MATTHEW WEINZIERL
Harvard University

Seesaws and Social Security Benefits Indexing

ABSTRACT The price indexation of Social Security benefit payments has emerged in recent years as a flashpoint of debate in the United States. I characterize the direct effects that changes in that price index would have on retirees who differ in their initial wealth at retirement and in their mortality rates after retirement. I propose a simple but flexible theoretical framework that converts benefits reform first into changes to retirees' consumption paths and then into a net effect on social welfare. I calibrate that framework using recently produced data on Social Security beneficiaries by lifetime income decile and both existing and new survey evidence on the normative priorities Americans have regarding their Social Security benefits. The results suggest that the value retirees place on protection against longevity risk is an important caveat to the widespread enthusiasm for a switch to a slower-growing price index such as the chained CPI-U.

The indexation of Social Security benefit payments may seem like an issue about which only an economist could get excited, but it has emerged in recent years as a flashpoint of debate in the United States. In his 2014 budget, President Obama proposed changing the price index by which retiree benefits are adjusted for inflation. In brief, the White House expected the change to lower the growth rate of benefits for all retirees, although at advanced ages that change would be offset by progressive "benefit enhancements." Because it would not be tied to an increase in the starting level of those benefits, the president's proposal was expected to reduce the total present value of benefits. It was explicitly intended to appeal to Congressional Republicans eager to reduce future spending on Social Security, but it was deeply unpopular with many of the president's fellow Democrats.[1] When negotiations on more general fiscal

1. The indexing change was intended to cover tax brackets, as well, and the president's proposal was thereby intended to be a net positive contributor to the government budget.

policy challenges yielded little progress over the ensuing year, the president removed the proposal from his 2015 budget. His spokesperson made clear, however, that changes to indexation were still on the table if included in broader budget deals.

While the overall fiscal implications of benefits-indexing reform have been widely discussed, this paper's contribution is to explore both the positive and normative aspects of its distributional consequences across the population of retirees.[2] In particular, I study the direct effects that changes in benefits indexing have on retirees who differ in two important ways: initial wealth at retirement and mortality rates after retirement. I propose a simple but flexible theoretical framework that converts benefits reform first into changes to retirees' consumption paths and then into a net effect on social welfare. To provide quantitative results I use recently produced data on the net worth, benefit levels, and mortality risks of Social Security beneficiaries by lifetime income decile. Finally, I introduce survey evidence on the priorities Americans have for Social Security in general, a first step in pinning down the normative implications of these effects of indexing reform. The specific questions I use in the survey take a novel form that may be useful for estimating normative preferences across a wide range of policy issues.

In brief, I find that a useful metaphor for thinking about the direct effects of indexation on heterogeneous retirees is a playground seesaw, where two facts about retired households in the United States push down on opposite ends.

Pushing down on the left end of the seesaw (that is, toward a path of increasing real benefits) is the large majority of Social Security beneficiaries, who worry about outliving their private assets and having to rely nearly exclusively on those benefits to fund expenditures late in life. In fact, a core purpose—and achievement—of Social Security is to prevent the elderly from falling into poverty as they age (Englehardt and Gruber 2004). As has long been understood, for instance in the work of Martin Feldstein (1987), benefits that rise in real terms over retirement and are therefore back-loaded later in life will provide valuable protection against longevity risk for retirees with positive private wealth (throughout this paper, I assume

2. The effects of moving to a slower-growing price index, such as the chained CPI-U, on the paths of benefits across retirees have been analyzed by a number of researchers. See CRFB (2013) and Olsen (2008), for example. These prior analyses did not translate the effects on benefits into implications for consumption or welfare, and they did not compare alternative indexing schemes—the two main contributions of this paper.

that private annuitization of wealth outside of defined-benefit pensions is unavailable). A faster-growing price index will therefore generate welfare gains through its effects on these households.

Pushing down on the other end of the metaphorical seesaw (toward a path of decreasing real benefits) sit the poorest retiree households, who also have the highest mortality rates (see Duggan, Gillingham, and Greenlees 2006). Most directly, the poorest retirees sit on this end because, with little wealth at the start of retirement, they benefit less from Social Security's effective annuitization. A more subtle reason is that a faster-growing price index that back-loads the present value of benefits has the effect of redistributing, through an actuarially unfair adjustment, some of the total value of benefits away from poorer retirees when mortality is inversely related to income. To the extent that these retirees are the ones who most need support from Social Security, a faster-growing price index thereby generates welfare losses.

In other words, heterogeneity across retiree households means that any given reform to benefits indexing generates effects with exactly opposite welfare implications. In this way, benefits-indexing policy inevitably has distributional consequences and, as will be evident when President Obama's proposal is considered, may even be used to pursue distributional goals.

It is important to emphasize from the start that this paper focuses on the direct effects of benefits-indexing reform on retirees, abstracting from a number of general equilibrium effects and other factors that matter for the optimal path of benefits and that, therefore, ought to be part of a comprehensive evaluation of indexing reform.[3] Most prominently, changing the path of benefits may affect individuals' labor effort and saving decisions during their working lives, but my calculations hold households' preretirement behavior fixed. Similarly, I do not consider the implications of benefits-indexing reform for the accumulation of the economy's capital stock, and I abstract from the controversial possibility that benefits paid earlier will yield gains to those households that can achieve a higher rate of return in the private investment market than they obtain from the natural rate of return of a pay-as-you-go Social Security system (see Feldstein

3. In principle, as suggested to me by Martin Feldstein, the design of the optimal path of benefits and the identification of an ideal price index are separate tasks. If one believes the path of real benefits has been chosen optimally in current policy, such a separation is natural. In this paper, I explore the question of how proposed price indexes affect the path of real benefits and, therefore, retirees and social welfare.

1987, 1990). Finally, technological change, especially in the context of medical care for the elderly, may affect the optimal response of policy to an increase in real benefits and therefore matter for the choice of indexing.[4] This paper's omission of these factors is not meant to imply that they can be ignored. Instead, I omit them to better focus on one piece of that broader question.

This paper also abstracts from several complications specific to the Social Security system that may matter for the results but that would make the analysis and intuition for the results substantially less straightforward. In particular, I do not model spouses' joint decisions about benefits or surviving spouses' decisions about benefits options, instead treating the *household* as the unit of analysis; I do not allow for early or late retirement, instead having all households retire at the same age; and I do not include the disability benefits portion of Social Security in the analysis. Microsimulation models that capture much of the complexity of the actual Social Security system, for instance the MINT model[5] as described by Karen Smith and Melissa Favreault (2013), may be useful for including these features in future analyses.

Which side of the seesaw carries more weight? I show that the answer depends on both positive factors, about which there is some good existing evidence, and normative factors, about which there is very little evidence. Concerning the positive factors, I show that a large majority of retirees are likely to sit on the left end of the seesaw, that is, favor a steeper path of benefits that effectively annuitizes more of a given retiree's total wealth. Moreover, the simulations below suggest that the gains to the poor from front-loading benefits are much smaller, whether measured in consumption or individual utility terms, than the gains to the majority of retirees from back-loading benefits. These positive results suggest that the direct effects on retirees from front-loading benefits, such as by switching to the chained CPI, are likely to generate a net loss of welfare. That is, they will generate such a net loss unless society puts a strong normative priority on those retirees who are the poorest and shortest-lived relative to the rest of

4. Suppose that advances in medical care for the elderly allowed them to purchase a higher quality of life at a lower real cost, such as through the introduction of a new product. Their real benefits would rise in this case, but so too would their ability to generate extra welfare with additional resources. In that case, it may be important to target benefits toward households with high marginal utilities of consumption, not those with low values of real benefits.

5. MINT stands for Modeling Income in the Near Term; as of 2013 this model was in its seventh iteration.

the population, in particular relative to poor and middle-class retirees who outlive their life expectancy at retirement. In other words, the results of this paper suggest that the value retirees place on protection against longevity risk is an important caveat to the widespread enthusiasm for a switch to a slower-growing price index such as the chained CPI-U.

To explore the normative aspects of this question, I consider two classic normative criteria and generate novel opinion survey evidence on the relevant preferences of Americans. The two classic criteria would endorse opposite reform proposals as simulated here: that is, the utilitarian criterion would endorse back-loading, while the Rawlsian criterion would endorse front-loading. Survey respondents put equal value on increasing benefits to poor retirees who die young and poor retirees who outlive their life expectancies, and they put substantially less value—perhaps even negligible value—on increasing the benefits of average retirees. These results are inconsistent with either a utilitarian or a Rawlsian criterion on its own, but applying them to the simulated reform results suggests that back-loading of benefits is likely to generate net welfare gains, at least in its direct effects on retirees, as is the case under the standard utilitarian criterion.

From a policy perspective, the net positive welfare implications for retirees of the direct effects of moving to a faster-growing price index might be expected to produce political support for such a reform. However, two political realities make that support less likely, namely, public opposition to benefit reductions and pressure from some policymakers to lower total Social Security spending.[6] To see why, note that such a reform automatically means either a decrease in initial benefits for retirees (if total spending is held fixed) or an increase in total spending (if initial benefits are held fixed). Taking those political realities into account, the results of this paper shed some light on the specific reform President Obama proposed in his 2014 budget. That proposal, which was designed to reduce total spending by maintaining initial benefit levels but slowing their growth rate, used "benefit enhancements" at advanced ages to protect some of the effective annuitization that front-loading would otherwise have sacrificed. As I show below, the progressive design of those benefit enhancements meant that they would provide this protection largely to lower-income households.

The president's proposed reform would thus simultaneously achieve the positive effects of front-loading on the poorest, shortest-lived retirees and

6. Of course, the indirect effects of reform not included in this paper's analysis, such as the effects on private saving and capital accumulation, may also explain resistance to reform.

the positive effects of back-loading on the poorest, longest-lived retirees, and it would bring a substantial net welfare gain under the utilitarian, Rawlsian, or survey-based normative criterion.[7] Of course, that reform would generate losses as well, reducing the well-being of the higher-income half of the retiree population and therefore potentially having disincentive effects that would reduce its appeal in a more comprehensive analysis. The president's proposal thus illustrates the inherent connection between benefits-indexing policy and the redistributional role of Social Security.

The paper proceeds as follows. Section I reviews how Social Security uses indexing today, lays out the seesaw metaphor described above, and briefly summarizes the empirical literature behind the factors at each end. Section II presents a simple model that allows us to analyze these direct effects on retiree households with a small set of positive and normative parameters. Section III simulates a version of that model using U.S. data and considers three prominent indexing reform proposals: the chained CPI-U, the CPI-E (an experimental series calculated by the BLS "using households whose reference person or spouse is 62 years of age or older"), and the chained CPI-U augmented with late-in-life "benefit enhancements" as proposed by President Obama in 2014. Section IV presents novel, but far from definitive, opinion survey evidence on the normative components of the model and uses that evidence as well as conventional normative criteria to provide suggestive welfare evaluations of the direct effects of the three policy options. Section V extends the analysis to include several aspects omitted from the baseline case, and section VI concludes.

I. Background and Key Considerations

The current Social Security system uses indexing—that is, adjustments of nominal values over time—in three ways. First, it scales the income earned during a beneficiary's working years into current dollars when calculating the value at retirement of his (or her) total accumulated Social Security earnings. Second, it indexes the bracket points of the progressive function that converts scaled lifetime earnings into a monthly benefit. Third, it indexes benefits upon retirement. For the first two instances of indexing, the current system uses a wage index; for the third it uses the CPI-W, the consumer price index for urban wage earners.

7. Note that the front-loaded element of the proposal adds to its appeal under the Rawlsian criterion, but not under the utilitarian or survey-based criteria.

These three instances of indexing can be seen as serving different purposes. The first, which I will call *earnings indexing,* is most naturally seen as trying to capture the natural rate of return of the pay-as-you-go (or "unfunded") Social Security system, which is closely related to the growth rate of nominal wages.[8] The second, which I will call *brackets indexing,* tries to preserve the desired progressivity (across lifetime earnings levels) of the system despite changes in the wage distribution and nominal wages. The third, which I will call *benefits indexing*, tries to protect the real value of retirees' benefits over time, although as emphasized throughout this paper it also has implications for the effective progressivity of the system due to differences in mortality by lifetime income levels. It is benefits indexing that has been the focus of public debate, and it will be my focus in this paper as well.[9]

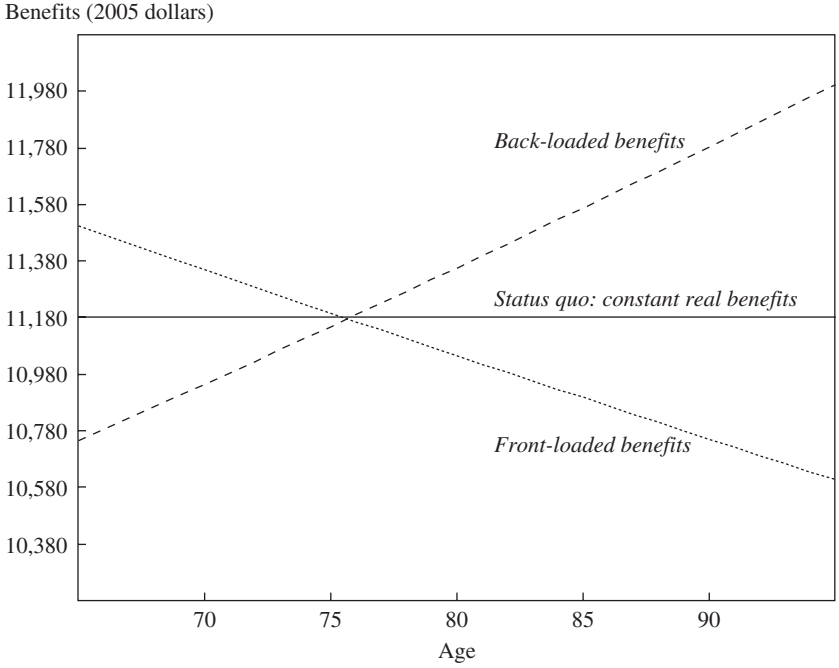
I.A. Budget-Neutral Benefits Indexing

This paper's baseline analysis focuses on budget-neutral benefits-indexing reforms. By "budget-neutral" I mean that the expected present value of benefits (across all individuals in an age cohort) is unaffected by the way benefits are indexed. Therefore, in the analysis below, in which I consider a shift to an index that causes a steeper rise in benefits over time, I adjust (down) the starting value of benefits for all beneficiaries in the cohort by the factor required to keep the expected present value of total benefits the same. The assumption of budget neutrality is not necessary, but it allows us to focus on the direct effects of the time path of benefits rather

8. Earnings indexing could, in principle, serve many purposes. Because the life cycle path of earnings varies systematically with the value of lifetime earnings, the choice of indexing will tend to favor some earners over others. One could try to use that choice, therefore, as a new optimal tax instrument that would relax the classic efficiency-equality trade-off. Similarly, one might try to take advantage of the effect that expected earnings indexing has on the extent to which workers view the payroll tax as a tax, rather than as a form of saving. These are purposes that can be more directly pursued by adjusting the history-dependent redistributive elements of Social Security, such as the replacement rates in each income bracket.

9. There appears to be little interest in reform to the other two uses of indexing. Even the Bowles-Simpson proposal (NCFRR 2010), which suggests changing the bracket points to increase progressivity, does not change the methods of earnings indexing or brackets indexing. This is somewhat unfortunate, in that changes to the method of earnings indexing hold substantial promise for more closely aligning the current system with its true "natural" rate of return. In particular, earnings indexing could include changes to projected beneficiary-worker (dependency) ratios and aggregate life expectancies. Brackets indexing would be a simple way of implementing a limited version of the inequality adjustments suggested by, for example, Shiller (2003).

