4). Thus, the Trustees and others have long argued, acting sooner will enable less drastic change (Board of Trustees OASDI, 2015; CRFB, 2015).

4.1.1. Are the trust funds “real”?

Many budget experts and economic commentators, however, would agree (in more gentle words) with Allan Sloan’s (2010) description of the trust fund logic as, “ridiculous...
The problem is the nature of the assets in the trust funds. It is not just a technical dispute because it puts in question most of the arguments for planning ahead.

In some countries, such as Canada and Sweden, national pension funds are allowed to invest surpluses in private instruments such as bonds and equities (OECD, 2016; Severinson and Stewart, 2012; Tedesco and Shecter, 2016). In the United States, however, the Treasury is forbidden to invest surpluses in the private economy because that looks like Socialism. So it holds the surpluses as federal debt instruments. The critics' point is that when Social Security spends interest or spends down the principal, Treasury has to get the money somewhere, which means from taxes or borrowing. But if there were no trust funds, it would do the same – so the trust funds would seem to be irrelevant.

The 2001 President's Commission to Strengthen Social Security was able to quote many authorities to this effect. It cited President Clinton's OMB (“(t)he existence of large Trust Fund balances... does not, by itself, have any impact on the Government's ability to pay benefits”) the Social Security public trustees, CBO, GAO, and the Congressional Research Service to the same effect (President's Commission, 2001: 17-19). This view has also become conventional wisdom for many journalists (e.g. Jackson, 2011; Montgomery, 2011).

Yet if the trust fund is useless, then what is the benefit of the early action demanded by advocates of long-term budgeting? There are two arguments, both typified in GAO’s (2017: 22) statement that, "acting soon would allow any adjustments to be smaller and spread across more generations of participants and be phased in so that affected individuals could have time to adjust their retirement planning." The first presumes that earlier action will make it easier to pay benefits in 2034, but if trust fund balances don’t count, that isn’t true. As for the second, if the benefit cut in 2034 will be the same either way, then phasing it in ahead of time hurts anyone who retires before 2034, but it doesn’t help anyone who retires then and after. If the cut will be the same either way, they are better off waiting and seeing if it happens – after all, it might not. Moreover, a rational retiree in 2025, if told that the government would cut their benefits each year leading up to 2034 so as to help “phase in” the pain of the cut, would say, “no thanks, let me have the full benefit up to 2034, and if I want to I’ll save the difference each year.” Having the government give her less earlier surely won’t help her cope.

Aside from the fact that it doesn’t support the policy its advocates promote, scepticism about the trust funds is also a highly dubious financial and political analysis. The 1938 Advisory Council explained why:

“The United States Treasury uses the money realized from the issuance of these special securities by the old-age reserve account in the same manner as it does moneys realized from the sale of other Government securities. As long as the budget is not balanced, the net result is to reduce the amounts which the Government has to borrow from banks, insurance companies, and other private parties. When the budget is balanced, these balances will be available for the reduction of the national debt held by the public.

“Members of the Advisory Council are in agreement that the fulfillment of the promises made to the wage-earners included in the old-age insurance system depends upon, more than anything else, the financial integrity of the Government. The members of the Council, regardless of differing views on other aspects of the financing of old-age insurance, are of the opinion that the present provisions regarding the investment of the moneys in the old-age reserve account do not involve any misuse of these moneys or endanger the safety of these funds.”
Past surpluses reduced federal borrowing from, so debt owed to, the public. Less debt (interest rates being equal) means lower interest spending in the future. Lower spending on interest means more money available to pay for other government activities.

The situation is analogous to a person preparing for retirement. She could invest any extra cash so that future unearned income could replace earned income. But she could also pay off her mortgage, reducing future housing expenses. Nobody would say that reducing future costs by paying off the mortgage is “useless” as a way to prepare for retirement. Reducing future government interest costs is exactly the same principle. In 2018, the trust fund interest earnings account for what the government does not have to pay to outsiders because of the previous surpluses.

In short, the balance sheet effects of accumulating surpluses are not at all “ridiculous.” But those who want to call trust fund surpluses fake also make a further, political, argument. They claim Social Security’s surpluses made the deficit look smaller than it really was, so Congress and the President simply spent extra money (or cut taxes more) in the rest of the budget (CRFB, 2011b; Munnell, 2005; Spivak, 2010). In a separate paper (White, 2012b: 13-25) I give extensive reasons to doubt that claim. The economic modelling (Smetters, 2003, 2004) cited to support the argument has major technical flaws (Hungerford, 2009) and is based on a strange construction of budget process and history. The historical record provides no evident examples of the supposed fiscal illusions from Social Security surpluses reducing budget responsibility. One cannot even make a case that the 2001 tax cuts were enabled by Social Security surpluses, as CBO’s forecast at the time still showed an on-budget surplus after the revenue losses. There is much better reason to agree with Alicia Munnell (2005: 3), who judged that, “neither administrations nor Congress pushed less hard to restore balance on the non-Social Security side because of the surpluses in the trust funds.” Therefore the surpluses are highly likely to have reduced federal borrowing.

In the United States, the history of budget-making as I interpret it suggests that the trust funds are real, and some portion of future costs can be pre-funded by building them up. This is an argument for long-term financial planning for Social Security. The next questions are whether costs and revenues can be forecast accurately, and how likely it is that policy-makers could agree on long-term plans.

4.2. Accuracy and acting on projections

How accurate, then, are long-term projections of Social Security costs and revenues? Clearly there is some uncertainty, which is why the actuaries create high-, low-, and intermediate cost forecasts. Estimates depend on assumptions about dozens of factors. Just a few include birth rates, death rates, immigration rates, the ages of immigrants, trends in disability, growth in national income, how that growth is distributed between labour and capital, what portion of the labour share is above and below the maximum for OASDI contributions, interest rates and labour force participation.

Part of the future is relatively predictable. Spending should rise substantially relative to payroll tax revenues and GDP because the number of beneficiaries will rise relative to the number of workers paying taxes. The effect of other factors, however, is much less predictable, and that has been visible from changes over time in the projections of long-term actuarial balance.

The 1983 legislation projected actuarial balance over 75 years at 0.02% of taxable payroll (so, barely). By 1985 the estimate declined to (-0.41). By 1994, the projected trust fund balance
was down to (-2.13), with the reserve estimated to be depleted by 2029. But in 2008 the projection of the actuarial balance had improved to (-1.70), and depletion of the reserves was not expected until 2041. Then the projections worsened again, so that by 2014 the projected long-term shortfall was (-2.88), and the reserve was estimated to run out in 2033.

Some of the decline in projected actuarial balance over most of this period is simply due to time. Each year’s projection is for 75 years; each successive estimate includes one less year when the large baby boom cohort is contributing much more than it is collecting. Yet the variation from year to year has been much larger than can be explained by timing alone. The deterioration in the forecast long-range balance in the six years between 1988 and 1994 was 1.55 percentage points – close to as large as the 1.84 point improvement from the 1983 legislation (Board of Trustees OASDI 2015, Table VI.B1). The trustees’ explanations (Board of Trustees OASDI, 2015: 165-167) show how the wide variety of factors involved in estimates allow for this kind of change.

We can get another view of variation in forecasts by looking at spending projections as a share of GDP. Figures 1 and 2 give two views of how forecasts have changed over time. The 2050 forecast, highlighted in Figure 1, may look fairly stable – but the estimate made in 2008, at 5.81% of GDP, was almost a full point lower than the 1999 projection of 6.79%. Figure 2 shows projections for a series of other years, and how they have changed over time. If we were to look at the numbers behind the chart we would see, for example, that the current predicted shortfall in 2035 is 1.24% of GDP, but projected spending in that year was 6.96% of GDP in 1999 but only 6.02% today.

