JANUARY 2020

Evidence-based retirement policy: Necessity and opportunity

William G. Gale
Brookings Institution

Benjamin H. Harris
Results for America

Claire Haldeman
Brookings Institution

This report is available online at: https://www.brookings.edu

Contents

ABOUT THE AUTHORS

William Gale is a senior fellow in Economic Studies at Brookings.
Benjamin H. Harris is the Chief Economist and Senior Adviser with Results for America and a visiting associate professor at the Kellogg School of Management.
Claire Haldeman is a senior research analyst in Economic Studies at Brookings.

DISCLOSURE

The authors did not receive financial support from any firm or person for this article or from any firm or person with a financial or political interest in this article. None of the authors is currently an officer, director, or board member of any organization with a financial or political interest in this article. The authors are not currently an officer, director, or board member of any organization with a financial or political interest in this article.

ACKNOWLEDGEMENTS

The authors thank Zac Coile, Maia Jachimowicz, and David Medina for helpful comments and feedback.

I. Introduction

Retirement saving plays an important role in the U.S. economy. Americans hold more than $18 trillion in private retirement accounts like 401(k)s and IRAs, while defined benefit pensions in the private and public sector hold trillions more. Social Security and Medicare comprise nearly 40 percent of the federal budget. The government also provides tax subsidies for retirement saving, and funds Medicaid, which covers elder long-term care. Retirement issues will only become more important in the future, as the population ages, the Baby Boom retires, lifespans increase, and health care costs rise.[1]

Yet despite existing research, policymakers do not have access to robust empirical consensus when making decisions that affect the retirement security of tens of millions of families. There are many major outstanding questions:

- How well are households preparing for retirement? Are Americans failing to accumulate enough wealth to support themselves in retirement, or are retirement saving shortfalls small and declining?[2]
- Why does consumption fall at retirement? Does this indicate that retirees aren't saving enough, or does it reflect their ability to secure the same quality of life with fewer expenditures?[3]
- Do tax-based saving incentives raise net wealth accumulation? Do the effects vary by saver characteristics and plan design? To what extent is the substantial flow of contributions into 401(k)s and IRAs a net addition to saving, as opposed to saving that would have been done anyway, in other forms?[4]
- What policies boost saving? Would improving financial literacy or mandating saving be effective?[5] Are there retirement saving programs that raise participation and increase overall wealth accumulation?[6]
- Why do households consistently make retirement financing decisions that do not appear to be in their own best interest? For example, why do households buy fewer annuities and reverse mortgages than expected?[7] Why don't more retirees boost expected lifetime benefits by waiting longer to claim Social Security?[8] Are

. . .

1. Gale (2019), Lee (2014), Poterba (2014).
2. Biggs (2019), Munnell, Hou, and Sanzenbacher (2018), Rhee (2013), Scholz, Seshadri, and Khitatrakun (2006).
3. Banks, Blundell, and Tanner (1998), Bernheim, Skinner, and Weinberg (1997).
4. Chetty et al. (2014), Engen and Gale (2000), Engen, Gale, and Scholz (1994, 1996), Poterba, Venti, and Wise (1995, 1996).
5. Chetty et al. (2014), Lusardi and Mitchell (2007).
6. Chetty et al. (2014), Madrian and Shea (2001), Thaler and Benartzi (2004).
7. Baily, Harris, and Wang (2019), Moulton and Haurin (2019).
8. Brown, Kapteyn, and Mitchell (2016).

these choices rationally dictated by either unobserved preferences or imperfect markets? Or are they irrational responses that are better explained by a variety of decision-making biases?

Obtaining better answers to these questions and using the insights they provide to guide changes in the American retirement system could improve living standards for generations of retirees and control the federal budget. But to obtain these answers, researchers and policy makers need better information. There are several major sources of data on U.S. households' saving and wealth, including several large-scale surveys administered by academic or public institutions. These data sets have proven useful for examining many questions but are not comprehensive or extensive enough to generate evidence that can conclusively address the major outstanding questions in retirement policy.

To generate compelling results, researchers need more than access to more comprehensive data— they must also employ better study designs. The research design most conducive to drawing causal inference is the randomized control trial (RCT), where subjects are randomly assigned to treatment and control groups.[9] For example, one of us (Gale) has used an RCT to study the homeownership outcomes of those who used Individual Development Accounts (IDAs).[10] The five-year analysis and the ten-year follow-up found that IDAs accelerate home buying but do not materially affect long-term rates of homeownership—a crucial assessment that helped shift advocates' efforts to help low-income households to other strategies.