Figure 1. Benefits Received under Three Paths with the Same Expected Present Value for a 65-Year-Old Male Retiree[a]



Source: Author's analysis; see text for data sources.
a. Based on a 65-year-old male retiree with median household earnings and average mortality rate.

than on their level. As I show in section V, reforms that are not budget-neutral can be analyzed using this paper's approach as well, and the main lessons are unaffected.

A simple but useful observation about changes to budget-neutral benefits indexing is that their effects on benefit levels are highly concentrated toward the beginning and end of retirement, as illustrated in figure 1.

To produce figure 1, I start with the initial annual Social Security benefit for the median retired household from the fifth decile of household lifetime earnings, as calculated by John Karl Scholz, Ananth Seshadri, and Surachai Khitatrakun (2006) (referred to hereafter as SSK) and shown in real (2005) dollars. I assume the "status quo" policy would provide a constant stream of real benefits at this level (in section V.A, I show that alternatives to this assumption do not change the lessons of the baseline analysis). As alternatives to a constant real benefit, I consider two paths: a front-loaded path, in which real benefits grow at an annual rate of –0.27 percentage points, and

a back-loaded path, in which they grow at +0.37 percentage points. These alternatives correspond to two prominent proposals for benefits indexing reform: namely the use of the chained CPI-U and the CPI-E indexes calculated by the BLS. Using average mortality rates from the Social Security Administration (SSA) and an annual discount factor of 0.96, I adjust initial benefits under these two alternatives to ensure that the expected present-value total cost of each path is the same as for the status quo.

Figure 1 makes clear the roots of the seesaw metaphor: the sizable differences in benefits early and late in retirement across benefit paths. Retirees who value Social Security's insurance against longevity risk, and especially those who come to rely on Social Security benefits because they outlive their private savings, will prefer the back-loaded benefits path. Retirees with little private wealth or high mortality risks, and especially those who do not survive to advanced ages, will prefer (or will have benefited most from) the front-loaded path. I now turn to a brief discussion of the existing literature on these two competing features in the retired population.

I.B. Evidence on Variation in Private Retirement Savings

An extensive literature has examined whether retirees enter retirement with sufficient assets to sustain their economic well-being. In general, the results have drawn a qualitative distinction between the status of the large majority of the Social Security population and approximately the bottom quintile of retirees. Reassuringly, most retirees appear to reach retirement with sufficient assets (both Social Security and non-Social Security) to smooth shocks, supplement Social Security benefits, and maintain what a rational life-cycle consumer with his or her lifetime earnings history would plan as an optimal path of expenditures. The bottom quintile, in contrast, enter retirement (or soon find themselves) almost entirely dependent upon Social Security benefits and other transfers. A number of studies have obtained such a result, and some have drawn policy implications from it. I briefly review several of them here, although of course the literature is too large for me to do this subject full justice in such a short discussion.

Rudolph Penner and Karen Smith (2010) summarize the findings of Smith, Mauricio Soto, and Penner (2009), who use Health and Retirement Survey (HRS) data from 1998 through 2006 to study the assets held by retirees, including the expected present value of their pension and Social Security benefits. They conclude:

> The net worth of those in the top quintile of the income distribution increased until age 85. . . . For those in the three middle quintiles, net worth began

> declining after age 70, but only very slowly. Evidently, the vast majority in this portion of the income distribution will die with a significant amount of assets. Few older households, including those with little income, used home equity to finance retirement consumption. The bottom income quintile never accumulated much wealth and spent their assets quickly, leaving them dependent on Social Security and whatever DB pensions they had earned. . . . Our results are reassuring for households in the top 80 percent of the income distribution, but the data indicate that the lowest income quintile quickly becomes almost wholly dependent on Social Security after retirement. . . . Reformers must be sensitive to the heavy dependence on Social Security in the lowest part of the income distribution. (Penner and Smith 2010, pp. 1–2)

Consistent with these findings, David Love, Michael Palumbo, and Paul Smith (2009) note that "it is reasonably well known that retirees in the bottom quintile of the income distribution (conditional on their age and marital status) rely almost exclusively on defined-benefit pension benefits, Social Security benefits, and other government transfers to finance spending" (pp. 191–208). Barbara Butrica, Joshua Goldwyn, and Richard Johnson (2005) also use HRS data and find that "individuals in the lowest income quintile consume between 99 and 107 percent of their after-tax income plus annuitized assets." Michael Hurd and Susann Rohwedder (2008) focus on preparedness for retirement across levels of educational attainment and estimate that 17 percent of married people and 36 percent of singles are not adequately prepared for retirement (meaning that they are likely to exhaust their wealth before death).

Finally, SSK use HRS data and a dynamic lifecycle optimization model to show that only a small minority of individuals are failing to save adequately to sustain desired consumption paths. While doing so, they find that Social Security wealth dominates for at least the bottom lifetime-income decile, arguably the bottom three deciles, of retirees.[10] As discussed below, I will rely on these authors' research for estimates of retiree wealth and Social Security benefits by income group.

I.C. Evidence on the Relation between Income and Mortality

It has become a staple of commentary on the fiscal health of Social Security that mortality differences across income groups matter for the true impact of a range of proposed reforms, such as the full retirement age (see Krugman 2012, for example). While the literature quantifying these

10. SSK find that the under-accumulation of wealth is driven not by lifetime income per se but by being single rather than married, because single retirees have systematically lower incomes. I do not distinguish between single and married households in this analysis.

mortality differences is less developed than the literature on wealth, I draw heavily on a few relatively recent contributions.

Gopi Shah Goda, John Shoven, and Sita Nataraj Slavov (2011a) use mortality estimates for the top and bottom halves of the earnings distribution to reach this dramatic conclusion:

> Under the assumption of constant mortality across lifetime income subgroups, the Social Security system is progressive regardless of the measure shown. However, a good deal of the progressivity is undone or even reversed when differential mortality is taken into account. The results are similar for both stylized earners at different points of the earnings distribution and actual workers' earnings histories. . . . Rather than analyzing the mortality differences between those in the top and bottom halves of the lifetime earnings distributions, we would have liked to have the information by lifetime income decile so that we could examine the mortality experience of the genuinely poor vs. those at other parts of the distribution. It seems likely that the extent of mortality inequality is even greater than reflected in the top half/bottom half analysis. (pp. 2, 8)

In fact, it appears that variation in mortality does widen at the extremes of the income distribution. James Duggan, Robert Gillingham, and John Greenlees (2007) use administrative Social Security data to show a consistently positive relationship between average age of death and lifetime earnings deciles.

Related to Goda, Shoven, and Slavov's suggestion that more disaggregated estimates would yield additional insights is the work of Hilary Waldron (2007), who uses SSA data to characterize life expectancy for men by income quartile at 5-year increments from age 60 to age 85. A complementary data source is provided by Barry Bosworth and Kathleen Burke (2014), who use the HRS to calculate life expectancy at age 55 for men and women as well as relative mortality rates for men and women ages 50–74 and 75+ by income quintile. Both of these sources show that retirees in approximately the bottom quarter (for example) of the lifetime earnings distribution have life expectancies 15 to 20 percent shorter than those in the top quarter prior to retirement. These gaps are larger than those between the second and third quarters of the income distribution, and Waldron's estimates suggest they are not narrowing over time. In section III, I will explain how I use these results.

II. A Partial Reform Approach to Optimal Benefits Indexing

In this section I lay out a simple formal structure through which to model how benefits indexing reforms turn into changes in the consumption paths of retirees and how these consumption changes can be aggregated into a

measure of social welfare.[11] As noted at the beginning of this paper, this analysis focuses on the direct effects of benefits-indexing reform on retirees in the context of heterogeneity in initial wealth and mortality risks, setting aside a number of other factors that matter for a more general approach to the topic of optimal benefits indexing. In particular, I abstract from any distortionary effects that changes to the method of Social Security benefits indexing might have on labor supply or on the savings decisions of households during their working lives.

In the model, there are I types of Social Security beneficiaries, indexed by $i \in \{1, 2, \ldots, I\}$ and equally prevalent at the time of retirement, $t = 1$. Type indicates the level of lifetime income y_i, the level of nonannuitized wealth $A_{i,1}$ available at $t = 1$, and age-specific mortality risks $m_{i,t}$. A more general model would not impose a one-to-one link between net wealth, mortality risk, and lifetime income, but the theoretical and (especially) empirical challenges to the analysis are substantially reduced with this assumption. The probability of individual i being alive at age t is $\Pi_{\tau=1}^{t}(1 - m_{i,\tau})$. Because the use of private annuities in the United States is quite limited (see Brown and others 2001), I assume annuitization of A_i is unavailable or unappealingly costly.

Upon reaching retirement, each beneficiary receives streams of real-valued Social Security benefits denoted $\{B_{i,t}\}_{i,t}$ and (possibly zero) defined-benefit pension benefits denoted $\{P_{i,t}\}$ for type i at age t (note that all quantities in this paper's analysis are real, not nominal, unless otherwise stated). In the status quo policy, I assume that this stream is constant in real terms, so that $B_{i,s}^{SQ} = B_{i,s}^{SQ}$ for all ages s,t. A reform to the method of benefits indexing generates a stream of changes in benefits that I will denote $\{dB_{i,t}\}_{i,t}$. Note that I treat $B_{i,t}$ as an after-tax benefit, implicitly assuming that reform to benefits indexing does not change the tax rates on retiree benefits.

Although in principle a reform could take a wide range of forms, in this paper I am especially interested in one class of reform:

$$B_{i,t} = \lambda B_i^{SQ}(1 + \pi)^t \text{ for } t \in \{1, 2, \ldots, T\}, \tag{1}$$

such that

$$\sum_i \sum_t \beta^t \left[\prod_t (1 - m_{i,t})\right](B_{i,t} - B_i^{SQ}) = 0. \tag{2}$$

11. I consider a relatively narrow set of reforms that deviate only slightly from the status quo policy, so it is natural to use this so-called "partial reform" approach of Guesnerie (1977), Feldstein (1976), and more recently Saez and Stantcheva (2014).

where β is the uniform discount factor in the economy (I consider heterogeneity in β in section V). This class of reforms scales the initial benefit level by the factor $\lambda \geq 0$ and grows that scaled benefit by the rate π each year, such that the total present-value cost of benefit payments is the same in the status quo and reform policies. For example, a reform that increased the initial benefit level and then reduced the rate of growth in real benefits would have $\lambda > 1$ and $\pi < 0$.

Individuals solve a standard utility-maximization problem once they reach retirement. They use their accumulated assets and their streams of Social Security and defined-benefit pension benefits to fund consumption in each period they are alive, and they obtain time-separable utility from that consumption. Note that there is no uncertainty in the utility they obtain from spending in the future.[12] Utility is zero when the individual is not alive and there is no bequest motive. (In section V, I show that the results are robust to adding a bequest motive.) Individuals are subject to the (real-world) constraint that they cannot borrow against future Social Security or defined-benefit pension benefits. Formally, individual i solves

$$\max_{\{c_{i,t}\}} E[U_i] = \sum_t \beta^t \left[\prod_t (1 - m_{i,t}) \right] u(c_{i,t})$$

subject to a constraint that (non-Social Security) net worth must be nonnegative at all points during retirement:

$$A_{i,t} \geq 0, \text{for all } t \in \{1, 2, \ldots, T\},$$

where $A_{i,1}$ is given and

$$A_{i,t+1} = (A_{i,t} + P_{i,t} + B_{i,t} - c_{i,t})(1 + r),$$

where $(1 + r) = \beta^{-1}$ is the annual return (net of taxes) that an individual may earn on net wealth.

Note that in this model, were households able to fully annuitize their wealth, they would choose a constant consumption level throughout retirement. Without such full annuitization, mortality risk will cause the household's optimal consumption path to decline throughout retirement until

12. Medical expenditure shocks have been shown by many previous researchers to be important for retirees' decisions and welfare. Though such shocks are not included in this paper, upward shocks to the marginal utility of spending at later ages would likely increase the appeal of the back-loaded benefits streams.

reaching the level of annuity benefits (provided by Social Security and defined-benefit pensions, if applicable). After that point, the household will be dependent on these benefits to fund consumption.

The expected change in social welfare from reform $\{dB_{i,t}\}_{i,t}$ is evaluated as the weighted sum of the welfare values of the consumption changes it causes. In particular, social welfare is denoted *W,* so the change in social welfare from the stream of changes in benefits is

$$dW = \sum_i \sum_t \left[\prod_t (1 - m_{i,t}) \right] \beta^t \frac{dc_{i,t}}{\{dB_{i,t}\}_{i,t}} g_{i,t}, \tag{3}$$

where $\beta^t dc_{i,t}/\{dB_{i,t}\}_{i,t}$ denotes the present value of the change in consumption by type i in year t in response to the change in policy, and $g_{i,t}$ is the marginal social welfare value of a present-value unit of consumption for a beneficiary of type i in year t.

It is important to note that these $g_{i,t}$ parameters can take essentially any values, although Pareto efficiency would require them to be non-negative. This flexibility enables us to use a wide variety of welfare criteria, including those inferred from public opinion, to evaluate policy reforms. An alternative formal approach would locate the welfare costs from the low lifetime utility of the shortest-lived, poorest retirees in their own utility functions, perhaps by having their utility be a highly concave function of total consumption in retirement or some other version of time nonseparability.[13] In that approach, helping those retirees would be a matter of insurance, not redistribution (this logic is related to the justification Rawls offers for the maximin priority). It is far from clear that individuals have such preferences, however, so I take the approach that granting large weights to those worst-case outcomes is a normative decision by society, not a feature of individual preferences.

III. Simulated Effects of Benefits-Indexing Reform Proposals

To simulate the effects of benefits-indexing reform, I need to specify functional forms and parameter values for the preceding section's model, determine the values of the model's key empirical inputs, and choose candidate reform policies.

13. I thank Robert Hall for suggesting this discussion.

III.A. Functional Forms and Parameter Values

The per-period utility function takes the familiar form of constant relative risk aversion:

$$u(c_{i,1}) = \frac{1}{1-\gamma}\left((c_{i,1})^{1-\gamma} - 1\right), \tag{4}$$

where $\gamma = 3$ following SSK.[14] I also follow SSK in setting the annual discount factor $\beta = 0.96$, and I assume that the return to saving $(1 + r) = \beta^{-1}$.

III.B. Data on Initial Wealth, Benefit Levels, and Mortality

To determine the key empirical inputs to the model, I use estimates drawn from the existing literature on Social Security. I divide the population of retiree households into deciles by lifetime income, so $I = 10$ and each type $I = \{1, 2, \ldots, 10\}$ corresponds to a lifetime income decile. The use of ten types is made possible by recent empirical work estimating household wealth, benefits, and mortality data at that level of disaggregation. Some of those data are not available by gender, so I treat the household as the unit of analysis throughout.

For the initial wealth and benefits levels of retirees, I rely on SSK, which is a carefully and uniquely detailed source of these data; no other source of which I am aware provides both median overall (non-Social Security) net worth and median (present value) Social Security and defined-benefit pension wealth data by lifetime income decile. This level of detail is especially important for capturing the right end of the seesaw: for example, data that divide the population into quintiles, or that group households by point-in-time income rather than lifetime income, can obscure the difficult position in which the lowest decile of retirees appear to find themselves.

To infer annual benefit amounts for both Social Security and defined-benefit pensions, I use average mortality rates (for men) in the United States and the same real interest rate r as in SSK to calculate the constant

14. This value for γ is toward the upper end of typical ranges for this parameter, which measures the degree of the individual's risk aversion. Although a high value for γ may be appropriate if retirees are generally more risk-averse than the average person, I have also run the analysis assuming that $\gamma = 1.5$. All qualitative results described in the baseline case hold there as well, although the welfare gains generated by the Hybrid Progressive Reform are smaller in size. Intuitively, with less concave utility from consumption, that policy's redistribution is less valuable in terms of social welfare.

real benefit amounts that yield SSK's reported wealth figures by lifetime earnings decile (in their table 2). Of course, SSK's data are not perfectly designed for my purposes. Most obviously, the average age of their sample is 55.7 years, several years prior to typical retirement age. Ideally, one would have data at age 62 or 65. While it is possible that the last few years prior to retirement differentially affect retiree households by income decile, a comparison of the SSK data with calculations by Love, Palumbo, and Smith (2009) for (point-in-time) income quintiles suggests that this is not likely to be a serious concern.