**Figure 1.** *Projected OASDI Spending in 2050 as percentage of GDP, by year of Forecast*

Hence the future burden of Social Security on the economy is not so easy to predict, and no one should imagine long-term budgets for Social Security can be made with the same kind of confidence as annual budgets for agencies. These overall figures, moreover, hide the greater volatility of the Disability Insurance programme.

DI is subject to some of the same long-range demographic dynamics as OASI, because older people are more likely to become disabled. The effect shows up in the numbers about fifteen years earlier both in the U.S. (CBO, 2010; Fritze, 2015) and other countries (OECD, 2010). Yet DI eligibility is much more complex. Administrative discretion matters more, and
there are many more arguments about supposed fraud and abuse (people getting benefits who should not) or inequity (people not getting benefits who should). There are two levels of choice, which vary less systematically than demographics: people have to choose to apply, but the disability determination process also must decide they are eligible. Both are influenced in part by economic conditions (Merline, 2012), for legitimate reasons. DI spending therefore has fluctuated not only with underlying economic conditions, which influence applications, but with the administrative process, which varies with legislative changes, budgetary resources, administrations’ priorities, and among states. States manage the application process in different ways and so with different initial results.

The number of applications and awards for DI therefore varies substantially more than for OASI. Increases and decreases in both have been related to policy changes, especially a crackdown on both applicants and current beneficiaries that was legislated in 1980 and implemented with perhaps unexpected zeal by the Reagan Administration (Derthick, 1990; Pear, 1982; Social Security Administration, 1986). That crackdown included an increase in Continuing Disability Reviews, designed to identify persons who had been awarded benefits but whose medical condition had improved and so no longer needed them. This programme, as administered, was widely accused of going too far, and a backlash resulted in legislation in 1984 that changed some of the terms of the 1980 Disability Amendments (Derthick, 1980; Social Security Administration, 1986). Discretionary funding for administration also affects the programme’s financing. The amount of Continuing Disability Reviews has varied substantially over the years even though CDRs on average lower total federal spending (CBO, 2010).

To summarise, DI is an entitlement but involves administrative discretion that means it does not operate remotely as automatically as OASI. It involves a much less certain risk, and most people will not end up collecting. Nobody who is paying into DI expects to collect even 50 years later – they’re supposed to be retired and on OASI by then. The promise is that DI will protect you if you get disabled soon, or at least sooner than your expected retirement age. If a change in eligibility rules is defensible, then it normally should be
implemented quickly. If current rules reward people they should not, or prevent aid to people who should get it, that should be fixed sooner rather than later. These differences between DI and OASI are likely to also distinguish disability pensions from old-age pensions in any country. For example, the issues regarding disability benefits, entitlement and administration have been far more complex and conflictual than old-age pensions in the Netherlands.\(^{56}\)

4.3. The weaknesses of the best case

OASI is the best case for long-term budgeting. Individual voters have reason to look ahead and – if they’re not told the trust funds are fake, and trust politicians to treat the assets as real – perhaps agree to pay for benefits in the future with earlier contributions. Because discretion is limited, administration is not so strongly related to spending, as long as it is not badly underfunded. Any efforts to change long-term outcomes also require specific decisions about those rules, so conscious and visible connection of details to totals. Estimates are uncertain enough that policy-makers should not respond often to changed projections. There can even be honest mistakes.\(^{57}\) But over many decades the Social Security Actuaries have earned a reputation for integrity. Building up a trust fund is a policy choice to lean in a more fiscally contractionary direction, but only modestly so and should not be viewed as especially biased.

Yet even in the case of U.S. Social Security, the conventional approach to long-term planning is excessive. 75-year forecasts even for OASI are not especially credible, and DI does not belong in those forecasts. Its spending is driven too much by short-term economic developments and how the programme is administered. Medical understandings change, and so decisions about disability change, in ways that cannot be anticipated. Administration is especially important for the value of efficiency, as both accepting applicants who perhaps should not get benefits and turning down people who perhaps should can reduce the net benefit of the programme. So it would make a lot more sense to view DI in a much shorter time frame. A third problem applies to current arguments that entitlements as a whole should be subjected to caps over a 20-year or 30-year or longer time period. Ironically, a long-term plan that separates entitlements out from the rest of the budget and sets a cap on entitlement spending would ignore how extra trust fund revenue can reduce spending in the rest of the budget.

The most severe problem with proposals to focus on long-term financing of Social Security, however, is that adding some sort of larger set of long-term budget controls or solvency rules to what already exists within Social Security will not solve the problem of how to get agreement among U.S. policy-makers about the means. There might be situations in which a long-term focus enables political leaders to compromise with each other, because very long-term measures would not excite their constituents. The 1983 legislation’s provision to raise the retirement age over forty years seems like an example. But in order for that kind of buffering from voters to work, the sides have to want to compromise. In 1983 a majority of the House and Senate thought raising the retirement age was a reasonable policy.\(^{58}\) Under present circumstances, looking at the long-term does nothing to reduce conflict. If anything, the choice between accelerating and delaying action is just a matter of tactics.

DI would have gone into deficit in 2016, and substantive response was postponed by changing the fund allocation between OASI and DI, rather than addressing DI on its own. The Trustees have argued for many years that the trust funds were not in long-term actuarial balance, and since 2013 have projected exhaustion within 20 years. From this
perspective, long-term budgeting for OASI is possible but is not being done; and the time for long-range budgeting of DI is past.

So why has there been no action on OASI or DI in recent years? It is not because DI beneficiaries are powerful “claimants.” They are poor, disabled people – which may be why U.S. policies to help the disabled are among the least adequate among all OECD nations (OECD, 2010: 85-87). Instead, policy leaders in the two parties disagree intensely about how to respond to DI’s financing issues. Republicans believe the programme is rife with abuse and encourages people who aren’t really disabled not to work. Democrats believe these claims grossly misrepresent benefit levels, the difficulty of claiming benefits, and the medical status of beneficiaries (Fritze, 2015; Kessler, 2015; Ruffing, 2015). Democrats believed they could not convince a Republican-controlled Congress, so sought to postpone the issue by reallocating funds from OASI; House Republicans tried to forbid re-allocation through a House rule (Romig, 2015; Ruffing, 2015). Eventually they backed down, and the Bipartisan Budget Act of 2015 included a three-year reallocation from OASI to DI, which was projected to build up the DI fund enough to pay full benefits into 2022. The reallocation was immediately condemned by the Heritage Foundation as robbing USD 150 billion from Social Security (Greszler, 2015). Not much room for agreement there.

The OASI story is much the same. Virtually the entire Republican party insists that only spending cuts are acceptable. More and more Democrats have concluded that a steady decline in other sources of retirement income means Social Security benefits should be increased, not cut (Altman and Kingson, 2015; Hiltzik, 2016). Even without advocating expansion, Democrats have emphasised raising payroll taxes and especially the cap on taxable payroll. Republicans, if they are willing to be specific at all, favour privatisation and forbid tax increases. Self-described centrists propose approaches that hardly anyone likes – in the Bowles-Simpson case, also relying heavily on benefit cuts (Blahous, 2010; Ruffing, 2011). The current OASI conflict is precisely the kind of situation in which attention to totals does not cause action because of disagreement about the details, for there is no majority to meet solvency targets in any specific way.

In this situation, any attempt to force immediate action is fundamentally partisan. To put this another way, time is more on the Democrats’ side. They can reasonably project that OASI’s funding is far more likely to be improved by tax increases in the future than it is today. In the early 2030s the prospect of major benefit cuts will seem much more real, and the elderly will be a much larger portion of voters – particularly in midterm elections (such as 2030 and 2034). Under these circumstances, raising taxes – including even taxing “unearned income,” or raising the threshold so high earners pay on more of their earnings – seem much more likely than they are today (Arnold, 2015).

What, then, could any efforts to further enforce long-term budgeting for Social Security accomplish? A plan might include some sort of automatic policy changes triggered by changes in projections. Yet the fight over such a plan will look just like the fight over more straightforward policy changes, and there will be no more reason to expect compromise. Or, the plan could include some sort of commission to submit proposals in response to projected shortfalls. Yet there is no reason to expect that approach to be any more successful than the 2011 commission – or various other efforts over the years. If policy-makers do not have some underlying basis for agreement, then addressing issues that do not have to be dealt with yet (because they are long-term) can only increase conflict, giving more to fight about. Postponing will be much more attractive. And so it has been.
5. American Medicare

Within long-term projections of the federal budget, Medicare is the major cause of future spending increases. For instance, the 2015 forecasts (Board of Trustees OASDI Table VB3) showed OASDI growing from 4.98% of GDP in 2015 to 6.03% in 2060, while Medicare was projected to grow twice as much, from 3.53% in 2015 to 5.7% in 2060. Indeed, health care spending is widely considered a key aspect of budgetary sustainability across all rich democracies (OECD, 2015b; White, 2014). If we combine Medicare with Medicaid and other federal health care programmes, it seems clear that the United States does not have a “deficit” or “entitlement” problem, but a health care cost problem (Aaron, 2007). Because it is the single most important factor in long-term projections of “unsustainable” budget deficits, whether long-term budgeting for Medicare is desirable or possible must be central to the case for long-term budgeting in the U.S. overall.

The Medicare programme is remarkably complicated. It involves separate insurance for hospital benefits (HI, Part A), medical benefits (SMI, Part B), and pharmaceutical benefits (Part D); payroll funding for Part A alone; enrollee premiums for Parts B and D; private insurance either as a supplement to the traditional package (“Medigap” plans) or as a replacement (“Medicare Advantage”), and shared responsibility with the state-based Medicaid programmes for costs of impoverished and disabled beneficiaries, the “dually eligible” (Cohen et al., 2015; Kaiser Family Foundation, 2015, 2016; Jacobson, Huang and Newman, 2014; Brown et al., 2014; Newhouse et al., 2015). This article cannot come close to identifying all the relevant budgetary twists that result from this complexity. Yet one thing is clear: long-term budgeting for Medicare is a truly terrible idea. Some of the reasons are unique to the U.S, but other difficulties will apply in any system.


Healthcare spending in the United States, as a share of GDP, far exceeds that in any other country. Even though government provides only about half of total U.S. healthcare spending, its share of GDP, at over 8%, is greater than in all but a few OECD countries.60 Yet the U.S. covers its population less successfully than any other rich democracy, being the only one with no national guarantee of health care services.61

Higher U.S. total spending is not caused by more generous benefits (especially for the uninsured!). Nor is it caused by higher use of services (Angrisano et al., 2007; Farrell et al., 2008; Ginsburg, 2008; Papanicolas et al., 2018; Squires and Anderson, 2015). Nor is it caused by greater prevalence of disease or worse risk factors, as American disadvantages on factors such as obesity are offset by other factors, such as low tobacco use and a younger age distribution. Nor does it buy better outcomes. In fact, the proportion of Americans that dies from conditions that would be amenable to healthcare is higher than in at least 23 other countries, including all those that might be expected to have comparable economic capacity and political responsiveness (Gay et al., 2011; Nolte and McKee, 2012).

Although Medicare has controlled spending growth more effectively than have private insurers (MedPAC, 2015: 9),62 the peculiar division between government and private finance in the United States makes population ageing seem like a much bigger budgetary problem in the U.S. than in other countries. The dynamics that will increase Social Security beneficiaries’ share of the population have the same effect on the Medicare share. This enrolment growth is a major factor in projected Medicare cost increases. (e.g. CBO, 1998, Tables 4-2, 4-5). Yet OECD analyses do not show aging having comparable effects on public
spending for health care in other countries (de la Maisonneuve and Martins, 2013), and it's important to understand why. One reason is that U.S. health care costs per person are so much higher than in any other country, so effects of aging compound on a much higher spending base. But the major reason is that aging in the United States shapes eligibility, while in other countries it only affects cost per person. In the United States, when a person turns 65 she may shift from private budgets to Medicare. Her entire cost is an increase in government spending. In other countries, the increase in spending when a person turns 65 is just the difference between her costs that year and the year before. Therefore, if all Americans were “entitled” to the same public or semi-public coverage, the effect of aging on the budget would be much lower!

Moreover, when projections show Medicare spending rising quickly over time, so being “unsustainable,” then if the same assumptions about spending growth per capita were applied to the entire health care system, it too would be “unsustainable.” For example, in 2001, when Medicare spending was expected to rise to about 8.5% of GDP by 2075, the same assumptions projected total national health care spending of 38% of GDP (Technical Review Panel, 2000: 39). This seems likely to be at least as much of a challenge to private budgets as the Medicare figure would be to the federal budget.

These facts help explain why focusing on long-term Medicare costs is a seriously flawed problem definition. Considering the overall failings of the U.S. approach to financing health care, and the relative success of Medicare compared to private insurance, treating Medicare as the problem is a bizarre policy analysis. There are strong arguments for reforming the U.S. health care system as a whole, rather than focusing on the Medicare budget (Aaron, 2007; Aaron et al., 2008).


In spite of all the exceptional characteristics of American healthcare finance in general, and Medicare in particular, Medicare still offers general lessons about long-term budgeting for health care. The most fundamental involves spending forecasts. Former programme administrator Bruce Vladeck in 1998 described the whole enterprise of forecasting costs beyond a ten-year horizon as “an exercise in comparative fantasy,” and with good reason (Pear, 1998).

Part of the problem is that short- to medium-term forecasting is also quite difficult. Policy can change outcomes, but they also depend on unpredictable behaviours within the health care system. The bottom line of Figure 3 shows projections for spending in 2000 that were made from 1994 on. The 1994 projections were more than 40% higher than actual spending (3.26% of GDP vs. 2.29%). In this case one major factor was the unexpected effectiveness of a campaign against fraud and abuse (Rosenblatt and Rubin, 2000).

Spending on the Part D drug benefit, enacted in 2003 and implemented in 2006, also illustrates the difficulty of forecasting. As CBO (2014: 5) reported, its estimate of spending for fiscal year 2013 was nearly twice the result (USD 99 billion vs. USD 50 billion). Total spending over the period from 2006-2013 was 36% less than the original estimate. Part of the difference is explained by lower enrollment than expected (CBO, 2014: 11), but more was due to unexpected trends within the pharmaceutical market, which affected spending by all payers (CBO, 2014; Hoadley, 2012). Not only did many brand-name drugs go off patent, allowing substitution by lower-priced generics, but, less predictably, “new brand-name drugs (which
tend to be more expensive than older brand-name therapies) were introduced at a slower rate than in the late 1990s” (CBO, 2014: 2). The trend through 2013, however, could not be used to predict the future. In 2014 Part D payments jumped by 8.3%, with the Medicare trustees (2015: 106) projecting a 15.1% increase in 2015. What happened? Prescription drug spending in the entire system accelerated sharply in 2014, largely “caused by increased spending on new medicines (particularly for specialty drugs such as those used to treat Hepatitis C), a smaller impact from patent expirations than in previous years, and price increases for brand-name drugs.” (Martin et al., 2016).