A different concern is that the SSK data are relatively old, focused on the 1992 HRS wave. William Gale, John Karl Scholz, and Ananth Seshadri (2009) attempt to address this concern and show that their core findings are largely unaffected by considering later cohorts (though they do not reproduce the estimates needed for this paper for later waves). Finally, recent work on the progressivity of the overall Old Age, Survivors, and Disability Insurance (OASDI) program has noted that, in the words of the CBO (2006), "the progressivity of Social Security is driven mainly by disabled-worker and auxiliary [survivor] benefits" (p. 4). While this paper focuses on the retirement portion of benefits, for which SSK's estimates are well suited, indexing reform's implications for disability benefits may be of interest as well. (Note that SSK implicitly includes disability benefits after retirement age has been reached, since disability benefits are automatically converted to retirement benefits at that point).

Table 1 shows the median (non-Social Security) net worth A_i, annualized defined-benefit pension benefit $P_{i,t}$, annual benefit level under the status quo Social Security system B_i^{SQ}, and present-value Social Security wealth, all in 2005 dollars and by household lifetime earnings decile. To be clear, all of these estimates are from SSK other than B_i^{SQ}, which is inferred from SSK's Social Security wealth estimates.

Standing out from table 1 are the small initial net wealth holdings A_i of the lowest deciles of the lifetime-earnings distribution.[15]

To estimate mortality rates by income decile, I rely on recent work by Bosworth and Burke (2014), who calculate relative mortality rates by lifetime-income quintiles and gender in the HRS for the age range 50–74. From the SSA's current period life table, I also have average mortality rates

15. Though not included in the baseline analysis, in Section 5 I show the robustness of the main results to including in the simulations a means-tested (by income) transfer payable to all individuals, modeled on the Supplemental Security Income program of the United States.

Table 1. Net Wealth and Initial Benefits
2005 dollars

Lifetime earnings decile, type i	*Initial net worth at retirement* (A_i)	*Annual defined-benefit pension payment* ($P_{i,t}$)	*Annual Social Security benefit payment* (B_i^{SQ})	*Present value of benefit payments (SS wealth)*
1 (lowest)	6,938	0	3,086	36,357
2	35,383	0	4,965	58,499
3	60,336	0	6,813	80,263
4	104,069	1,259	9,122	107,471
5	124,883	2,504	11,180	131,722
6	172,543	3,872	14,016	165,138
7	178,416	4,756	15,717	185,175
8	231,727	6,058	17,831	210,076
9	313,594	8,883	19,273	227,063
10 (highest)	545,321	10,779	23,273	281,206

Source: Author's analysis; Scholz, Seshadri, and Khitatrakun (2006).

by age and gender. Combining these data sources, I adjust the SSA's overall average mortality rates at age 65 by a vector of scalars to approximately match Bosworth and Burke's mortality patterns by income quintile. Bosworth and Burke also report relative mortality rates for the age range 75+, indicating some convergence of mortality rates across income quintiles as retirees age. To roughly match this convergence, I calculate mortality rates after age 65 so that each decile's mortality rate approaches linearly the average gender-specific mortality rate by age 119, the SSA life table's terminal age (the results change very little if I assume no such convergence in relative mortality). Table 2 shows the resulting one-year mortality rates for each decile, by gender, at 10-year increments from age 65 through age 95.

The calculated mortality rates in table 2 show the dramatic negative relationship between lifetime earnings and mortality rates, especially early in retirement. These rates roughly match existing related estimates along a number of dimensions. For example, shown at the bottom of table 2 are the SSA's official average mortality rates for each gender at each age; these rates match the calculated values for the sixth decile in all cases.

III.C. Reform Proposals

I consider three reform proposals, two informed by recent experience with chained CPI-U and the experimental CPI-E series and a third based on President Obama's proposal that is a hybrid of the first two. Figure 2 shows historical data for the December values of three price indexes, the chained

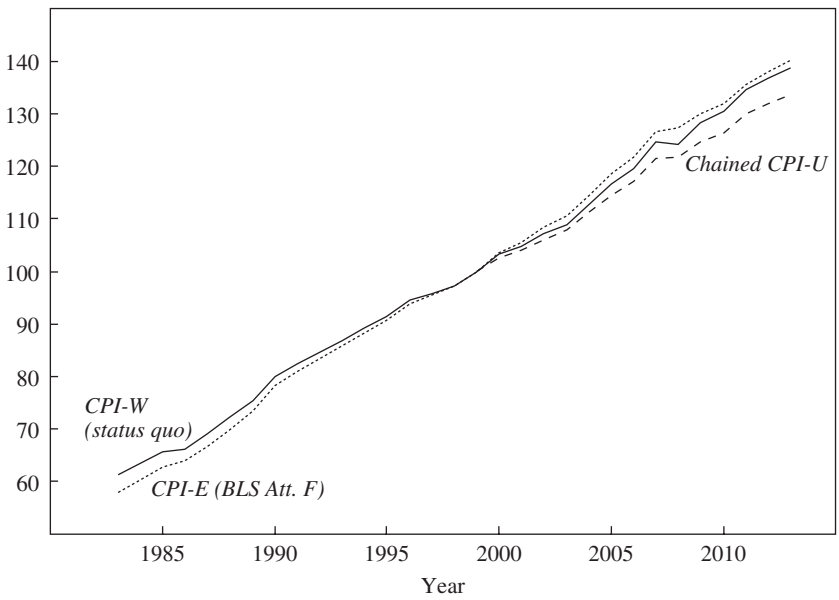
Table 2. One-Year Calculated Mortality Rates

Percent

Lifetime earnings decile, type i	*Men's age*				*Women's age*			
	65	*75*	*85*	*95*	*65*	*75*	*85*	*95*
1 (lowest)	2.2	4.9	12.3	30.4	1.5	3.6	9.6	25.4
2	2.0	4.6	11.8	29.4	1.4	3.4	9.2	24.6
3	1.9	4.3	11.2	28.4	1.2	3.0	8.4	23.0
4	1.9	4.3	11.2	28.4	1.1	2.9	8.0	22.2
5	1.9	4.3	11.2	28.4	1.1	2.8	7.8	21.8
6	1.6	3.8	10.2	26.4	1.0	2.7	7.6	21.4
7	1.5	3.6	9.6	25.4	1.0	2.7	7.6	21.4
8	1.3	3.3	9.1	24.4	1.0	2.7	7.6	21.4
9	1.2	3.0	8.5	23.5	0.9	2.3	6.8	19.8
10 (highest)	1.1	2.8	8.0	22.5	0.5	1.6	5.2	16.6
Social Security Administration average	1.6	3.8	10.2	26.4	1.0	2.7	7.6	21.4

Source: Author's analysis; Bosworth and Burke (2014).

Figure 2. Historical Data on Three Price Indexes, 1980–2015



Source: Author's analysis; see text for data sources.

CPI-U, CPI-E, and the currently used CPI-W, with each index set equal to 100 in December 1999. The chained CPI-U was first reported by the BLS in 2000, while the BLS has backcast the experimental CPI-E through 1983.

The figure makes clear that chained CPI-U has risen more slowly than CPI-W over the 2000–2013 period, although year-to-year changes are not always smaller. By the end of the 14 available years, chained CPI-U is approximately 4 percent lower than CPI-W, for an average annual gap of –0.27 percentage points.[16] Over the same period, CPI-E has been nearly identical to CPI-W. That stands in stark contrast to its more rapid growth from 1983 through 2000, when it exceeded CPI-W by an average annual rate of 0.37 percentage points.

The first reform policy—Back-Loaded Reform—is designed to mimic the CPI-E's behavior in the 1983–2000 era, having benefits grow at a faster rate than the status quo. Specifically, I set $\pi = 0.0037$ for Back-Loaded Reform, implying a steeper path of benefits and a smaller initial benefit than in the status quo.

The second reform policy—"Front-Loaded Reform"—is designed to mimic the chained CPI-U's behavior since its origination, having benefits grow at a slower rate than the status quo. Specifically, I set $\pi = -0.0027$ for Front-Loaded Reform, implying a flatter path of benefits and a larger initial benefit than in the status quo. This reform has received much attention in public debates, since the chained price index is generally viewed as addressing upward bias in the traditional CPI. It may be of interest to note that the National Commission on Fiscal Responsibility and Reform (NCFRR 2010), commonly known as the Bowles-Simpson commission, recommended a shift to chained CPI-U for Social Security benefits indexing.

The third reform policy—"Hybrid Progressive Reform"—is designed to match the proposal made by President Obama. In that proposal, the chained CPI-U would be used to index benefits, but so-called benefit enhancements would phase in for individuals when they reached ages 75 and 95, each time eventually raising their benefits by five percent of the average retiree's benefit over a 10-year phase-in period. This reform combines features of the two others, but it also includes a substantial increase in the progressivity of Social Security benefits. The source of this increase is the use of the average retiree's benefit, rather than each individual retiree's benefit, in the calculation of the benefit enhancement. Because the average benefit is approximately four times greater than the lowest decile's benefit and half

16. Thanks to Alan Viard for noting an error in my calculation of this value for π in an earlier draft of the paper.

Table 3. Parameters of Three Reform Policies

Policy	π^a	λ^b
Status quo	0.0000	1.000
Back-Loaded Reform (CPI-E)	0.0037	0.970
Front-Loaded Reform (chained CPI-U)	–0.0027	1.022
Hybrid Progressive Reform	–0.0027 + benefit enhancements for ages 76–85 and 95+	1.011

Source: Author's analysis; see text for data sources.
a. π determines the slope of the path of benefits; see text for details.
b. λ determines the starting level of the path of benefits; see the text for details.

as large as the top decile's (see table 2), the first 10-year benefit enhancement would effectively raise benefits by 20 percent for the lowest decile retiree and by 2.5 percent for the highest decile retiree.

The president's proposal thereby illustrates how the debate over indexing is closely linked to the broader debate over progressivity. Note too that this proposal's redistributive impacts make it more likely to affect labor supply during households' working lives—effects from which this paper abstracts throughout (see Liebman, Luttmer, and Seif 2009 for evidence on how labor supply responds to Social Security benefits changes).

Table 3 summarizes these proposals and shows the equilibrium value of λ that satisfies the government budget constraint when I simulate the economy's response to each policy. Technically, to obtain these values I set π for each reform policy and have the simulation guess a value for λ. All individuals maximize their utilities given these parameters and the data on benefits, net wealth, and mortality. The simulation searches for a value of λ that satisfies the government's budget constraint, as shown in equation 2.

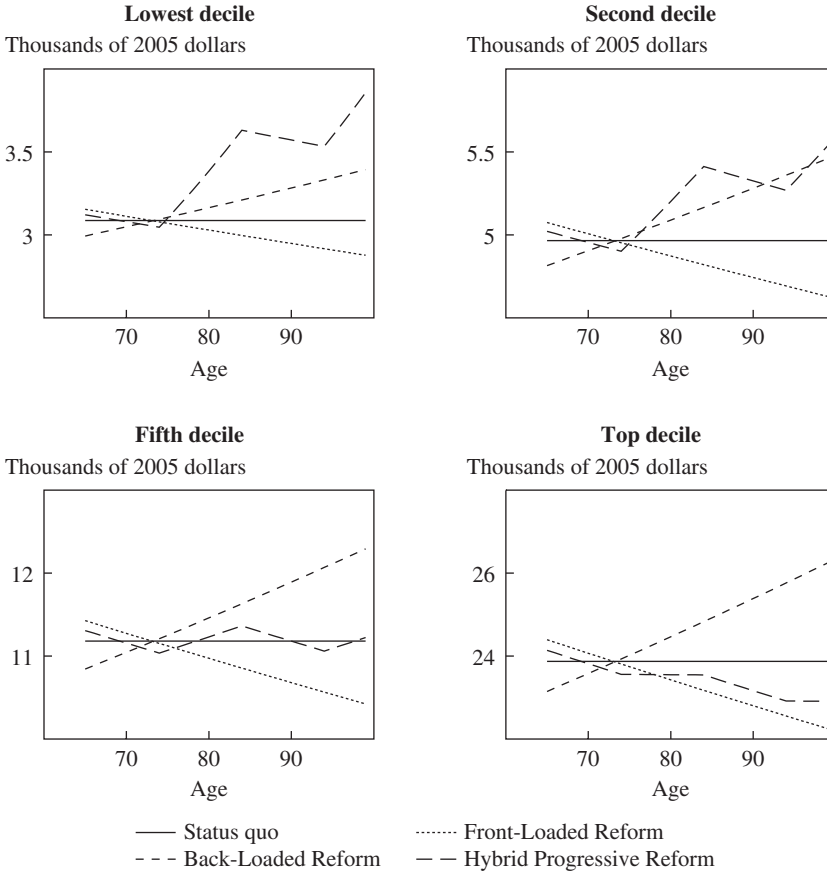
III.D. Simulated Effects of Reform

Now I turn to the effects of these reform policies.

FRONT-LOADED AND BACK-LOADED REFORMS. I begin by showing the policies' effects on real benefit payments in figure 3. The four panels of figure 3 show results for the lowest and second-lowest income deciles, the fifth income decile, and the top income decile. In each panel, I show the benefit paths under the status quo and the three reform policies at each age.

The subplots in figure 3 for the Front-Loaded Reform (dotted lines) and Back-Loaded Reform (dashed lines) closely resemble figure 1, of course. In fact, because I apply the same λ, π pair to all benefit paths in each

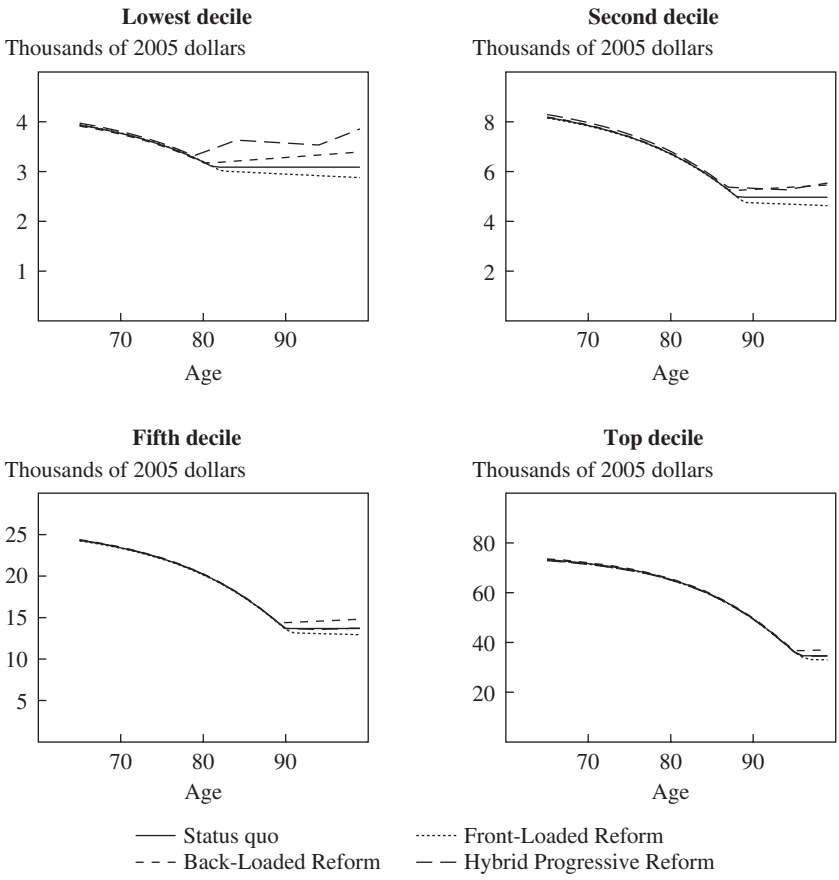
Figure 3. Annual Benefits Paths under Three Reform Proposals for Four Lifetime-Income Deciles



Source: Author's analysis; see text for data sources.

reform, the figure shows that the effects of reform are quite similar across income deciles. The effects of the Hybrid Progressive Reform (dash-dotted lines) differ dramatically across deciles, as lower-income retirees gain from its redistribution toward them. It is also apparent how the Hybrid Progressive Reform takes on a zigzag shape that combines the two other reforms, providing higher benefits at the start of retirement than the Back-Loaded Reform provides, and providing higher benefits at the end of retirement than Front-Loaded Reform does. In fact, compared to the status quo policy (solid lines), it achieves both higher initial benefits and higher final

Figure 4. Annual Consumption Paths under Three Reform Proposals for Four Lifetime-Income Deciles



Source: Author's analysis; see text for data sources.

benefits for low-income retirees, reflecting its substantial redistribution of benefits from higher to lower deciles.