I do not want to suggest policy-makers could do nothing about such developments. They certainly could regulate prices for new drugs more strictly. Absent much stricter short-term cost control measures, however, Part D experience illustrates the difficulty of even short- and medium-term estimates.

Long-term forecasts are volatile from year to year for four reasons. First, short-term experience can change the baseline on which long-term estimates compound (Sommers, 2010). The immediate good news in the late 1990s, as the other lines in figure 3 show, drove down longer-term forecasts as well as the forecast for 2000. Second, short-term experience may cause estimators to re-think their judgments about underlying dynamics. Third, the assumptions used for longer-term projections are rather arbitrary, and may be changed because none of them are particularly compelling. Last, legislation can change policy and virtually mandate changes in the forecasts – whether that legislation is sustainable or not.

All four factors explain the dramatic variation shown in Figure 4, which gives the projections for Medicare spending as a share of GDP in 2050 made from the 1994 Trustees Report through the 2015 Report. Readers can see how Medicare’s estimates depend on far more than demographics or basic economic trends by comparing Figure 4 to the Social Security forecasts in Figure 1.

The dramatic fall from projecting 2050 spending of 8.64% of GDP in the 1994 report to 4.79% in the 2000 report was largely due to the real life trends that lowered spending as of
2000. At that point, however, the Trustees changed an arbitrary assumption about long-term trends. Until then, the actuaries figured they had no basis for projecting spending after 25 years, so assumed spending per enrollee would decline to equal the growth rate of per capita GDP. The 2000 Technical Review Panel recommended raising the assumption to growth 1% higher than per capita GDP, and adopting that assumption raised the 2050 projection to 6.0% of GDP. The long-term nature of the changed assumption is reflected in the lines in Figure 3, in which projections for 2050 and beyond rose much more steeply in 2000 than projections for earlier years. As Figure 3 also shows, all projections increased substantially in 2004, due to the effect of a policy change: creation of the Part D benefit.

The sharp drop in spending estimates as of 2010 is due in part to the short-term measures in the Patient Protection and Affordable Care Act, as amended by the Health Care and Education Reconciliation Act (PPACA hereafter). PPACA included provisions that immediately reduced spending, and raised revenues, in reasonably predictable ways. The big shift in the long-term projections as of 2010 however, resulted from language in the law about how payments to health care providers would be calculated “in 2011 and later” – that is, forever. The law said that instead of updating payments to each category of provider according to changes in the underlying “prices that providers must pay to purchase the goods and services they use to provide health care services,” the Medicare programme would raise its providers’ prices about 1.1% per year more slowly, on the theory that the productivity of Medicare services could improve at the same rate as productivity increases in the whole economy (Medicare Trustees, 2010: 2, 8). Thus the legislation essentially offset the actuaries’ previous assumption that in the long-term spending per capita would rise by one percentage point over the trend of per capita GDP. The U.S. health care system could surely do a better job of capturing technological improvement in the form of greater value for payers rather than higher income for the medical industry. Yet nobody should pretend to know if the productivity adjustment to prices can be enforced for ten years, never mind forever.

More recent forecasts provide an even better example of the arbitrary nature of long-term estimates. The 2015 Medicare Access and CHIP Reauthorization Act (MACRA) overturned projected major reductions in Medicare’s payments to physicians, which had been part of baseline estimates even though they were not about to happen. CBO (2015c)
projected the change in physician payment would raise (estimated) Medicare spending by USD 175 billion over ten years. Yet MACRA enacted some annual restraints on increases in physician fees, which were assumed to apply forever. As a result, the next Trustees report projected slightly smaller Part B spending beginning in 2059 (Medicare Trustees, 2015: 71, 98). CBO, however, still makes its long-term projections by assuming that “excess cost growth per beneficiary” [over growth of per capita GDP] “will trend smoothly to a rate of 1.0 [percent] between 2027 and 2046” (CBO, 2016b: 16). In 2015 CBO stopped publishing projections of Medicare spending more than 30 years in the future, so is not officially disagreeing with the Medicare Actuaries about longer-term trends; but it is making a different, equally arbitrary, assumption. It is easier to see why both assumptions are wrong than why either would be right.

5.3. Why health care cost control is a short-term problem

The evident unlikelihood of credible long-term estimates is reason enough to reject long-term budgeting for Medicare. Yet developing credible long-term measures is equally difficult.

For Medicare and for any national health care programmes, spending depends on how the programme is administered. In the United States, that means an immense variety of choices about payment rates and rules for thousands and thousands of services. It also depends on decisions outside the government, especially the continual expansion of the medical realm in modern society. This is not so much “technology” as the medicalisation of problems. The medical industry seeks to expand spending not only through traditional political action, but by promoting new services and drawing attention to problems that were not previously viewed as medical, such as Attention Deficit Disorder (for further discussion see White, 2014).

Spending control is made more challenging by the fact that within the health policy world there is a plethora of arguments that long-term savings can be achieved from an agenda of delivery or payment reforms such as electronic health records, “paying for performance,” and “integrating care” through “Accountable Care Organizations”. Distinctly unimpressive results from such efforts (Goldsmith et al., 2015; Nichols et al., 2017) does not seem to reduce advocacy for such measures (as in Nichols, 2017). In one telling example, leading health economists criticised CBO’s forecasters for insisting on “traditional evidence” that measures would succeed (Cutler, Davis and Stremeckis, 2009: 10). Ideas that have better records of success are not as attractive to analysts as these policy unicorns (Vladeck, 1999), and ideas that have records of success, such as price regulation, have bitter enemies (the people who lost incomes). Ideas that have not reduced spending meet less fervent opposition from providers, who may even nod solemnly and say those are really good ways to make health care “higher value” (White, 2013b).

The basic politics and economics of health care provide a further obstacle to long-term policies. Health care cost control is a continual war, a contest for territory (money) between powerful adversaries. On one side are the payers – governments, sickness funds, private insurers. On the other side are the providers, in the U.S. a sixth of the national economy. They are some of the smartest and most accomplished people in any country; they have massive resources; at least the physicians have some significant credibility among voters. That does not mean the medical industry cannot be beat in political battles. Even in the U.S. the relative cost-control success of Medicare shows that governments have power too. In many other countries, the payers are better-organised, more purposeful, and the
medical industry has less independent power. But it is a battle that has to be fought again and again. The providers will attempt to bust any spending control method, and eventually budget-makers will need to adjust their tactics in response.

None of this is to say there can be no merit to experimenting with unproven cost control methods in the hope that they will someday “bend the cost curve.” Any policy analyst will have ideas for measures that cost more up-front but could be useful in the long run. Some may be worth trying. Budgeting should not have such a short-term perspective that useful investments are crowded out.

Yet unless policy seeks budget savings simply by cutting benefits, there is no credible way to “save” in the long run. And cutting benefits based on claims about total spending decades in the future is not likely to be desired by voters, either now or in the future. It is not likely to be representative, will be based on highly questionable forecasts, is likely to involve measures designed to hide effects (such as vouchers), and so seeks to avoid relating details to totals. Measures that seek to control spending in other ways will have the same questionable forecasts and, if they include details such as long-term assumptions about payment rates, relate details to totals only in the sense that because the details are made-up, the totals are made-up too.