The effects on consumption paths chosen by retirees in the model are much more variable across income deciles. In figure 4, I plot consumption under the status quo and these three reforms following the same structure as above:

Figure 4 shows the pattern of declining consumption until private assets are exhausted, as discussed in section 1, which obtains due to the lack of private annuitization. A similar pattern is found in Hurd and Rohwedder (2008).

To interpret figure 4, it may be helpful to focus first on a comparison of the two simplest reforms—Back-Loaded Reform and Front-Loaded Reform—postponing a consideration of the Hybrid Progressive Reform until later. For these two reforms, two prominent features stand out. First, while a household's chosen consumption paths are hardly distinguishable across benefits-indexing methods early in retirement, they sharply diverge when they exhaust their non-Social Security wealth. Remarkably, all deciles experience a substantially higher median path of consumption in these later years under the Back-Loaded Reform than they do under the Front-Loaded Reform, despite the latter's inability to generate substantial increases in consumption earlier in retirement. Second, the ages at which households exhaust their private assets and become dependent on Social Security benefits rise substantially with lifetime income.

One potentially puzzling nuance related to the first of these features is that for all but the bottom decile, consumption is slightly greater at all ages under the Back-Loaded Reform than it is under the Front-Loaded Reform. The key intuition for this result is that the Front-Loaded Reform provides less insurance against longevity risk than the Back-Loaded Reform. Therefore, households choose to consume less of their private assets in order to self-insure against longevity risk, offsetting the mechanical increase in benefits at early ages that the Front-Loaded Reform provides.

For the lowest-decile households, however, consumption is greater at early ages under the Front-loaded Reform than under the Back-Loaded Reform. Two factors explain this exception. First, these households have little wealth and high mortality rates. Thus, the effective annuitization provided by the back-loaded path enables only small increases in consumption out of their private assets early in retirement, in contrast to higher-decile retirees. A second, more subtle reason is that the front loading that comes from using a slower-growing price index is not actuarially fair. To see why, note that the Front-Loaded Reform allocates the total value of benefits through a uniform proportional adjustment to status quo benefits. Thus, it causes a redistribution of resources from low-mortality to high-mortality retirees, increasing the consumption of lower-income retirees. Note that this factor provides a second reason why consumption paths do not rise for higher-decile retirees under the Front-Loaded Reform.

As for the second prominent feature of figure 4, consistent with prior research I find that most retirees exhaust their private assets only late into retirement, while a substantial share of lower-income retirees depend on Social Security benefits throughout much of retirement. Overall, only 18 percent of individuals exhaust their non-Social Security,

non-defined-benefit assets in this simulation, also consistent with prior research. For example, Love, Palumbo, and Smith (2009) calculate what they call "annualized comprehensive wealth," which is the value of a retiree's total resources divided by his or her remaining life expectancy at any given age. In their research they find that "in (real) dollar terms, the median household's . . . real annualized wealth actually tends to rise with age over retirement" (p. 191). In my simulations, I find consistent patterns, with annualized wealth calculated this way being greater 15 years into retirement than at the start and positive until at least age 90 for retirees in the third income decile or higher. At the same time, lower income decile retirees exhaust their non-Social Security wealth much earlier. For the lowest decile, in these simulations non-Social Security wealth is nearly exhausted 15 years into retirement and is less than the level of annual benefits only 8 years in.

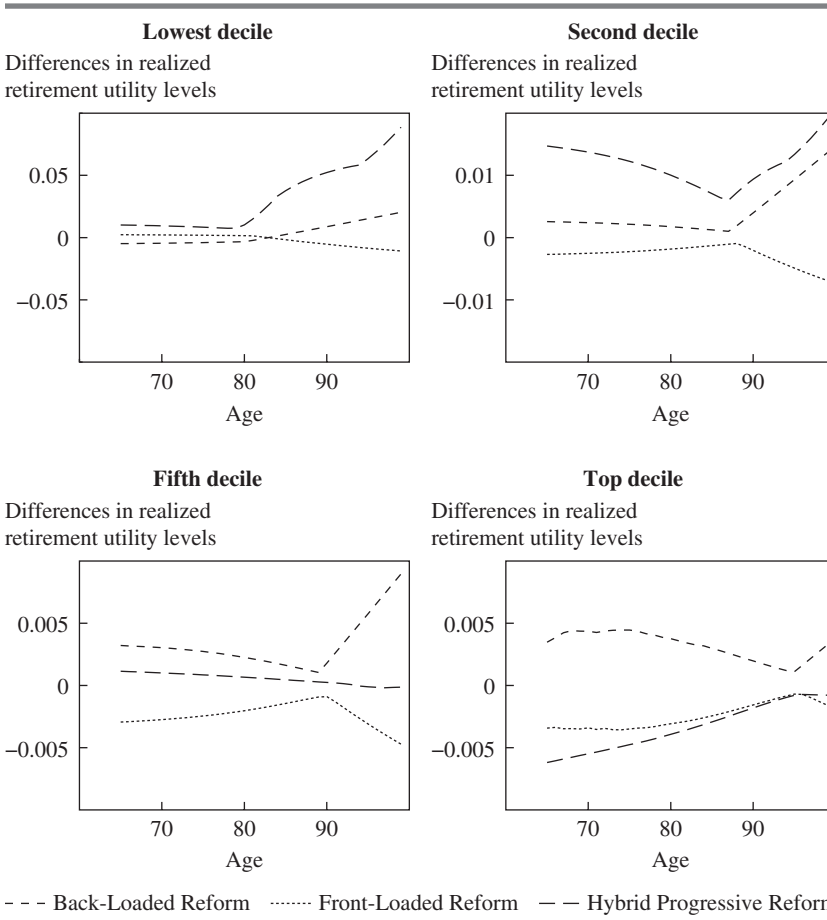
Figure 4 appears to make a strong case in favor of Back-Loaded Reform relative to Front-Loaded Reform, and that case looks all the stronger if one converts these results on consumption paths into changes to expected utility during retirement. At retirement, all deciles—even the lowest—prefer Back-Loaded Reform to the status quo and prefer the status quo to Front-Loaded Reform in expected utility terms.

The seeming dominance of Back-Loaded Reform is not airtight, however, because it generates losses for the poorest, shortest-lived retirees relative to Front-Loaded Reform or the status quo. To examine this feature of the reforms, I calculate each individual's change in "realized retirement utility": the change in total utility during retirement for an individual from decile i who lives t years under each policy. I then convert these changes, which are in units of utility, into consumption equivalents by calculating the percentage change in the total present value of consumption during retirement that, when multiplied by the marginal utility of consumption in the last year of life for a retiree, yields the given change in realized retirement utility. Figure 5 shows the results.

Figure 5 shows the extent to which the seesaw apparent in benefit paths translates into a similar shape in realized retirement utilities.

For all but the lowest decile of households, the Back-Loaded Reform generates higher realized utility than the Front-Loaded Reform no matter the age of death, but especially at later ages, when its ability to provide longevity insurance has its greatest value. The same preference holds for the poorest households who live beyond approximately age 82. In other words, most retirees sit squarely on the left end of the seesaw when it comes to these direct effects of benefits-indexing reform, preferring a steeper path of

Figure 5. Differences in Realized Retirement Utility Levels for Four Lifetime-Income Deciles[a]



Source: Author's analysis; see text for data sources.

a. Retirement utility levels are converted into consumption equivalents. Realized retirement utility for type i and age t is the total utility obtained in retirement for a household of income i who lives to the age t.

benefits with a lower starting point. However, the poorest households, who die earlier in retirement, prefer Front-Loaded Reform over Back-Loaded Reform, as shown in the top left panel of figure 5. That is, they sit on the right end of the seesaw and prefer a flatter benefit profile.

HYBRID PROGRESSIVE REFORM. I now turn to a consideration of the Hybrid Progressive Reform illustrated in figures 4 and 5. The Hybrid Progressive Reform generates very different consumption effects across deciles:

while its path lies below that of the status quo at all ages for the top decile retiree, it exceeds all other paths at all ages for the bottom decile retiree. These differences reflect both its combination of the two other reforms and its extensive redistribution of benefits, since it can achieve wide-ranging improvements for low-income retirees at the cost of a general decrease in consumption for higher-income retirees. As would be expected, these implications for consumption translate into gains (relative to the status quo) in realized retirement utility for every retiree in the bottom two (three, in fact) lifetime income deciles and losses for every retiree in the top five deciles.

In the next section, I explore how one might convert these heterogeneous results across retiree households into net welfare implications.

IV. Welfare Criteria and Net Welfare Implications of the Direct Effects of Reform

As summarized formally above in equation 3, in this paper I calculate the net welfare effects of reform by multiplying discounted changes in consumption from the status quo by two factors: the population proportion of individuals who survive to enjoy that consumption, and a non-negative welfare weight $g_{i,t}$. The weight $g_{i,t}$ measures the value society puts on a marginal increase in consumption for a household of type i at age t, relative to all other households. Because each reform has some retirees who gain and some who lose, their net welfare implications depend on how those welfare weights vary across the population of retirees.

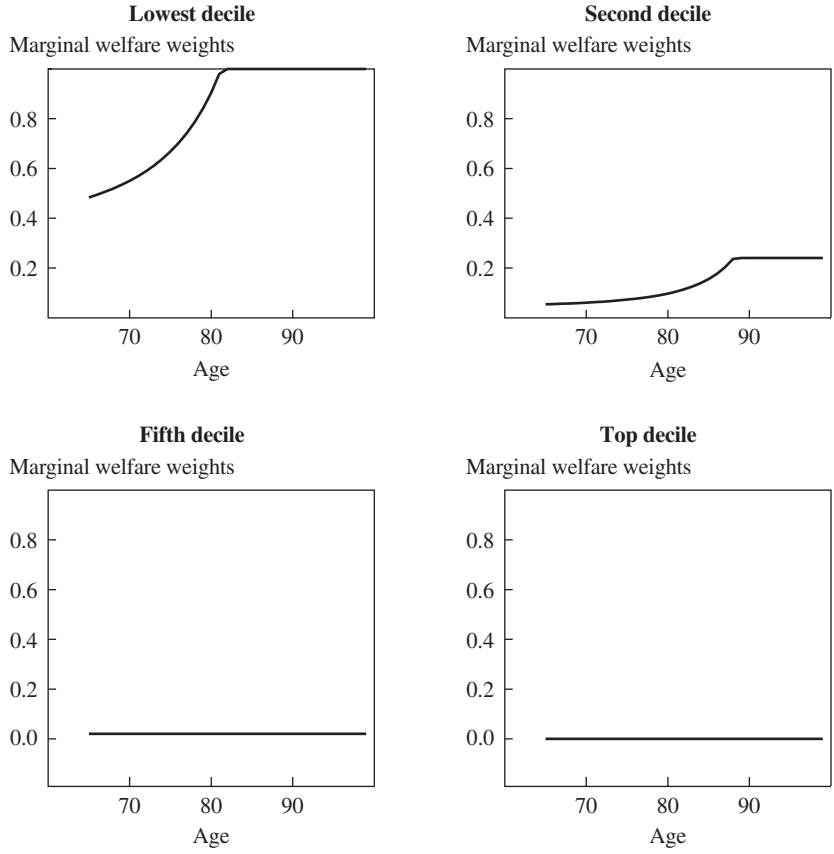
IV.A. Two Familiar Principles

The conventional approach to normative evaluation in much of applied public economics research is to rely on well-known principles with roots in political philosophy, the two most commonly used being the simple-sum utilitarian criterion and the so-called Rawlsian or maximin criterion.[17] These two criteria have especially clear implications for the welfare weight $g_{i,t}$. Indeed, it turns out that the choice between them well illustrates the policymaking challenge posed by benefits-indexing reform.

Under the simple-sum utilitarian criterion, the change in welfare dW is the sum of the experienced annual utility changes across all individuals. In the language of the general formula (equation 3 above), this option sets

17. See Weinzierl (2014) for a critique of this conventional choice in the optimal tax literature.

Figure 6. Marginal Welfare Weights for Four Lifetime-Income Deciles, Utilitarian Criterion[a]



Source: Author's analysis; see text for data sources.
a. Marginal welfare weight $g_{i,t}$ equals the marginal utility from consumption at age t for a retiree of type i.

$g_{i,t} = \acute{u}(c_{i,t})$, so that society puts greater weight on the annual consumption changes of individuals with lower consumption levels (and thus higher marginal utilities of consumption). Figure 6 shows the $g_{i,t}$ for the same four deciles as in previous figures, given the consumption levels in the simulated status quo economy from the previous section and scaled so that the maximum $g_{i,t}$ equals one.

The vertical axes in figure 6 all have the same scale, making it clear that the utilitarian criterion puts much greater weight on consumption changes for lower-income retirees than other retirees, and in particular

on consumption changes at advanced ages for those households, a period when their consumption levels fall to the level of their Social Security benefits. Note that these differences are especially large given my assumption (following SSK) that $\gamma = 3$, a value that is toward the upper end of conventional ranges for that parameter. If I use $\gamma = 1.5$, the marginal weights on the fifth decile rise to around 0.20.

The Rawlsian criterion prioritizes the well-being of the worst-off member of society. It therefore sets $g_{i,t} = 1.00$ on the consumption change of the retiree in the lowest-income decile who lives only one year in retirement—that is, the retiree with the lowest overall utility in retirement—and $g_{i,t} = 0.00$ on all other consumption changes.[18]

The net welfare implications of each benefits-indexing alternative under the Rawlsian criterion are immediately apparent from examining figure 5 for $i = 1$, $t = 1$, which shows the effects of each path on the total utility in retirement for the retiree from decile 1 who dies in the first year of retirement. From that figure, it is clear that the Rawlsian criterion would endorse the Front-Loaded Reform over the status quo and both over the Back-Loaded Reform.

The net welfare implications under the utilitarian case are not immediately clear, since that criterion puts substantial weight not only on the same worst-off retiree that drove the Rawlsian results but also on poor retirees who live long into retirement and spend down their private assets. To calculate the change in social welfare under the utilitarian criterion, I multiply the relevant $g_{i,t}$ values by discounted consumption changes adjusted for survivorship and take the sum, as in equation 3. The utilitarian criterion turns out to endorse the Back-Loaded Reform over the status quo and both Back-Loaded Reform and the status quo over the Front-Loaded Reform: in other words, exactly the opposite order as under the Rawlsian criterion.