The history of Medicare spending projections shows that they should not be a guide for action. They are too volatile, too sensitive to unpredictable events, and too reliant on highly dubious assumptions. Reducing spending in the short run however, can have dramatically favourable effects on long-term spending estimates, as occurred after 1997. Better yet, it saves money quickly and these effects – so long as policy-makers don’t go out of their way to reverse them – seem more reliable than any assumptions about what will be done 10 or 20 years in the future. Sherry Glied and Abigail Zaylor (2015: 1), based on their own analysis of this history, draw the same conclusion. “The ability to make significant programmatic changes in the short-term,” they write, “…and the difficulty of forecasting costs over the long-term suggest that policymaking can be most effective by addressing immediate needs.”

6. Summary and conclusion: the search for budget responsibility

The campaign for long-term budgeting in the United States is promoted by highly prominent individuals and organisations within the budgeting community. They have asserted that binding decisions about programmes and revenues should be made with the intent of meeting some sort of targets for the totals 30 or more years in the future. Possible results that far or further away are described as crises that need to be addressed immediately. The advocacy is so broad that it nearly constitutes a conventional wisdom.

In this analysis I have sought to remind readers that good budget processes serve a wide range of goals. Long-term budgeting, based on the logic and evidence discussed in this article, is inimical to increasing efficiency, relating details to totals, transparency, accountability, the values of representative government more generally, accuracy and honesty. It is designed to serve its sponsors’ economic theories and overall political priorities over others. Long-term budgeting might serve the goal of economy, but even that claim may often be false, because the policies involved are not themselves sustainable.

The operation of budget processes depends on how they solve budgeting’s core decision-making quandary, which is how to match social preferences about totals to social preferences about details. Social preferences are inconsistent even when individual preferences are not. The most basic problem is not how “guardians” can protect the budget
against “claimants.” It is how to resolve conflict between different views of how to fit totals and details together. Therefore common political arguments about how to budget are based on a misguided view of the problem. Advocacy for long-term budgeting is an example, because it is based in part on a belief that making the totals seem like more of a crisis will strengthen guardians. This does nothing to dampen conflict and enable compromise between competing views.

In the U.S., old age pensions (OASI), offer the strongest case for making long-term decisions within a broader budget process. Long-term commitments are part of the programme’s promise, and so their financing should be part of its decision-making. Compared to either taxes or other spending programmes, OASI is designed in a way that allows the least inaccurate estimates, and poses fewer issues of transparency or efficiency.

Yet even the case for long-term decision-making about OASI, unfortunately, has been made in logically questionable ways. It requires giving credence to pre-funding through the trust funds, so is contradicted by common claims from many of those who claim to be budget guardians, that the trust funds are fake. Unfortunately also, if building up trust funds is agreed to be helpful, that still does not solve the problem of how to do so – from whom to collect extra money in advance (if anyone) and which benefits to reduce so as to allow greater accumulations. It was difficult enough to forge agreement in 1982-83; it seems impossible now. Any efforts to “reform” budgeting to create a doomsday mechanism to force earlier action will face the same problem, only redefined as how to agree on the mechanism.

Within American Social Security, Disability Insurance illustrates how policy dynamics should determine attitudes towards long-term projections. Nobody should imagine that DI can reasonably be budgeted far into the future. DI is more similar to Unemployment Insurance (UI) than to OASI. Although some of its spending demand is predictable in the long run, other factors, including economic cycles but also medical norms and administrative processes, are much more short-term and unpredictable. In the U.S., defining DI as a long-term problem, in tandem with OASI, has made it easier to avoid budgetary responsibility. If it were viewed as a freestanding short-term programme like UI, its financing would already have been addressed.

Long- or medium-term caps on discretionary spending are another form of long- or medium-term budgeting. American experience with these measures suggests a few lessons. They have become the default method of deficit reduction because it is possible to adopt caps without specifying details. Hence they privilege economy (in the short run at least) over transparency, accountability, and efficiency. As a way to manage the government they can only do harm, as the specific details of programme operations are given much less weight in budget decisions. When the time comes to make details fit the cap, this may lead to intense conflict that can only be resolved by either fraud or busting the caps. Enforcement has depended very much on each year’s political conditions, rather than on the supposedly binding nature of the caps. None of these patterns suggest that long-term caps on discretionary spending meet the goals of budgeting as normally understood. All suggest that using discretionary caps as backup enforcement for long-term budget plans is a bad idea.

Long-term Medicare spending projections, combined with assumptions they will not be financed, explain much of projected increases in U.S. federal government deficits and debt. Indeed, any U.S. long-term budget problem is actually a health care cost control problem. Yet long-term budgeting for Medicare cannot be responsible.
This paper has shown the immense variation in Medicare spending projections. Both positive and negative shifts in the projections have often involved arbitrary assumptions about long-term spending trends – such as whether the long-term rate of increase in per-enrollee spending will be the same as or more than the growth of per capita GDP, or for how long the programme could enforce fee updates that are below the rate of general inflation. Trends in Medicare costs per enrollee depend partly on policy and partly on behaviour by patients and the medical industry. The medical industry is in the business of outflanking or overpowering spending control policies. In spite of these pressures, public policy can restrain growth in spending per capita. Experience in other countries has been far more successful than in the United States, and Medicare has been more successful than private insurers. There is no way, however, to make policies to resist the providers that can be expected to work for long periods of time. Therefore the most practical way to constrain Medicare costs in the long run, while attending to other budgetary values, is to adopt short-term measures that actually work, and so reduce the base on which future increases compound. Then do it again, again, and again.

Long-term risks are a reason to make both annual budgets and other policies with careful attention to these risks (remembering they are risks, not certainties). Like managers of any household, political authorities should try to anticipate serious problems down the road. Yet it is not at all obvious that the annual budget process is the place to address those challenges. In general, they are discreet challenges associated with specific policies. Some, such as the effect of global warming on nations with extensive coastlines, or possible needs for higher military spending, are barely being considered by budget experts at the moment. Others, like pension and health care costs, are major policy issues in their own right. In the case of health care, viewing the problem as long-term is misguided. It should be addressed year after year in annual budgeting. In the case of pensions the extreme importance of benefits to many voters, the ability to direct specific revenues towards those programmes, and the need for credible long-term promises all make the annual budget process a questionable place to address the policy. Budget reform is much less important than simply building the coalition to enact stand-alone legislation.

In the United States, the campaign for long-term budgeting is not only unwise but dangerous. Budgeting has always been hard. At the moment it is almost impossible. There is a political war, and the stakes are what kind of government the United States will have, for whom. In this context the call for long-term budgeting has two major flaws beyond its normal problems. First, it largely supports one side, by defining government as the problem. The December, 2016 budget process “reform” proposed by House Budget Committee Chair Price, his parting statement before his departure to serve (briefly) as Secretary of Health and Human Services, illustrates the link (Kogan and Reich, 2016). Second, it provides no reason for anyone to compromise. Rather than cooling the temperature of political debate, the campaign for long-term budgeting raises the temperature, by providing one more source of heat, one more “crisis.”