As this result implies, the extent of back-loading most preferred under the utilitarian criterion may be substantially larger than that implied by a switch to the CPI-E. Solving for the utilitarian-optimal π and λ (that is,

18. The Rawlsian priority as modeled here is an extreme case of a social objective in which weights on individuals decrease in their lifetime utility, rather than their annual utility—see Pestieau and Ponthiere (2012) and the comments on this paper by Aleh Tsyvinski for discussions. A related pattern for MSWWs (in this paper's framework) weights consumption changes by the retiree's total utility in retirement raised to a negative exponent (e.g., as if one were taking the marginal utility of total consumption in retirement). Such weights can generate a preference for front-loading if the curvature over total retirement utility is steep enough, because the weights in that case approach Rawlsian weights. For less steep curvature, back-loading is still preferred.

for the class of reforms formalized in equation 1), I find that $\pi = 0.012$ (about three times the rate increase from the switch to the CPI-E) and $\lambda = 0.91$ maximize the total expected utility of all retirees at retirement. (I cannot use the marginal welfare weights approach in this case because the changes are too large.)

The contrast between the Rawlsian and utilitarian rankings suggests that there may exist a mixture of the two that would be consistent with the status quo policy being chosen as optimal. In fact, if I put a weight of 0.91 on the Rawlsian weights and 0.09 on the utilitarian weights, the planner prefers the status quo policy to both the Back-Loaded and Front-Loaded reforms.[19] The large implied weight on the Rawlsian weights in the status quo makes sense in light of the finding that back loading is preferred by most agents. That is, for the status quo policy to be optimal the planner must have a large weight on the worst-off retiree.

The Hybrid Progressive Reform, however, is the most preferred of these policies under both the utilitarian and Rawlsian criteria. By combining the two other reforms' positive implications for the poorest retirees, the Hybrid Progressive Reform outperforms them both. That is, the proposal's front loading in the early years benefits the worst-off retirees, increasing its appeal under the Rawlsian criterion, while its back loading through benefit enhancements brings utilitarian gains. Both of these benefits are substantially augmented by the redistribution pursued under this reform, while under both criteria the corresponding negative effects on the top half of retirees are given very little weight. I can quantify the potential gain from this reform under the utilitarian criterion by calculating the uniform proportional increase in consumption, across all retiree types and ages, that would generate the same increase in social welfare as does this reform over the status quo. That "consumption-equivalent" gain is 0.75 percent of consumption for retirees in the case of the Hybrid Progressive Reform. For comparison, the Back-Loaded Reform generates a gain of 0.12 percent of consumption for retirees, and the Front-Loaded Reform generates a slightly smaller size loss. Recall that all of these calculations abstract from a number of indirect effects of benefits-indexing reform on households, including on their labor supply and saving decisions during their working lives.

Of course, these conventional criteria might not match true social preferences. To explore this possibility, I will now take an empirical look at society's normative priorities for Social Security.

19. Thanks to Aleh Tsyvinski for suggesting this analysis.

IV.B. Evidence on Prevailing Normative Priorities for Social Security

This section presents some novel survey evidence on the American public's priorities for Social Security that I generated using Amazon's Mechanical Turk (M-Turk) interface. The way in which I elicit marginal social welfare weights $g_{i,t}$ through this survey may prove useful to other researchers interested in using a positive approach to normative questions.[20]

The survey was completed in August 2014 by 150 members of the Amazon Mechanical Turk worker population from the United States who demonstrated good past performance on prior M-Turk tasks. Respondents had up to 15 minutes to complete the survey, and they were asked to enter their M-Turk identification number as well as a completion code at the end of the survey for verification purposes. The respondents completed the survey in less than 7 minutes on average. They were paid $2.50 for answering the survey, for an average hourly rate of $23.00.

Mechanical Turk is admittedly an imperfect tool, as it does not provide a representative sample of Americans. That said, it has proven to be a popular alternative to surveys costing orders of magnitude more (and which have their own problems with representativeness), and analysis by subgroup can provide some reassurance regarding the robustness of the results to sample composition. John Horton, David Rand, and Richard Zeckhauser (2011), after studying the use of Mechanical Turk, reach this finding: "Online experiments, we show, can be just as valid—both internally and externally—as laboratory and field experiments, while often requiring far less money and time to design and conduct" (p. 399).

The survey has three parts. The first part tests whether respondents understand and can perform simple calculations related to the concepts of percentages, averages, and life expectancy. The third part of the survey asks respondents to self-report their political views and demographic traits (age, gender, education, and economic status).

I gather data on normative priorities for Social Security from the second part of the survey. Respondents are given a one-sentence (official) description of Social Security, told that policymakers must decide (among other things) how much in benefits to pay out to different retirees, and then told they will be asked a couple of questions to get their "opinions

20. A growing literature in public economic theory has considered using positive evidence on prevailing normative priorities, rather than exogenously specified normative criteria, to inform evaluations of policy. See Gaertner and Schokkaert (2012) for an overview of "empirical social choice" research. Weinzierl (2014) and Saez and Stantcheva (2014) are recent examples applied to tax policy.

on how policymakers should make this choice." They are then shown the following screen:

Please consider the following situation.

Suppose the Social Security system has raised some extra tax revenue that must be allocated among the three retirees described below. Please assume that these retirees worked equally hard during their working years and saved equally well for retirement.

Please rank these retirees in terms of who ought to receive an increase in his Social Security benefits, where #1 is the retiree you think ought to be the first to receive an increase and #3 is the retiree you think ought to be the last. (Drag the retiree descriptions to change their ranks).

- John is 65 years old. He has just retired and is expected to live to age 70. While he was working, his income was in the bottom 10% of incomes, and he currently has about $10,000 per year to spend.
- Robert is 75 years old. He retired at age 65, when he was expected to live to age 83. While he was working, his income was in the middle 10% of incomes, and he currently has about $25,000 per year to spend.
- William is 90 years old. He retired at age 65, when he was expected to live to age 81. While he was working, his income was in the bottom 10% of incomes, and he currently has about $9,000 per year to spend.

The three retirees in this first question represent three important points in the joint age-income distribution.[21] In particular, John represents a very-low-income individual with a short life expectancy, the point given particular priority by the Rawlsian criterion. William is also very low income but has lived a long life, giving him a greater overall utility level than John but leaving him with a smaller current (according to the survey) level of consumption. Thus a utilitarian would allocate more to William, while a Rawlsian would allocate more to John. Finally, Robert is a middle-income individual approaching his expected lifespan. He is much better off than either of the other retirees and provides a simple way for us to gauge how quickly marginal welfare weights decline with well-being.[22] This first question is largely intended to get respondents to engage with the descriptions of these retirees. Nevertheless, the responses may be of interest. William is rated first by 62 percent the respondents, John by 29 percent, and

21. These names were the most popular names, according to the Social Security Administration's names database, for boys born in 1949 and 1924.

22. The consumption levels indicated in the survey implicitly include other transfers for John and William, such as SSI and SNAP. See section V for a discussion of how these programs relate to this paper's analysis of Social Security.

Robert by 9 percent. The preference for William directly casts some doubt on the possibility that a Rawlsian criterion will emerge from the survey evidence.

The key questions for this paper's purposes come next, when respondents are shown a series of screens starting with one like the following, tailored according to which retiree the respondent ranked last in the previous question. The screen reproduced below is shown to a respondent who had ranked Robert last in that question.

On the next several pages, we'll ask for your opinion on some specific options for changing these retirees' benefits.

When making your choices, please imagine that you are a policymaker trying to choose what is best. Ignore any effects these options might have on the rest of the economy, and focus on the effect each option has on the retirees.

For your reference, we'll copy the descriptions of the retirees on each page.

- John is 65 years old. He has just retired and is expected to live to age 70. While he was working, his income was in the bottom 10% of incomes, and he currently has about $10,000 per year to spend.
- Robert is 75 years old. He retired at age 65, when he was expected to live to age 83. When he was working, his income was in the middle 10% of incomes, and he currently has about $25,000 per year to spend.
- William is 90 years old. He retired at age 65, when he was expected to live to age 81. While he was working, his income was in the bottom 10% of incomes, and he currently has about $9,000 per year to spend.

Which of the following would you prefer?

- ○ Increasing Robert's benefit by $100
- ○ Increasing John's benefit by $100

If the respondent chooses Robert over John in this question, he or she is reminded (by the computer) that Robert was ranked last in the earlier question, and he or she is asked to make the choices consistent. Then, the following choice appears:

Which of the following would you prefer?

- ○ Increasing Robert's benefit by $100
- ○ Increasing John's benefit by $75

If the respondent chooses John over Robert, he or she then faces the same choice but with the increase for John set at $50 and then at $25. After that, or whenever the respondent chooses Robert over John, he or she then

faces a similar set of choices between benefits increases for Robert and William.

This series of questions is designed to allow the direct inference of marginal welfare weights.[23] To see how, suppose a respondent ranks Robert last and (implicitly) assigns marginal value g_{Robert} to Robert's consumption. In that case, a $100 increase in Robert's benefits provides a benefit to the respondent (acting as a policymaker) of $100g_{Robert}$. The respondent is then asked to choose between this gain and alternative gains. Suppose the respondent chooses the $50 increase for William (but not the $25 increase) over the $100 increase for Robert. Then, one can infer that $100g_{Robert} > 25g_{William}$ and $100g_{Robert} < 50g_{\text{William}}$, implying that $g_{Robert}/g_{William} \in [0.25, 0.5]$. Similarly, one can calculate a range for $g_{Robert}/g_{\text{John}}$ for each respondent, indicating the profile of relative welfare weights across these retirees.

Components of these questions are designed to counteract some potential confounding influences on the respondents. I ask respondents to "ignore any effects these options might have on the rest of the economy, and focus on the effect each option has on the corresponding retiree." This request is intended to minimize the extent to which respondents consider the efficiency costs of raising different amounts of extra revenue for the benefits increases. I also ask them to "imagine that you are a policymaker" in the hopes that it will cause the respondents to take a considered, objective perspective. Inconsistent answers across the ranking question and the series of choices cause error messages to appear, preventing the respondent from making errors in interpreting the questions. Finally, the wide range of potential relative valuations implied by the choices (from 1.3 to 4.0 in each case) is intended to reduce concerns that respondents would default to equality and thereby imply smaller differences between g weights than is accurate, as might be natural in other designs (such as splitting an amount between the retirees).

A number of potential risks remain with survey evidence of this kind. One risk is that respondents may not be accustomed to thinking about these policy choices in terms of indifference points, which seem natural to most economists but which reverse the intuitive idea that the respondent would like to grant his or her preferred retirees greater increases, not smaller ones, than his or her least preferred retiree. Of course, more general concerns about how the questions are framed and whether the survey

23. An earlier version of the survey used sliders to elicit the same information. Though preferable in many ways, the slider interface appeared to confuse respondents, who often implicitly assigned lower weights to retirees they preferred.

primes respondents toward any particular outcome also apply to this specific survey.

The results of the survey for the relative weight on Robert versus John and William are consistent with the weights implied by the utilitarian criterion. The median choices across all respondents imply a range of values for both $g_{Robert}/g_{William}$ and g_{Robert}/g_{John} of [0.00, 0.25]. That is, these median responses indicate a very small value for the welfare weight on Robert relative to both John and William, consistent with figure 6 that shows a negligible weight on Robert under the utilitarian criterion. Of course, it is possible that the true relative weight put on Robert lies closer to 0.25 than to zero (which the survey cannot pin down). It is also possible that respondents systematically took the mental shortcut provided by choosing that John or William receive an increase—no matter how small—rather than Robert, which would bias the estimate of the relative value of g_{Robert} toward zero. Of the 116 respondents who ranked Robert last in the first question, approximately three-quarters (85) chose the benefit increase for both John and William in all cases.

At the same time, the results of the survey for the relative weights on John and William are not consistent with those implied by either the utilitarian criterion or the Rawlsian criterion on their own. Specifically, the median choice across all respondents implies that $g_{John}/g_{William} = 1.00$, so that respondents put similar value on benefits increases for John and William, contrary to both the utilitarian preference for William (which figure 6 suggests would approximately set $g_{John}/g_{William} = 0.5$) and the Rawlsian extreme preference for John. In fact, the mean choice among those who ranked Robert last, which is in general too sensitive to outliers to be a useful measure of preferences in this survey, implies that $g_{John}/g_{William} = 1.05$ with a standard error of 0.05 (the means are very large for those who ranked William last and very small for those who ranked John last).

Both of these sets of results hold across virtually all subgroups. They hold for respondents ages 18–25, ages 26–40, ages 41–64, and ages 65+; for both male and female respondents; and for respondents who place themselves on the political left, middle, and right. The only exceptions are across race and household income, where for (the small number of) black and high-income respondents, the relative weight on Robert is in the range [0.25, 0.5].

One possibility suggested by these results is that respondents' moral reasoning reflects a mixture of these two standard criteria. Such a mix can easily generate $g_{i,t}$ values for $t = 1$ and $t = 35$ for the lowest-income decile

$i = 1$ that are very similar, in keeping with the survey evidence on John and William. The same mix yields extremely small values for the welfare weight on the "average" retiree (that is, where $i = 5$ and $t = 10$), in keeping with the survey evidence on Robert.

When I apply these weights to the reform options, the rankings and consumption-equivalent welfare gains and losses are the same for all reform proposals as under the utilitarian criterion. The costs of the Back-Loaded Reform for the worst-off retirees are not large enough to offset the gains that reform generates for the poor retirees who outlive their private assets, so back-loaded benefits, such as under a switch to the CPI-E, are preferred to front-loaded benefits, such as under a switch to the chained CPI-U. Under this criterion, the Hybrid Progressive Reform, by combining the Back-Loaded Reform's appeal to long-lived poor retirees with the Front-Loaded Reform's appeal to the short-lived poor retirees, dominates the policy ranking, reflecting survey respondents' low concern for consumption decreases among better-off retirees.

V. Extensions to the Baseline Analysis

In this section, I extend the analysis above along a number of dimensions. Although each extension somewhat modifies the baseline results, the basic seesaw metaphor continues to apply, as do the trade-off between the effects on the vast majority of retiree households and those on the worst-off retirees and the likely net welfare impacts of the direct effects of benefits-indexing reform. To simplify the discussion, I focus on the effects of these extensions on the Back-Loaded and Front-Loaded reform policies.

V.A. Myopic Households

The fully rational, foresighted utility-maximizing household modeled above may not represent all, or even most, retirees' consumption and saving behavior. In particular, though the evidence reviewed in section I suggests that the problem of myopia does not apply to most retirees outside of the lowest income deciles, it may be worth knowing the effects of assuming that some share of retirees have difficulty delaying consumption early in retirement. Feldstein (1985, 1987) made clear the importance of considering myopic households in determining the optimal path of benefits.

Table 4. Impatience by Income Decile

	Bottom	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*Top*
$\hat{\beta}^a$	0.90	0.91	0.92	0.93	0.94	0.95	0.96	0.96	0.96	0.96
β^b	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96

Source: Author's analysis; see text for data sources.
a. $\hat{\beta}$ is the intertemporal discount factor used by impatient retirees.
b. β is the discount factor used to calculate social welfare.

To gauge the effects of this myopia, I consider a model in which retirees from the bottom through sixth income deciles choose consumption at each age using a lower discount factor β than the true discount factor β upon which their utility depends, as in table 4:

That is, individuals use β to choose their consumption paths, but the utilitarian evaluation of their overall retirement utility uses β (as it is assumed the individuals would if they were able to adopt a disinterested perspective). Of course, this is only a crude version of this extension to the baseline model; a more sophisticated model would have heterogeneity in impatience within deciles.

The results are similar to the baseline results, but more extreme. That is, the gains from Back-Loaded Reform are larger for the majority of households that value its insurance against longevity risk—a feature even more beneficial in a setting where households have difficulty saving. For example, households in the second decile ($i = 2$) that survive 30 years into retirement see a 30 percent larger gain (in utility terms) from Back-Loaded Reform in this setting than in the baseline. At the same time, the gains from front loading are even higher for those retirees with short ex-post lives and few initial assets. For example, the shortest-lived household in the bottom decile has more than twice the gain from front loading in this setting than from the baseline. Moreover, the shorter-lived half of households in the second income decile now gain from Front-Loaded Reform (whereas they lost in the baseline case), since their impatience causes them to benefit more from the higher initial benefits and their limited assets make the appeal of back loading small. Their (impatient) consumption of an even higher share of the front-loaded benefits means that, when they (ex post) do not survive later into retirement, their realized utility during retirement was even higher than in the patient case.

On balance, under the utilitarian criterion the increase in the gains to the majority of households outweighs the increase in the losses to a few households, such that the net welfare impacts of the direct effects on retirees are

more positive for back-loading and more negative for front-loading in this model than in the baseline case. Similarly, the difference between the two policies under the Rawlsian criterion, which ranks the front-loading policy ahead of the back-loading policy, also grows.

V.B. Budget Non-Neutrality

Thus far I have imposed budget neutrality to disentangle the effects of changing the shape of the time-path of benefits from the effects of changing the expected present-value of those benefits. However, much of the energy in the policy debate over benefits indexing is due to the likelihood that choosing a more slowly growing price index, such as the chained CPI-U, would generate savings for the Social Security program.

The approach taken above could be readily modified to include a requirement that reform lower the expected present-value cost of benefits. To illustrate this, I reduce all status quo benefits by 10 percent and impose the same restriction on reform policies as before, namely that they have the same expected present-value total cost of benefits. This variation changes the baseline results very little, the same households lining up on either end of the seesaw as in the baseline case and the same net welfare implications obtaining. The intuition for these results is that the relative effects of the reforms are largely unaffected by the shift in their total value. Once all are adjusted to provide 10 percent smaller total benefits, the Back-Loaded Reform continues to provide better longevity risk protection than the modified status quo or Front-Loaded Reform, while the poorest, shortest-lived retirees continue to prefer the Front-Loaded Reform, which still provides greater benefits early on. These results support the argument that the level of benefits and the shape of benefits may be analyzed separately.

V.C. Non-Constant Status Quo Benefits

I have also assumed, thus far, that status quo benefits are constant in real terms. In reality, there is considerable debate and uncertainty over whether they are increasing or decreasing in real terms. Goda, Shoven, and Slavov (2011b) argue that current benefits indexing, and even the faster-growing CPI-E, fail to provide enough protection against the rising costs of medical expenditure among retirees both as they age and over time. Specifically, they calculate the real Social Security benefit net of medical expenses and show that it grew more slowly from 1983 to 2007 than did a price index of nonmedical goods and services (so the real nonmedical purchasing power of Social Security beneficiaries declined). On the other hand, fixed-basket

price indexes such as the CPI-W are susceptible to the well-known problem that they overestimate the inflation faced by individuals due to quality changes and substitution away from expensive goods and services (see Boskin and others 1996).

To test the sensitivity of my baseline results to this assumption, I consider two alternatives.

First, to study the possibility that the CPI-W underestimates the inflation faced by retirees, I assume that the CPI-E is, in fact, the correct price index for retirees. This means that Back-Loaded Reform now has $\pi = 0.00$, and I set its $\lambda = 1.00$ to impose budget neutrality as in the baseline case. The status quo now has $\pi = -0.0037$ and $\lambda = 1.031$, while Front-Loaded Reform has $\pi = -0.0064$ and $\lambda = 1.054$. Visually, the benefit paths are as illustrated in figure 7.

Although the benefits paths in figure 7 look quite different from those in figure 3, the relative effects of reform on households are remarkably similar in this variation on the baseline analysis. Essentially the same households benefit from Back-Loaded Reform and from Front-Loaded Reform, and to very similar degrees. The results on the net welfare implications of reform are very similar as well under both the utilitarian and the Rawlsian criterion.

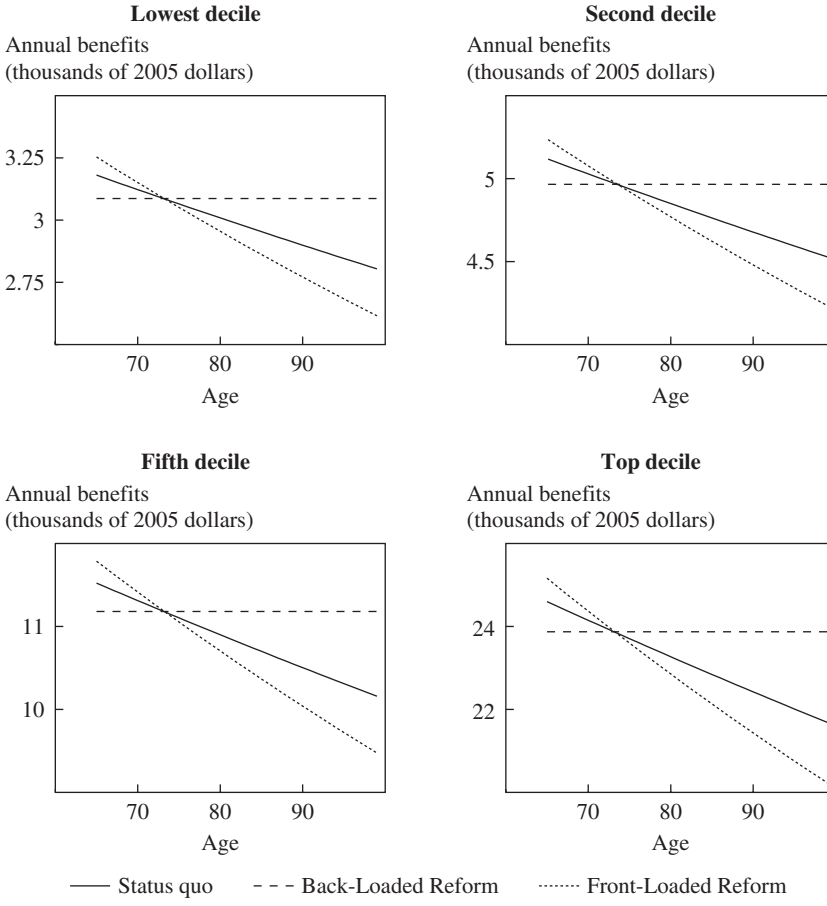
Second, to study the possibility that the CPI-W overestimates the inflation faced by retirees, I assume that the chained CPI-U is the correct price index for retirees. This means that Front-Loaded Reform now has $\pi = 0.00$, and I set its $\lambda = 1.00$; the status quo now has $\pi = +0.0027$ and $\lambda = 0.978$; and Back-Loaded Reform has $\pi = +0.0064$ and $\lambda = 0.947$. The baseline results are robust to this variation, as well.

V.D. Bequest Motive

The retirees in the baseline model have no reason to retain wealth other than longevity risk. In much of the existing literature explaining retiree wealth dynamics, a bequest motive is used as an ingredient to explain the retention of substantial assets late into retirement. As noted earlier, the simulations in this paper generate paths for what Love, Palumbo, and Smith (2009) call "annualized comprehensive wealth," which fit well with what appears in the data. Nevertheless, it may be valuable to understand the robustness of my results to the existence of a bequest motive, given its prominence in previous, more sophisticated simulations of retiree behavior.

To test this, I have retirees value any assets left at death as if those assets were consumed by their heirs in the next period, multiplied by a factor

Figure 7. Annual Benefits under Status Quo and Three Reform Paths for Four Lifetime-Income Deciles[a]



Source: Author's analysis; see text for data sources.
a. In this analysis, the Back-Loaded Reform provides a constant real benefit.

scaling the strength of their bequest motive, using the same utility function specification (equation 4) as was used for the retiree while he or she was alive. That parameter, κ, is such that the top deciles have a relatively strong bequest motive, as illustrated in table 5.

As would be expected, the resulting simulation has no retirees die with zero assets (whereas in the baseline simulation 19 percent of retirees die with zero assets. This likely overstates the degree of bequeathing done at

Table 5. Bequest Motive by Income Decile

	Bottom	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*Top*
κ^a	1	1	1	1	1	10	20	30	40	50

Source: Author's analysis; see text for data sources.
a. κ measures the strength of the bequest motive, as discussed in the text.

the bottom of the income distribution and understates it at the top. Nevertheless, the results are informative in that the changes from the baseline simulation are quite minor outside of the pattern of asset holdings. That is, the top decile retirees now die with assets equal to about three times their consumption late in retirement. The sets of households that gain and lose from each reform are largely the same as in the baseline analysis, as is the ranking of policies under either the utilitarian or Rawlsian criterion. One minor but interesting difference from the baseline is that high-decile retirees raise their consumption less later in life under the Back-Loaded Reform than in the baseline case. Intuitively, with a bequest motive these retirees save more of their private assets until death.

Of course, the nonsurviving retiree leaves assets unspent in the front-loaded policy, and if those assets were reclaimed by the government, the difference between the policies would diminish. In reality, the U.S. government raises very little revenue from the taxation of bequests, and none from households for which Social Security benefits materially change their accumulation of assets, so I assume that the direct fiscal costs of benefits are not offset by any posthumous taxation. Similarly, I do not consider the value inheritors place on bequests—a topic analyzed in Feldstein (1990)—to retain this paper's focus on the direct effects of reform on retirees.

V.E. Additional Transfers

Throughout the analysis I abstract from additional transfers made to poor retirees. In reality, the very poor elderly receive support from the Supplementary Security Income (SSI) program as well as more targeted programs such as in-kind food stamps through the SNAP program. The baseline SSI transfer was approximately $10,000 per year for an elderly couple and $7,000 for an individual in 2005, according to the Social Security Administration. SSI benefits are displaced dollar-for-dollar by almost any income source, including Social Security benefits.

While including these transfers in the simulations above is technically straightforward, there are conceptual complications. In particular, in

reality these transfers are indexed for inflation just as are Social Security benefits. If one includes these transfers without adjusting their indexing approach—something the president's 2014 budget proposal suggested—changes to the Social Security benefits of the lowest-decile retirees are entirely canceled out by changes to their SSI benefits (although the government would save some money that could be allocated to other retirees). This mechanically neutralizes the benefits of Front-Loaded Reform. In fact, in simulations of the baseline model modified to include a guaranteed minimum benefit of $9,000 that falls dollar-for-dollar with Social Security benefit increases, the lowest-decile households are unaffected by any indexing reform, most other households lose from the Front-Loaded Reform, and all other households gain from Back-Loaded Reform. That is, the Back-Loaded Reform can produce a Pareto improvement in this case relative to the status quo. Another, simpler scenario is that such transfers would also be adjusted in any reform to benefits indexing, so that the net effect on beneficiaries of a reform to Social Security benefits indexing may be only partly offset, not offset at all, or even magnified. Because of this ambiguity, as well as the likelihood that including such transfer programs in the analysis would strengthen the results of the baseline, I chose to omit them from the main paper. Of course, a more comprehensive analysis that included a range of potential changes to these transfer programs would be valuable.

Finally, note that the omission of these transfers causes the marginal utility values of consumption for individuals in the lowest-income decile to be larger than if these transfers had been included. This factor will cause the baseline analysis to overestimate the appeal of Front-Loaded Reform and underestimate the appeal of Back-Loaded Reform.

VI. Discussion and Conclusion

For most purposes, the choice of a price index for Social Security benefits may seem to have small stakes. One exception, however, is its implications for retirees who rely on Social Security benefits to fund their consumption, either because their own resources are limited or because they outlive their expected life span. For these retirees, half a percentage point faster growth in benefits—approximately the difference between two of the most prominent proposals for indexing reform—turns into a 20 percent increase in benefits if they outlive their private savings. On the other hand, assuming budget-neutral reform, it also could mean benefits that are 7 percent lower at the start of retirement, when they are sure to be alive to receive them.

In this paper, I have outlined a flexible and relatively simple formal structure for modeling this trade-off in the direct effects of benefits-indexing reform on a population of heterogeneous retiree households. I brought to that model evidence from recent empirical work on Social Security, quantifying the effects of three prominent policy proposals. I gathered some new evidence on the priorities Americans appear to have for Social Security benefits, using a methodology that may prove useful more broadly. Finally, using that evidence, as well as conventional normative criteria, I have provided suggestive estimates of those proposals' effects in terms of social welfare.

The results of this analysis suggest that reform to a back-loaded benefits-indexing approach, such as the CPI-E, has substantial appeal, at least in its direct effects on retirees. Note that this is the opposite proposal to the one that has generated the most enthusiasm in Washington: namely, a switch to the slower-growing chained CPI-U. A back-loaded approach's ability to concentrate resources at later ages, when retirees face longevity risk and have exhausted their own resources, makes it the preferred approach for most retirees. While a normative criterion that concentrates priority on the worst-off retirees would therefore endorse a front-loaded reform, the standard utilitarian criterion and the criterion implied by the survey evidence in this paper prefer to back-load the path of benefits.

Political considerations make the case for back-loaded benefits-indexing reform extremely difficult, however. Such a reform would require either a reduction in initial benefits to retain budget neutrality or an increase in total spending on benefits to retain initial benefit levels. Both requirements are likely to be deal-killers in Washington.

In this context, the appeal of President Obama's 2014 budget proposal for a benefits-indexing reform that combines a shift to the chained CPI-U with benefit enhancements at advanced ages becomes clear. Such a proposal can capture the best parts of both of the simpler reforms: protecting both the poorest, shortest-lived retirees who would prefer front-loading and the large majority of retirees, especially those who live to advanced ages, who prefer back loading. It is important to note that the president's specific proposal combined this hybrid of front loading and back loading with an increase in progressivity, which might be achieved through other means, as the benefit enhancements at advanced ages were to be uniform across the lifetime-income distribution. In the simulations in this paper, it causes the top five income deciles of retirees to prefer the status quo to this reform, and the potential disincentive effects from

which this paper abstracts may therefore reduce this proposal's appeal. Nevertheless, if those disincentive effects are limited and the normative preferences of Americans resemble those of either the conventional utilitarian criterion or those implied by the survey results in this paper, in its direct effects on retirees the Hybrid Progressive Reform is likely to generate a sizable net welfare gain.

As this last result and the rest of the analysis in this paper have demonstrated, benefits-indexing reform is more than just a fiscal issue; its distributional implications and its possible role as a vehicle for redistribution make it a flexible and potentially powerful policy tool. That said, it is important to reiterate that this paper uses a simplified model that abstracts from a number of effects of shifting the time-path of benefits on household behavior and the general economic environment, as well as from complexities of the Social Security system and retiree household structure. My hope is that it puts that simplicity to good use, clarifying one aspect of the trade-offs involved in choosing a method of benefits indexing, and that further analyses will refine our understanding of the lessons learned here.

ACKNOWLEDGMENTS I wish to thanks the editors as well as my discussants, Martin Feldstein and Aleh Tsyvinski, and the many conference participants at the fall 2014 Brookings economic meeting, for helpful comments and discussions. Thanks also to Darren A. Rippy at the Bureau of Labor Statistics for sharing the data on CPI-E. I have no relevant material or financial interests to declare regarding the content of this paper.

References

Boskin, Michael J., E. Dulberger, R. Gordon, Z. Griliches, and D. Jorgenson. 1996. "Toward a More Accurate Measure of the Cost of Living." Final Report to the Senate Finance Committee, Washington, December 4.

Bosworth, Barry, and Kathleen Burke. 2014. "Differential Mortality and Retirement Benefits in the *Health and Retirement Study.*" Working Paper no. 2014-4. Center for Retirement Research, Boston College. http://crr.bc.edu/wp-content/uploads/2014/04/wp_2014-4.pdf

Brown, Jeffrey, Olivia S. Mitchell, James M. Poterba, and Mark J. Warshawsky. 2001. *The Role of Annuity Markets in Financing Retirement.* Cambridge, Mass.: MIT Press.