Notes

1. The work for this analysis was originally commissioned by the National Budgeting Roundtable, sponsored by the William and Flora Hewlett Foundation, and a version that focuses more on potential reforms in the United States was produced as White (2016). I thank the Roundtable leadership for permission to use some of the same material in this article.
2. CRFB is an advocacy organisation with a board of directors that includes many formerly leading officials in U.S. budgeting. Six were CBO Director; six were OMB Director; three chaired the House or Senate Budget Committee and two were Comptroller General.

3. The term “pay-as-you-go” usually refers to provisions to inhibit legislative changes, but in the context of the sentence must mean provisions that would change spending trends if events were different than expected.

4. One can see versions of this view in some of the proposals mentioned above, and it was expressed in the National Budgeting Roundtable meeting that discussed the first version of this paper.

5. This view was supported in February of 2016 by Jon Blondal, Head of the Division of Budgeting and Public Expenditures in OECD’s Directorate for Public Governance and Territorial Development, in a personal conversation. He suggested the closest equivalent would be Medium Term Expenditure Frameworks in a few countries, which to some extent allocate resources over the term of a government.

6. Thus principle 4 is, “Ensure that budget documents and data are open, transparent and accessible.” Principle 5 is, “Provide for an inclusive, participative and realistic debate on budgetary choices.” Principle 6 is, “Present a comprehensive, accurate and reliable account of the public finances,” and Principle 10 is, “Promote the integrity and quality of budgetary forecasts, fiscal plans and budgetary implementation through rigorous quality assurance including independent audit.” Principle 5 speaks to procedural fairness, mentioned just below.

7. This is the logic of the OECD’s call for creating MTEFs, with the model being the Netherlands, in which budget priorities and expenditure targets, subject to various adjustments, are negotiated as part of the coalition agreements when a new government is formed (Bos, 2008). In the U.S. the governing coalition tends to change every two years, with each congressional election, so the equivalent approach would be to move Budget Resolutions to a biennial cycle (White, 2009a).

8. As one OMB official described neutral competence in an interview in the early 1990s, “that meant if it was a Republican administration trying to minimise cost; if it was a Democratic administration, trying to maximise value for the money we had.”

9. These problems may help explain why OECD’s Recommendations (2015a: 9) call for “routinely presenting performance information in a way that informs” budget decisions, but using that information to only a limited extent – what some call “performance-informed budgeting.”

10. Thus both increased conflict which prevents timely allocations to agencies, and “sequester” processes that disrupt agency planning and administration, are bad budgeting on the management dimension.

11. Typical more recent examples include CRFB (2011a) and 10 Ex-Chairs (2011). Others include Democratic economists advocating for budget balance in 1980 (as described in White and Wildavsky, 1991), the rationale for the Reagan budget plans (Stein, 1984: 264, 266, 269; White and Wildavsky, 1991) and the pressures on President Clinton in 1993 (Woodward, 1994). Contrary experience both in the 1980s (with big deficits, modest growth but a long stock-market boom) and recent years (with historically low interest rates) has never shaken advocates of market expectations views.

12. Another good example of the right-wing view is Boccia (2014) – objecting to IMF (2014) because that document allowed for the possibility of revenue increases.

13. OECD 2015a, p. 10. Some of the other recommendations, such as comprehensive accounting (Recommendation 6) could also be considered part of the concern for honesty and accuracy.

14. The classic statement about repeated relationships generating trust is Wildavsky (1964). Roy Meyers (1994) provides an extensive but presumably not complete inventory of the deceptive tactics available even within the context of repeated interactions. White and Wildavsky (1991) provide numerous examples of deceptive budgeting being used to enable agreement.

15. The problem can be illustrated most easily with a smaller system, like a condo association. Imagine a condo association with ten units and monthly payments of USD 225, so annual revenue of USD 27 000. Imagine it has ten expenditures, each costing USD 3 000 per year: landscaping, cleaning public spaces, utilities for public spaces, heating and cleaning the pool, maintenance of the exercise room – whatever. So spending is USD 30 000, revenues are USD 27 000, and the deficit is USD 3 000. The simplest approach would be to raise the fees to USD 250 per month. But Mr. Simpson may say, “wait a minute, I don’t swim, let’s cut out spending for the pool.” That is a consistent position – but the other nine owners like to swim. Mr. Bowles may say, “I’m allergic to grass, let’s pave the yard and then we won’t have to pay to maintain it.” But the other nine think pavement is much uglier than grass. Mr. Cote thinks people can bring their own towels to the exercise room and shower in their condos, but the others think the showers and towels are worth
the money. In this scenario all ten owners object to raising the fee; all ten owners are intellectually consistent; none is a hypocrite; but 90% oppose any individual’s solution. A good budget process might then involve carefully reviewing each expense; determining if any efficiencies are possible; some bargaining among the owners, and perhaps a package of specific changes adding up to USD 1 500 per year and an increase of the monthly payments to USD 237.50. The example can be extended to allow for borrowing. In short, the condo association would go through an iterative process in which preferences about the details and totals were adjusted. Traditional budgeting works in this iterative manner.

16. This point is made from a different perspective by Robinson (2016: 46). White and Wildavsky (1989) provided examples of how the policies to meet the goal of balancing the budget in 1993 could seem highly unacceptable on the policy merits.

17. The Dutch approach to setting a Medium Term Expenditure Framework begins with consideration of details in the form of “bottom-up calculated levels of expected public expenditure in constant prices” (Bos, 2008: 30). So targets are based on consideration of effects on details.

18. Readers may compare U.S. institutions to the OECD Recommendation (2015a).

19. See the discussion below and White (1998b).

20. One classic example involved the “Alternative Minimum Tax” (AMT). For an explanation and history through 2012, when it was finally “fixed,” see www.taxpolicycenter.org/briefing-book/what-amt. For an analysis at one point during the saga, see CBO (2008).

21. The rhetoric of “tough choices” deserves a separate discussion. Briefly, budget hawks appear to believe “tough choices” to reduce deficits are substantively easy, since the costs of higher deficits are so severe, so only “tough” for cowards. If the policy costs of the details are taken seriously, “tough choices” are actually tough, so not obvious. For more discussion see White (2010).

22. The summary here is based on the much more extensive account in White and Wildavsky (1991), chapters 19 and 21. See also Leloup (2005).


24. The combined effects of the two bills on deficits for Fiscal Years 1998 to 2000 would have been roughly zero – if the discretionary caps were enforced, and they were not. See Table 1 in CBO 1997b and the summary table in CBO 1997c. The 1997 BBA was not about deficit reduction, but was a short-lived truce in the budget war between congressional Republicans and the Clinton administration, possible because the deficit was going away anyway.

25. The story is discussed in the Postscript to White and Wildavsky (1991); as we put it (p. 583), “our sources report that [OMB Director] Darman told the appropriators that he would get them a domestic increase, but he would have to hide it in the baseline so Congressional Republicans wouldn’t object.”

26. Joyce and Meyers (2001) emphasise Bud Shuster and the Transportation Committee, but the doubling of NIH spending was initiated by House Appropriations Labor/HHS Subcommittee chair John Edward Porter (R-IL), with support from party leaders.

27. The problem, as staff and members explained in interviews I conducted at the time, was that colleagues who supported the caps in principle or when voting on budget resolutions would not want the blame for the cuts in actual appropriations, so would vote against them. The bitter pill of “more spending” was sweetened a bit by increasing the volume (and more modestly the spending total) of earmarks for which members could claim credit. Republicans preferred to get credit by passing tax cuts, but Clinton would veto that (see White, 2015b).