Burdick, Clark, and Lynn Fisher. 2007. "Social Security Cost-of-Living Adjustments and the Consumer Price Index." *Social Security Bulletin* 67, no. 3.

Butrica, Barbara A., Joshua H. Goldwyn, and Richard W. Johnson. 2005. "Understanding Expenditure Patterns in Retirement." Working Paper no. 2005-03. Center for Retirement Research, Boston College, Boston, Mass.

Committee for a Responsible Federal Budget. 2013. "The President's Proposed Chained CPI Protections." CRFP blog, August 12. http://crfb.org/blogs/presidents-proposed-chained-cpi-protections

Congressional Budget Office. 2006. "Is Social Security Progressive?" Economic and Budget Issue Brief. Washington, December 15.

Duggan, James E., Robert Gillingham, and John Greenlees. 2007. "Mortality and Lifetime Income: Evidence from U.S. Social Security Records." Working Paper no. 07/15. Washington: International Monetary Fund.

Englehardt, Gary V., and Jonathan Gruber. 2004. "Social Security and the Evolution of Elderly Poverty." Working Paper no. 10466. Cambridge, Mass.: National Bureau of Economic Research.

Feldstein, Martin. 1976. "On the Theory of Tax Reform." *Journal of Public Economics* 6: 77–104.

Feldstein, Martin. 1985. "The Optimal Level of Social Security Benefits." *Quarterly Journal of Economics* 50, no. 2.

Feldstein, Martin. 1987. "Should Social Security Benefits Increase with Age?" Working Paper no. 2200. Cambridge, Mass.: National Bureau of Economic Research.

Feldstein, Martin. 1990. "Imperfect Annuity Markets, Unintended Bequests, and the Optimal Age Structure of Social Security Benefits." *Journal of Public Economics* 41, no. 4: 31–43.

Gaertner, Wulf, and Erik Schokkaert. 2012. *Empirical Social Choice: Questionnaire-Experimental Studies on Distributive Justice.* Cambridge University Press.

Gale, William G., John Karl Scholz, and Ananth Seshadri. 2009. "Are All Americans Saving 'Optimally' for Retirement?" Preliminary Working Paper. http://www.ssc.wisc.edu/~scholz/Research/Are_All_Americans_v6.pdf

Goda, Gopi Shah, John B. Shoven, and Sita Nataraj Slavov. 2011a. "Differential Mortality by Income and Social Security Progressivity." In *Explorations in the Economics of Aging,* edited by David A. Wise. University of Chicago Press.

———. 2011b. "How Well Are Social Security Recipients Protected from Inflation?" *National Tax Journal* 64, no. 2: 429–50.

Guesnerie, R. 1977. "On the Direction of Tax Reform." *Journal of Public Economics* 7: 179–202.

Horton, John J., David G. Rand, and Richard J. Zeckhauser. 2011. "The Online Laboratory: Conducting Experiments in a Real Labor Market." *Experimental Economics* 14: 399–425. DOI 10.1007/s10683-011-9273-9.

Hurd, Michael D., and Susann Rohwedder. 2008. "The Adequacy of Economic Resources in Retirement." Working Paper no. 2008-184. University of Michigan Retirement Research Center, Ann Arbor, Mich.

Krugman, Paul. 2012. "Life, Death, and Deficits." Op-ed, *New York Times,* November 15.

Liebman, Jeffrey B., Erzo F.P. Luttmer, and David G. Seif. 2009. "Labor Supply Responses to Marginal Social Security Benefits: Evidence from Discontinuities." *Journal of Public Economics* 93: 1208–23.

Love, David A., Michael G. Palumbo, and Paul A. Smith. 2009. "The Trajectory of Wealth in Retirement." *Journal of Public Economics* 93, nos. 1–2: 191–208.

National Commission on Fiscal Responsibility and Reform (NCFRR). 2010. "The Moment of Truth: Report of National Commission on Fiscal Responsibility and Reform." Washington. http://www.fiscalcommission.gov/sites/fiscalcommission.gov/files/documents/TheMomentofTruth12_1_2010.pdf

Olsen, Anya. 2008. "Distributional Effects of Reducing the Cost-of-Living Adjustments." Policy Brief no. 2008-03. Washington: Social Security Administration.

Penner, Rudolph G., and Karen E. Smith. 2010. "How Quickly Do Older Adults Spend Their Wealth?" Washington: Urban Institute.

Pestieau, Pierre, and Gregory Ponthiere. 2012. "The Public Economics of Increasing Longevity." *Hacienda Pública Española/Revista de Economía Pública* 200: 35–68.

Saez, Emmanuel, and Stefanie Stantcheva. 2014. "Generalized Social Marginal Welfare Weights for Optimal Tax Theory." Working Paper no. 18835. Cambridge, Mass.: National Bureau of Economic Research.

Scholz, John Karl, Ananth Seshadri, and Surachai Khitatrakun. 2006. "Are Americans Saving 'Optimally' for Retirement?" *Journal of Political Economy* 114, no. 4.

Shiller, Robert J. 2003. *The New Financial Order: Risk in the 21st Century.* Princeton University Press.

Smith, Karen E., and Melissa M. Favreault. 2013. "A Primer on Modeling Income in the Near Term, Version 7 (MINT7)." Washington: Urban Institute.

Smith, Karen, Mauricio Soto, and Rudolph G. Penner. 2009. "How Seniors Change Their Asset Holdings during Retirement." Working Paper no. 2009-31. Center for Retirement Research, Boston College, Boston, Mass.

Waldron, Hilary. 2007. "Trends in Mortality Differentials and Life Expectancy for Male Social Security–Covered Workers, by Socioeconomic Status." *Social Security Bulletin* 67, no. 3.

Weinzierl, Matthew. 2014. "The Promise of Positive Optimal Taxation." *Journal of Public Economics* 118: 128–142.

Comments and Discussion

COMMENT BY

MARTIN FELDSTEIN Matthew Weinzierl has produced a very interesting paper about the indexing of Social Security retiree benefits. He has done so by transforming the choice among price indexes into the more important question of whether retirees' real benefits should increase or decrease during the individual's retirement years.

He reminds us that the Obama administration proposed in its 2014 budget to substitute a more slowly increasing price index (the C-CPI-U) for the current CPI-W to be used in adjusting Social Security benefits during each individual's retirement years. That substitution would have caused benefits to rise more slowly during each individual's retirement years. Although this was seen as a way of reducing the present value of the benefits of each retiree generation, the substitution of a "slow index" for the traditional CPI could alternatively have been done in a revenue-neutral way by raising the benefit level at the start of retirement. Such a revenue-neutral shift to a "slow index" would benefit those with a low life expectancy. Conversely, the substitution of a "fast index" would benefit those with a higher life expectancy.

The administration's proposal was not accepted by Congress. An important reason was that it would also have applied the slow index to the bracket points of the income tax, with the result that individuals' tax liabilities would rise more rapidly. Although the change in Social Security indexing would affect each individual retiree with no cumulative effect from one age cohort to the next, the change in the tax brackets would be cumulative from one generation to the next. For example, the nominal dollar amount at which the taxpayer shifts from the 25 percent marginal tax rate to the next higher rate would be lower and lower over time, implying a higher and higher real tax at each income level. The

increase in tax revenue over the long term would therefore be much greater than the decline in Social Security outlays. It is not surprising, therefore, that the opponents of a sustained increase in tax revenue—as well as the defenders of the Social Security status quo—opposed the administration proposal.

But quite apart from the choice of the price index, the question of whether the real level of benefits should rise during retirement is a very interesting one. More specifically, for any given real present value of benefits, a case can be made that annual Social Security benefits should rise as a retiree ages. Here is why. In the absence of a market in actuarially fair annuities, individuals can save for their early retirement years but cannot be confident about their ability to maintain their standard of living if they live for an unusually long time. So a typical young worker would be better off knowing that he or she has greater responsibility for the early retirement years but that Social Security will be there for the later retirement years. The issue is more complex than that suggests, though, and I will return to the nature of that complexity below.

The optimal time profile of real benefits is a separate issue from the choice of an inflation index. Whatever optimal time path of real benefits is selected, it should be achieved using the price index that correctly measures the cost of achieving a constant real level of consumption.

In principle, this price index should reflect the mix of goods and services that older individuals consume. That would not be the same for people in their sixties as it is for those in their eighties. But given the inability of any price index to adequately capture the quality changes and the introduction of new products, this changing mix of the consumption bundle among individuals of different older ages may be of second-order importance.

A more important issue is whether the goal of Social Security should be to help the individual maintain a constant real consumption level or a constant level relative to the rising level of income and consumption among the broader population. An individual who retires at age 62 with a median income will see that level of income decline relative to the median income of the population as she ages. If she is a rational life-cycle planner, she might save in order to maintain that same relative level during retirement. Why would a retired professor want to see her standard of living decline relative to that of younger professors?

I will now ignore that question and turn to a more formal analysis of whether real benefits should rise during retirement. I will then discuss several realistic extensions of that analysis.

SHOULD SOCIAL SECURITY BENEFITS INCREASE WITH AGE? In 1987 I wrote a working paper for the National Bureau of Economic Research titled, "Should Social Security Benefits Increase with Age?"[1] in which I formalized the argument about the advantage of providing longevity insurance in an economy that lacks actuarially fair annuities. That paper begins by showing that if such annuities were available and earned a higher real rate of return than the implicit return in Social Security and if all individuals were rational life-cycle actors, it would be optimal to pay out the entire lifetime Social Security benefit at the start of retirement, allowing the individual to convert it to a private annuity that earns the higher rate of return. But in the more realistic case, if higher-yielding actuarially fair annuities are not available or if some individuals are myopic in their behavior, the paper argues that it would be wrong to front-load benefits in that way.

More specifically, in that paper I use a simple overlapping-generations model with the innovation of two retirement periods, representing young retirees and older retirees, in which everyone lives in the first retirement period but only a fraction live to the second period. With a conventional additive social welfare function, if all individuals are completely myopic and therefore do no saving, the level of benefits should be the same for the young and old retirees who are alive in each period, and therefore the benefits of older retirees should rise relative to their benefits when they were young retirees. Specifically, benefits should rise during retirement at the same rate as the earnings growth of the working population.

More generally, in that paper I showed (and I quote) that "the optimal relation between social security benefits and retiree age depends on balancing the advantage of providing an otherwise unavailable actuarially fair annuity against the lower rate of return earned in a pay-as-you-go social security system. The ability of compulsory social security programs to provide an actuarially fair annuity implies that benefits should increase with age while the lower return on social security contributions than on private saving implies that a larger fraction of total benefits should be paid early in retirement. In an economy that contains a mixture of rational life-cycle savers and completely myopic individuals who do no saving, it is optimal for benefits to decline during the earlier part of the retirement period and then to begin rising" (pp. 12–13).

1. NBER Working Paper no. 2200.

The analysis that I have just summarized looked only at the consumption of retirees and ignored the value of the *unintended* bequests to the next generation. I corrected this in a separate paper published in 1989.[2] It showed that there is a potentially important difference between the structure of benefits that would be preferred by the current population of workers and retirees and the structure of benefits that would maximize the steady-state level of social welfare. The provision of higher benefits to older retirees reduces the amount of saving that is individually optimal and therefore the level of unintended bequests. While those bequests may have no value to the retirees, they are clearly of value to the young workers who receive those bequests. The paper provided an explicit analysis of a case in which the current workers want benefits to increase with age while the Social Security system that maximizes steady-state welfare would provide higher benefits to young retirees than to the very old.

REALISTIC EXTENSIONS. Weinzierl's paper provides detailed calculations of how the choice of the price index can affect social welfare, using different ways of measuring the changes in social welfare that result from the alternative price indexes. As I have indicated, the time path of real benefits is an important issue but one that should be decided separately from the choice of the price index.

My own two earlier analyses used very simple models. I turn therefore to a few additional considerations that would be needed for any practical evaluation of the optimal path of real benefits.

First, consider the matter of *bequests.* One of the puzzling features of the annuities that individuals chose—puzzling at least to an economist—is that individuals very often chose annuities with a "10-year certain" feature. That means that if the individual dies during the first 10 years of the annuity, it will continue to pay out to the individual's heirs. Of course, this feature reduces the amount of the annual annuity payment but individuals nevertheless chose to give up income while they are alive in order to provide a bequest if they die early in retirement.

I recall testifying to the Senate Finance Committee about investment-based Social Security at a time when Senator Moynihan was the chairman of that committee and also the head of a committee on Social Security

2. "Imperfect Annuity Markets, Unintended Bequests, and the Optimal Age Structure of Social Security Benefits," Working Paper no. 2820 (Cambridge, Mass.: National Bureau of Economic Research, 1989). The paper was also published under that title in the *Journal of Public Economics* 41, no. 4 (February 1990).

reform. I explained to the senators why it was not rational for individuals to provide for bequests if they died before the annuity began or during the early years of the annuity. Senator Moynihan objected, saying that the only way that the poor would have some wealth would be to inherit it through bequests.

So an optimal Social Security annuity might allow for bequests, either from Social Security or from the individual's other assets or investment-based annuities.

Another matter worth considering is *endogenous saving.* Consumption during retirement depends on personal saving as well as the Social Security benefits. But the amount of saving that individuals do directly, as well as the amount that company pension plans provide, reflects the amount of Social Security benefits and, presumably, the time schedule of those benefits. So optimizing the time path of Social Security benefits requires analyzing how personal saving and private pensions will respond to those benefit rules.

A third consideration is *retirement decisions.* Under current law, Social Security benefits before age 72 depend on the amount that the potential beneficiary chooses to earn during the years when he is eligible for benefits. Those earnings respond to the Social Security rules. So, again, optimizing the level and time path of Social Security benefits requires taking those retirement and partial retirement decisions into account.

A fourth consideration is *marital status.* Social Security benefits now vary with marital status, providing more to married couples and potentially more to widows and widowers. Spending patterns also depend on marital status and, to some extent, on the living arrangements of widows and widowers. An optimal plan for Social Security benefits has to take into account the marital status and how benefit rules affect marriage and remarriage among widows and widowers.

And finally, there is the issue of *long-term care.* Although it is common to assume that individuals maximize a utility function in which each year is the same as the next, that assumption should change when thinking about retirees, especially because retirees should consider the possibility of needing some degree of long-term institutional care.

In conclusion, Matthew Weinzierl has made an important contribution by focusing attention on the optimal relation between retiree age and the level of retirement benefits. I think that this is separate from the choice of the price index to use for adjusting post-retirement benefits. But I hope it inspires him and others to explore the broader issues that should shape the optimal time path of Social Security benefits.

COMMENT BY

ALEH TSYVINSKI An important question that this paper by Matthew Weinzierl encourages one to think about is the choice of the social welfare function and, more broadly, about the normative foundations of policy. A broad point that I would like to make is that public finance economists have much to gain from carefully understanding the normative issues which are now more often studied by philosophers. The choice of the social welfare function is one of the key determinants of the effects of both the partial reforms considered in this paper and a large number of optimal taxation issues. Typically, the research on these topics considers a utilitarian social welfare function, with different weights attached to the utilities of various agents or groups of agents to capture the redistributive preferences of society.

However, relatively few papers place the choice of the welfare function at the center of the research question. This paper builds on Weinzierl's research agenda, in which he has considered the limitations of the pure utilitarian social welfare function in the context of optimal policy (Mankiw and Weinzierl 2010; Gelber and Weinzierl 2012; Lockwood and Weinzierl 2014).

Concerning the main issue of the paper—the welfare effects of partial reforms in an environment where agents have different mortality—I will start by following and summarizing the arguments of Pierre Pestieau and Grégory Ponthière (2012) on the normative foundations of policy in such environments. For simplicity, consider an economy populated by two types of agents: the long-lived agents who live for two periods and the short-lived agents who live for one period. The proportions of each agent are equal. Each agent has an amount of resources equal to w and a log utility over consumption in a period when they are alive; the utility is normalized to zero when the agent is not alive. The environment is deterministic, and there is no discounting across periods.

Consider first a competitive equilibrium in which the types are already realized. The consumption of the long-lived agent is then $c_1^l = c_2^l = w/2$; and the consumption of the short-lived agent is then $c_1^s = w$. The utility of the long-lived agent is higher than the utility of the short-lived agent:[1]

$$\log(w/2) + \log(w/2) > \log(w).$$

In other words, the long-lived agent already has an advantage in terms of utility compared to the short-lived agent. Even for the same amount of

1. This holds for $w > 4$.

resources that each type has at its disposal, the possibility of spreading consumption across periods yields higher utility for the long-lived agent.

Now, consider a utilitarian social planner who places an equal weight on the utility of each type. The planner's problem is as follows:

$$\max[\log(c_1^l) + \log(c_2^l)] + \log(c_1^s),$$

s.t.

$$(c_1^l + c_2^l) + c_1^s = 2w.$$

The solution to this problem is $c_1^l = c_2^l = c_1^s = \frac{2}{3}w$. Not only does the long-lived agent receive higher utility of consumption for the same amount of allocated resources (as in the case of the competitive equilibrium above), but in addition, the planner allocates twice as much in resources to this agent as to the short-lived agent ($\frac{4}{3}$ vs. $\frac{2}{3}$). In fact, the higher derived marginal utility from an extra unit of resources allocated to the long-lived agent is exactly the reason for the higher amount of resources allocated to him by the social planner. In other words, the bad luck of the short-lived agent that limits him to enjoying consumption for only one period also translates to a redistribution to the lucky long-lived agent.

Viewed from this ex-post perspective, the allocation may seem intuitively unfair. A society may want to undo one or both sides of the disadvantage of the short-lived agent: the lower utility or the lower amount of resources. It is more reasonable to expect that only the unequal distribution of resources may be corrected. A simple way for the planner to achieve this is to choose weights differing from unity on each group in the social welfare function with the goal of redistributing to the unlucky agents and equalizing the lifetime resources. Consider weights $\alpha_l = 1$ and $\alpha_s = 2$, so that the social planner's problem becomes:

$$\max \alpha_l[\log(c_1^l) + \log(c_2^l)] + \alpha_s \log(c_1^s),$$

s.t.

$$(c_1^l + c_2^l) + c_1^s = 2w.$$

The solution to this problem coincides with the ex-post competitive equilibrium problem and allocates the same present value of w of resources to each type.

Several questions arise with the fairness of this possible solution. One is to what extent the length of life is predetermined and is a result of genetic

lottery. Kaare Christensen, Thomas E. Johnson, and James W. Vaupel (2006) compare the longevity of twins and find that one-quarter to one-third of the longevity is determined by genes. Should the social planner then undo only a third or a quarter of the difference or all of the difference?[2]

Another issue is whether such a choice of the welfare weights makes sense in other contexts. Suppose that instead of simply solving a problem of allocating fixed resources, one needs to solve for an economy that also needs to produce. Consider an environment in which all agents live for one period but differ in their productive ability. For example, suppose there are two types: an able agent with a skill of one and a disabled agent with a skill of zero. The utilitarian planner, putting equal weight on each type of agent, chooses equal consumption across types but requires the able agent to produce (and, hence, to incur disutility of work). The able agent then gets a lower utility than the disabled agent. Following similar logic to that above, one must ask, should the society choose the welfare weights to reward effort in this situation, or simply to equalize the resources available for consumption?

There are several other possible normative prescriptions for dealing with differential mortality. Antoine Bommier, Marie-Louise Leroux, and Jean-Marie Lozachmeur (2011) study the design of Social Security with differential mortality and consider an environment in which agents are risk-averse with respect to length of life. This makes the welfare function more concave and essentially brings it closer to the model of the social planner who is inequality-averse. Marc Fleurbaey, Marie-Louise Leroux, Pierre Pestieau, and Grégory Ponthière (2013) propose a social objective that is the maximin on realized lifetime welfare (ex post) that allows for a compensation for unequal lifetimes.

Of course, one can always view the social planner's problem as a standard Rawlsian insurance behind the veil of ignorance, ex-ante before the types are realized. The social planner's allocation can then be decentralized as a competitive equilibrium. In this competitive equilibrium, a continuum of competitive insurance companies offer insurance contracts. Viewed this way, the competitive equilibrium allocations (and the identical social planner's allocations) seem intuitively fair, viewed either ex-ante or ex-post, as the higher lifetime allocation of resources to the long-lived agents is a choice made by the agents when they were in the identical position behind the veil of ignorance.

2. See Pestieau and Ponthière (2012) on how to analyze the implications of choices made to increase or decrease longevity.

In summary, I do not have a definite answer or preference regarding which social welfare function one should use in this context. I see pros and cons in each of the above arguments, additionally complicated by differences in assets, which Weinzierl considers in this paper. Taking this reasoning further: Is the difference in assets due to luck, or is it an effect of thrifty versus profligate behavior? If one takes the normative aspects seriously, should one not consider that the utility may be different at different ages, and that sick and healthy people may enjoy consumption differently? On balance, my personal preference is probably to use a utilitarian social function with some redistributive component.

But then the question remains, how to choose the welfare weights. Weinzierl provides partial answers. I am sympathetic to his first approach, essentially evaluating the linearized effects of policies using the marginal utilities of consumption of different agents adjusted for the length of the lifespan. I am more skeptical about his second approach, which uses survey responses. I also outline another possible approach that may be used to calibrate the welfare weights.

The first approach and its resulting figure 5 in the paper is in many respects similar to the standard evaluation of tax reform. That is, it considers an infinitesimal change in policy and evaluates the first-order effect of this change on the utilitarian social welfare function with equal weights. This change is equal to the marginal utility of consumption of each type, multiplied by the change in consumption. Figure 5 plots these marginal utilities of consumption. Of course, the change in policy is not infinitesimal, especially at the ends of the "seesaw." A better approximation is achieved near the pivot. In other words, at the end of the seesaw this approximation has the usual problems of using a local approximation for large changes.

The second approach that the author uses is to conduct a survey by asking a sequence of questions on preferences. It is difficult for me to be convinced of this method beyond some suggestive evidence for the redistributionary preferences of the users of Amazon Mechanical Turk. Panagiotis Ipeirotis (2010) shows evidence on the demographics of participants using Amazon Mechanical Turk and finds that approximately 50 percent of the workers are located in the United States and 40 percent in India. Weinzierl only considers American workers, but I am not quite sure that selecting out the American responses from this survey can be done even in principle, since the incentive for participants to mimic others' responses is high.

A potential alternative way to choose the welfare weights may be as follows. The author already computes competitive equilibrium allocations given the current policy. But he (or others) could also consider a social

planner restricted to using the same instruments, that is, the same benefit schedules, and who faces the same constraints, such as lack of annuities. It then would be possible to determine the implicit weights and the degree of the redistribution that the planner chooses. Given these weights and the constraints on the policies and market structure, the social planner's allocation would coincide with the optimum. This exercise would be similar to the normal implementation procedure used in optimal taxation literature, but instead of choosing the taxes that implement the planner's solution, it would choose the welfare weights.

REFERENCES FOR THE TSYVINSKI COMMENT

Bommier, Antoine, Marie-Louise Leroux, and Jean-Marie Lozachmeur. 2011. "Differential Mortality and Social Security." *Canadian Journal of Economics/ Revue canadienne d'economique* 44, no. 1: 273–89.

Christensen, Kaare, Thomas E. Johnson, and James W. Vaupel. 2006. "The Quest for Genetic Determinants of Human Longevity: Challenges and Insights." *Nature Reviews–Genetics* 7, no. 6: 436–48.

Fleurbaey, Marc, Marie-Louise Leroux, Pierre Pestieau, and Grégory Ponthière. 2013. "Fair Retirement under Risky Lifetime." PSE Working Paper no. 2013-31. Paris-Jourdan Sciences Economiques, Paris.

Gelber, Alexander, and Matthew Weinzierl. 2012. "Equalizing Outcomes and Equalizing Opportunities: Optimal Taxation When Children's Abilities Depend on Parents' Resources." Working Paper no. 13-014. Harvard Business School.

Ipeirotis, Panagiotis G. 2010. "Demographics of Mechanical Turk." Working Paper CeDER-10-01. NYU Center for Digital Economy Research.

Lockwood, Benjamin B., and Matthew Weinzierl. 2014. "Positive and Normative Judgments Implicit in U.S. Tax Policy, and the Costs of Unequal Growth and Recessions." Working Paper no. 14-119. Harvard Business School.

Mankiw, N. Gregory, and Matthew C. Weinzierl. 2010. "The Optimal Taxation of Height: A Case Study of Utilitarian Income Redistribution." *American Economic Journal: Economic Policy* 2, no. 1: 155–76.

Pestieau, Pierre, and Grégory Ponthière. 2012. "The Public Economics of Increasing Longevity." PSE Working Paper no. 2012-10. Paris-Jourdan Sciences Economiques, Paris.

GENERAL DISCUSSION Ricardo Reis found it striking that a paper on Social Security indexing did not include a discussion of inflation. He noted that while the various indexes of inflation produce similar results over long periods, over periods of one or two years some of them have quite different results. Which measure is used to index Social Security

benefits matters a lot to retirees, especially since they tend to have difficulty accessing credit markets and have little in savings. Additionally, Reis thought that the current measures of inflation—such as the CPI-E—do not accurately measure the consumption patterns of retirees because they neglect the ability to smooth their expenditures over time. They also do not measure things that are especially important to retirees, such as housing prices and interest rates.

Referencing the Bergsonian view of social welfare, Robert Hall argued that time should not appear in the social welfare function—that the social welfare function should only be a function of the time-zero expected utility of the people. In his opinion, the paper included an element of insurance theory which says that insurance should equalize marginal utility over different states of the world. He also noted that imposing linear preferences with respect to longevity is a very strong assumption, one that implies that marginal utility will be the same regardless of a person's age. In fact, he argued, marginal utility can rise or fall as one ages, and indeed for some people—including himself—the idea of surviving beyond the age of 80 is rather unpleasant. He concluded by suggesting that the paper be reorganized to distinguish between what can be accomplished first by means of insurance, such as back loading for safer late-period retirement, and then what can be answered only through social welfare weighting.

Commenting on the benefit bump-up for older workers that Matthew Weinzierl discussed in the paper, Janice Eberly said that the number of low-benefit workers who outlive their life expectancy is small, since low-benefit workers tend to have lower life expectancies. But, she added, there is a sizable group of women who fall into that group. Those women tend to have had low or intermediate labor force participation rates, especially those who are single mothers, whose number continues to grow. Projecting the size of these groups into the future is difficult, because women's labor force participation has been in flux, and even the Social Security actuaries struggle with it. Nevertheless, Eberly suggested, in analyzing the slope of benefits and the design of policies such as benefits bump-ups, the changing forces affecting the size and characteristics of the retirement-age female population are going to be very important.

Katharine Abraham characterized the main problem discussed in the paper as longevity risk and lack of an annuity market. There are many reasons why individuals do not purchase conventional annuities, including their high cost and the desire to retain assets in order to leave a bequest. In principle, "longevity annuity," which individuals could purchase relatively inexpensively when they were 60 or 65 and from which they would

start collecting benefits when they were 80 or 85, should be more attractive to many people. Abraham noted that there have been barriers to developing something like this as a commercial product, though some of these barriers have been addressed by recent Treasury regulations. If a private market for these products can be developed, it could complement the Social Security system in addressing longevity risk.

Caroline Hoxby agreed with other commentators that the paper began as one that addressed the indexing problem and turned into one that was about distributional problems. She thought the same tools should not be used to address both problems. Adding to Eberly's comments about single women in the Social Security system, she noted that there were other similarly interesting populations, such as immigrants who left the country before receiving benefits. Concerning the distributional aspect, she observed that there is a lot that later affects benefits which takes place before a worker reaches 65, such as the formulas that relate payouts to contributions from earnings.

Adding to the discussion on the private annuity market, Martin Baily said that there had been a large increase in the private purchase of annuities and there has been innovation in the type of annuity products offered by the market. Baily suggested that Medicare should be included in distributional analysis of retirement, since so many individuals receive Medicare. He concluded with an anecdote regarding the Australian pension system, which allowed individuals to withdraw the full value of their benefits at the time of retirement. This proved to be quite popular. Individuals quickly spent the entirety of their benefits, and then went on disability, which they could afford to do because the Australian disability system provided roughly comparable benefits to the retirement system.

David Romer argued that it was useful to focus on characteristics of the program that could not be easily addressed through means other than back loading or front loading. For example, there are other fairly easy ways of changing the program's progressivity. Thus, since back loading and front loading have other effects, one probably would not want to use them to address progressivity.

Romer saw two issues that back loading and front loading appeared uniquely able to address. One was the extent to which benefits are annuitized. This points strongly to the desirability of greater back loading. The other is that some people may simply spend their income each period, perhaps because of time-inconsistent preferences. For those individuals, a social planner would want to figure out their optimal consumption pro-

files and set the back loading or front loading of benefits to match that. It is not immediately clear what that implies about whether benefits should be more or less back-loaded than currently. Romer was also struck by the fact that, overall, the paper pointed to increased back loading as desirable, but the only major policy proposal in this area with any traction involves greater front loading.

Following Reis's comment on volatility and indexing, John Haldane questioned whether the price indexes were accurate enough to capture changes in consumption patterns. He cited a National Academy of Sciences panel that studied the Consumer Expenditure Survey (CES) and identified a number of problems. According to the National Academy of Sciences, the nonresponse and measurement errors have increased in the CES. That may cause the estimates of the inflation indexes to become more volatile. He wondered if the proposed chained CPI would be any less volatile.

Discussing Social Security disability insurance (SSI), Andrew Levin argued that SSI has served a different purpose than general Social Security, which was clearly designed as a program people contribute to out of their earnings in order to have a secure pension on retirement and is not "behind the veil." By contrast, SSI is funded out of general tax revenues and is intended to provide a sufficient level of income regardless of whether an individual has paid into the system. In Levin's view, SSI is crucial but it is not well designed. It has a sharp threshold, and it could be improved, for example by applying some of the principles of the Earned Income Tax Credit.

James Stock concurred with other commenters that the indexing and distributional challenges need to remain separated. He urged the economists in the room to return to basics: measuring inflation. As he saw it, there has been steady progress in the measurement of inflation. That progress has hit an institutional constraint, and more economists should focus on the technical matter of how to measure inflation before turning to the distributional policy question.

Matthew Weinzierl responded to the group discussion. He agreed that the question of what index to use and how distributional the system should be are different conceptual questions. He addressed the distributional question in the paper, because while the policy debate has focused on changing the index for a variety of reasons, it has not focused on the distributional question.

Agreeing that SSI is important, Weinzierl noted that people who receive the least in Social Security benefits are essentially covered by SSI; their SSI benefits are completely offset by their Social Security benefits. If the

index for Social Security were changed, it remained unclear what would happen to those Social Security benefits that supplant SSI. That has consequences for reform proposals.

Responding to Romer's comments on how individuals are weighted, Weinzierl said that there was not a clear answer. Nevertheless, concerning individuals who are myopic, he said that weighting by people's preferences makes the results in the paper stronger. He found that people who want the system to help them more at the end of life because they have run out of money actually run out of money faster when the system is back-loaded. And the people who consume their resources quickly because they have low life expectancies are helped even more by front loading because they can consume even more resources quickly.

Weinzierl agreed with Hall that the line between insurance theory and welfare theory had been blurred. He suggested that individuals would be weighted inversely to their life spans, whereas in the paper he had allowed individuals' weights to vary according to their ex-post experience. He agreed that it might be beneficial to separate out the weights for individuals' expected life spans from their experiences.