28. See GAO 2002 (p. 34) and CBO 2003 (pp. 114-16). For some reason CBO and GAO have very different figures for Fiscal Year 2000: GAO says USD 30.8 billion and CBO USD 44 billion.

29. Author’s calculation from data in CBO 1997b, p. 90; CBO 2003, Table A-2.

30. Author’s calculation from data in CBO 2011a.

31. Savings from these cuts were assumed to lead to further savings from borrowing less and paying less interest.

32. These were CBO’s estimates from CBO 2011b, Table 2. For attempts to make sense of the BCA and its terms, see CBO, 2011a, c; Kogan, 2012, 2013; Austin, 2014; Driessen and Labonte, 2015.

33. Author’s calculation from data in CBO 2011a,c.
34. To trace the use of the OCO designation as a share of defense spending over time, one can consult reports on the office of the Department of Defense Comptroller, at http://comptroller.defense.gov/budgetmaterials/budget2016.aspx.

35. I used data from Driessen and Labonte, 2015; CBO, 2011a; CBO, 2011b.

36. For an overview see Driessen and Labonte, 2015. These varied measures addressed more than discretionary spending (for ARRA the discretionary aspect was a small portion of the law) and each had its own peculiar politics. On the two Bipartisan Budget Acts see CBO, 2015a; Greenstein, 2015; House Committee on Rules, 2015; Krawzak, 2015; Parrott et al., 2013; Tucker, 2015; “Winners and Losers” 2015. The final decisions implementing some implications of the 2015 BBA are reported in Brown et al., 2015.

37. ATRA also included massive deficit increases relative to the baseline because of its revenue provisions, so in context whatever it did to the sequester was decimal dust. We should remember that the baseline was not politically credible. Some portion of the Bush tax cuts would have been continued; they certainly would not have all expired at once; and provisions like the A.M.T. had to be fixed. Nevertheless, ATRA certainly could have reduced revenues less than it did, so can hardly be seen as an example of budgetary responsibility.

38. There was also some controversy about a change in Social Security payment rules, which was viewed by most Social Security experts as correcting a minor loophole and didn't amount to enough money to be mentioned in offset summaries. See NCPSSM 2015.

39. I have not discussed here, but should at least note, the further complications involved in defining categories within the caps. This can lead to a lot of games at the boundaries, such as putting breast cancer research into the defense category, or redefining the nature of “security.” Any efforts to separate “investment” will be especially subject to manipulation and conflict. Reischauer (1993) provides useful commentary on the problem.

40. There is a lot of controversy about Social Security, as well as an extensive literature. My own overview of the history and issues is White (2003). It does not seem to me that much has changed, in terms of the positions taken and their merits, since then; and much of the discussion that follows is based on that work.

41. The report estimated that the population aged 65 or over would be 9.3% in 1960 and 10.0% in 1975. The actual figures were 9.2% in 1960 and 10.5% in 1975. See Committee on Economic Security, 1935; United States Census Bureau, 1975. Common claims that Social Security was designed for the 1930s and is in that sense outdated are falsehoods. For critique of some more egregious examples, see Altman and Kingson, 2010.

42. The comment was reported in a memorandum written by Luther Gulick, which can be found at www.ssa.gov/history/Gulick.html Note that compulsory contributions were also intended to ensure that individuals took some responsibility for their own retirement, rather than rely entirely on others who were “more prudent” (Musgrave, 1986: 69-70; Penner, 1994: 4).

43. The Advisory Council also assumed that eventually some of the programme would be paid for from general revenues, on the grounds that it would help citizens other than the direct beneficiaries. For example, helping older people retire would help younger people find employment, and maintaining income for the elderly would help the economy.

44. Self-employed individuals pay both parts, but the equivalent of the employer contribution can be deducted from calculated income for income tax purposes. The history of payroll tax rates can be seen at www.ssa.gov/oact/progdata/oasdiRates.html.

45. The argument I make here views the trust funds with much less scepticism than I expressed in White and Wildavsky, 1991 (pp. 315-317). My excuse is that at that time I was focused on short-term budgeting, and had yet to investigate the history of Social Security financing. For a much more extensive analysis see White, 2012b. Readers interested in technical operations of the funds should see Pattison, 2015.

46. In 1949 the trust fund end-of-year balance was more than 16 times that year’s expenditure; in 1954 it was more than five times as high; 1957 was the first year in which expenditures exceeded payroll tax contributions to OASI, so trust fund earnings became part of financing the costs. Calculations here and the data for descriptions that follow are from Social Security Administration 2015, Tables 4.1 and 4.3. The best account of policymaking through the 1977 legislation is Derthick (1979).

47. The trust fund ratio (end of year balance to year’s expenditure) was projected to rise to 2.58 in 2000, 3.90 in 2005, 5.0 in 2010, and 5.44 in 2015 before declining slowly over the following 45 years.
48. The caveat is that some legislation since 1983 did improve the programme’s financing. For instance, the 1990 Omnibus Budget Reconciliation Act increased revenues by adding new state and local employees to the programme. Projected trust fund ratios based on the 1983 Amendments are taken from 1984 Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds, Table 33, p. 78, and from the same table from the 1985 Annual Report. For how the figures changed over the years, see Table 4.A3 in Social Security Administration, 2015.

49. Actually they said that the imbalance was 2.68%, but a 2.62% change would, for various reasons, be sufficient. I’ve used the 2.68% estimate because it allows for a cushion equal to one year’s spending in 2089. The Trustees (2015: 6, 12) also estimated that the programme could be put into 75-year actuarial balance by cutting benefits 16.4% immediately, while if benefits were cut on a pay-as-you-go basis, that would be a 21% cut in 2034, and 27% in 2089.

50. The same could have been done with Social Security surpluses, though it would have made only a modest difference. For discussions see Aaron and Reischauer (1998); Van de Water (2014); White (2003).

51. Also, CBO’s estimates differ from the Trustees’. Compare CBO (2015b), Exhibit 2, to Board of Trustees OASDI (2015), Table VI.G4.

52. A good recent example is the controversy over Senator Rand Paul’s claim (Kessler, 2015) that, “if you look like me and you hop out of your truck, you shouldn’t be getting a disability check. Over half of the people on disability are either anxious or their back hurts – join the club. Who doesn’t get a little anxious for work and their back hurts? Everybody over 40 has a back pain.”

53. CBO (2010: 2) gives a good example. A deaf person might receive accommodations from some employers but not others. When the economy turns sour, his disability might mean he cannot work. The Americans With Disabilities Act may have led to more people becoming eligible for DI because they could find work for a while (Merline, 2012).

54. These kinds of behavioral and administrative uncertainties led the Social Security Board Actuaries in the 1930s to believe it was very difficult to produce reasonable cost estimates for a programme (Kearney, 2005/2006).


56. Readers of this journal can surely think of many citations about this issue; one for U.S. readers would be Fultz (2015).

57. CBO made a quite surprising error recently; see CBO (2015b) and the note at the bottom of page 16, “text and table corrected on February 10, 2016.” The explanation is at www.cbo.gov/publication/51232. Arguably, this is an example of how having multiple analyses helps encourage accuracy. Then again, the SSA actuaries had not made the mistake to begin with.

58. As described in White and Wildavsky (1991: 323-325), some senior House Democrats supported the idea but did not say so publicly because they wanted to make House Rules Committee Chair Claude Pepper think he could defeat it, so he would allow a vote on the floor.

59. See also the comparison between average DI and OASI benefits at Social Security Administration (2016a, Table 2), and the discussion in CBO (2010).

60. By my calculations, from data at www.oecd.org/els/health-systems/health-statistics.htm, U.S. public spending was below Belgium, Denmark, France, Germany, Japan, the Netherlands, and Sweden. The U.S. might be above some of those countries if the figures counted capital investment, or the budgetary effects of the tax expenditure which encourage employers to offer health care benefits. In 2013 U.S. total spending as a share of GDP was 16.4%, compared to 11.1% for the next highest countries (OECD, 2015c: 165, 167).

61. Among OECD member nations, Greece and Poland also have sizeable portions of the population not covered, while the benefit package in Korea and Mexico is for a smaller portion of costs than is covered for the insured in the United States. See OECD (2015c), pp. 20, 25.

62. Medicare’s better performance over time is due to its better ability to restrain prices. A recent study (Wallace and Song, 2016) provides evidence that per capita costs go down when people switch from private insurance to Medicare, for the same reason.
63. Conventional “wisdom” in general exaggerates the effects of aging, in spite of more recent disagreement from almost every major study. For summaries of the literature see sources cited in White (2007), the discussion in de la Maisonneuve and Martins (2013), and the discussion in OECD (2015b). This is not to say aging does not matter, especially if a country is at the far end of the aging distribution and has had strong cost control policies for the system as a whole, as is the case for Japan. The United States is the opposite case.

64. White (2013a) is an overview of the PPACA that attempts to put it in perspective for international readers.

65. Health care policy wonks debate such issues as whether health care is subject to “Baumol’s Disease,” or why technology might drive spending up. For one review of such issues see White (2014).

66. The issue here was the Sustainable Growth Rate Mechanism (SGR), and its technical failings. For explanations of the policy and its problems, see Farb (2015); Fontenot, Brandt and McClellan (2015); and Laugesen, Wada and Chen (2012). I would add only that such measures for volume-adjusted fee constraints can work much better if properly designed; see White (2013b).

67. These small effects are seen in Figure 4: the slight uptick in the 2025 line in 2015 and the slight downtick of the 2065 line.

68. Even a skeptic like myself. I would like the federal government to pay for medical education so physicians do not feel they need to make huge incomes to pay off their huge debts. Then I’d ratchet down fees over time to collect more savings than the cost of education.

Bibliography


Butler, S. (2016), Testimony Before the Committee on the Budget, United States House of Representatives, 9 June.


Corson, J.J. (1940), Reasons for the 1939 Amendments to the Social Security Act, Memorandum to staff of the Bureau of Old Age and Survivors Insurance, 10 January, www.ssa.gov/history/reports/1939no3.html.


House Committee on Rules (2015), Bipartisan Budget Act of 2015, Section-by-Section Summary, docs.house.gov/meetings/RU/ RU00/CPRT-114-RU00-D001.pdf.


Long-Term Budgeting: A Cautionary Tale from U.S. Experience


ANNEX 1

**Likely differences between traditional and long-term budgeting for pursuing goals of budgeting**

**Democratic Accountability and Transparency**

Traditional: Decisions are overseen by current voters and include budgets that affect them. Information can be better or worse, depending on institutional arrangements. Multiple effects, both totals and details, can receive comparable attention. Long-term commitments on some details might be made more credible with dedicated funding. Credibility based on funding may be seen as “binding” future decisions in a way that reduces democratic control by future voters. May reduce “fiscal space” so flexibility to respond to new needs in the future.

Long-term: Policies about totals, if “enforceable,” presume decisions must be made in advance without consulting voters to whom they will apply. So policy-makers who make these decisions are not accountable to the voters who will be affected. Aggregated decisions (caps, targets) will involve less transparency about the details involved than in traditional budgeting. Decisions to reduce projected long-term spending by reducing commitments (e.g. privatisation) may create “fiscal space” in future but also will constrain future choices about those programmes – e.g. privatisation is not easily reversed. Policy is less transparent to the extent it is justified by less credible information.

**Efficiency and Economy**

Traditional: Allows detailed analysis of agency work plans. Information is relatively up-to-date. Some bias against capital investment that could achieve savings (e.g. favours leasing over purchase). But can focus on efficiency.

Long-Term: Information will be less accurate, and work plans make little sense; Could give credit for savings from investment – if the claims are accurate; Must emphasise economy much more than efficiency, since information to judge efficiency will not exist.

**Economic Policies and Balance**

Traditional: Could pursue multiple theories – household management, Keynesian demand management; targeting national savings rates, targeting “size of government,” satisfying supposed demand by the financial markets. Some of these theories are more credible than others, but all are contested. Critics believe it is biased towards deficits and so excess stimulus, insufficient savings, larger government, and unhappy financial markets. All of these judgments are in some sense in the eye of the beholder.
Long-term: Difficult to pursue Keynesian demand management, as such flexibility requires exceptions to the targets, and provision for those exceptions will be viewed as evading controls. Intended by its advocates to reduce deficits, increase savings, limit government and calm financial markets – so enforce specific economic policies. Actual effects are uncertain, though bias to limit totals for taxes and spending appears most likely to succeed.

Setting Priorities

Traditional: The routines of annual budgeting are incremental. So change in priorities is an exception, rather than the rule. May be related to events or to changes in political control of the government, so in that sense representative. If there is a bias, it is in favour of the status quo.

Long-term: In principle, it requires that the government at one point in time be able to bind decisions of future governments, so can short-circuit representation. It is normally proposed in order to change priorities, generally against entitlement programmes. Yet in the United States, medium-term (5- or 10-year) policies have been biased against discretionary programmes. There are reasons to believe long-term approaches are biased against raising revenues. So in practice, long-term budgeting tends to be less neutral in terms of policy priorities.

Relating Details to Totals

Traditional: Totals and details are considered in relation to each other, in an iterative process. Some recent movement towards more of a medium-term top-down approach, but most versions begin with some assessment of the set of concerns about details. Measures to increase efficiency (e.g. “scrub the estimates”) are designed to help resolve conflicting preferences.

Long-term: Strongly top-down, as long-term estimates for details make little sense. Exceptions require that spending or taxes be based to a great extent on relatively predictable factors such as demographics. Much of the traditional work of budgeting (e.g. search for management efficiencies, creating work plans) becomes irrelevant.

Managing Conflict

Traditional: Very difficult because of conflicting preferences, both among political forces and between details and totals. Might be dampened through incrementalism, but incrementalism may be more a result than cause of containing conflict. Occasionally can be reduced by obscuring consequences. Divided government makes it worse.

Long-term: In the United States, designed to increase conflict by making policy-makers confront “tough choices.” This normally means making the deficit look bigger or more frightening, which in turn should make it harder to meet goals for totals with acceptable details. However, a long-term focus can backfire by giving opportunities to postpone choices to a promised future. This tends to be associated with deviations from the final value...

Honesty

Traditional: Depends on professional budget staff with norms of honesty, transparency in the sense of clear statements of plans, multiple checks on projections and subsequent auditing. Also encouraged by repeated interactions, so dishonesty that is discovered can be
sanctioned in the next year’s decisions. Nevertheless there can be a wide range of gaming behaviours.

Long-term: Advocates claim it will be more honest because directs attention to concerns that would otherwise be ignored. But could be less honest because a) enables greater vagueness about details; b) projections involve greater uncertainty so multiple checks can create confusion more than control; c) auditing is irrelevant; d) errors or misrepresentations would be discovered much later, so the prospect of sanctions for misrepresentation is much less immediate, if it exists at all.