But RCTs are not always easy to implement, and researchers often focus instead on "quasi-experimental" research designs or "natural experiments" induced by policy changes or other exogenous events. These research strategies certainly have advantages, but at times the results are hard to interpret because it is unclear what would have happened in the absence of the policy. In any case, the results of well-designed studies are harder to ignore. They can have greater impact on policymakers' decisions because, unencumbered by contorted methodology, the findings inspire greater confidence and understanding in researchers and lawmakers alike.

In the absence of such robust studies, it hardly surprising that almost no retirement policymaking is rooted in evidence; programs simply continue indefinitely with little or no Congressional oversight. Federal tax expenditures for retirement saving—which totaled $252 billion in 2018—have never been formally evaluated, while the Social Security Administration devotes less than 1 percent of its administrative budget to research and evaluation.[11] Because public institutions do not formally evaluate their own programs, academic and think-tank economists provide most of the existing analyses, which are limited by the accessible data, as noted above. Using federal dollars to most efficiently improve retirement security requires building consensus around what makes effective policy—an

. . .

9. The 2019 Nobel Prize in Economics was awarded to MIT professors Abhijit Banerjee and Esther Duflo and Harvard professor Michael Kremer for their work in advancing RCTs in developing countries.

10. IDAs are specialized savings accounts designed to subsidize certain behavior, in this case purchasing a home. See Mills et al. (2009) and Grinstein-Weiss et al. (2013) for further analysis.

11. Joint Committee on Taxation (2018).

impossible task without robust and transparent research methods, empirical replication, and the comprehensive data these processes rely upon.

Strengthening the link between expert consensus and political action could create greater demand for this invaluable data. The federal government has already taken important steps towards implementing a more evidence-based policymaking infrastructure through the *Foundations for Evidence-Based Policymaking Act*, passed earlier this year, which creates new government entities devoted to data sharing and evaluation, while also dramatically increasing researchers' access to administrative data.[12] In addition, in recent years policymakers have adopted some evidence-based policy mechanisms, like Pay for Success and tiered grantmaking, designed to funnel federal dollars to programs with demonstrated effectiveness.[13] These initiatives present opportunities to learn about and improve existing retirement programs, but they also generate a host of new issues, like who decides what constitutes good evidence. But despite these programs, the vast bulk of federal dollars and tax expenditures are awarded without an explicit connection to evidence.

II. Research Issues and Policy Levers

Public policy influences private saving, but it is not always clear exactly *how*. Here we focus on four types of policies that could each be a part of successful retirement policy: savings mandates, tax incentives, financial literacy, and design architecture. In each case, we highlight how better answers will require more data of better quality and more conducive to robust research methods. In addition, we address gaps in knowledge around markets for private insurance products and older workers' labor—both of which can be materially impacted by public policy decisions.

A. Mandates

There is limited evidence on how mandated saving would affect households' overall wealth accumulation. Social Security, of course, already requires mandated contributions, but evidence on how Social Security affects saving in the U.S. is mostly dated, inconclusive, and based on non-robust methods.[14] More recent studies that exploit policy changes in the British and Italian public pension systems found significant substitution between mandated pension contributions and private saving.[15]

But countervailing evidence comes from a study of Danish saving behavior. The study used data on income, employment, and wealth for the entire population of Denmark over a 15-year period. With this data, the study was able to demonstrate—among other things—that a mandate to contribute 1 percent of wages to a retirement savings account brought

. . .

12. For more details on the legislation, see Results for America (2019b).

13. Results for America (2015).

14. Barro (1978), Darby (1978), Feldstein (1979), Munnell (1974).

15. Attanasio and Brugiavini (2003), Attanasio and Rohwedder (2003).

about an increase in net saving of almost 1 percent of wages, even among those who were already saving more than 1 percent of their earnings.[16] The authors highlight the role of passive savers: those who do not change their other saving behavior in response to a change in policy. The study was powerful because the high-quality data captured with compelling precision people's reactions to exogenous national changes in savings rules.

Whether these results can be applied to the United States is a matter up for debate.[17] But only better U.S. data can facilitate the randomized controlled trials necessary to address this uncertainty.

B. Tax-Based Saving Incentives

The United States provides tax incentives for people to accumulate wealth in retirement savings accounts like IRAs and 401(k)s. While it is clear that households make substantial contributions to these accounts, the overall impact on net wealth accumulation is less clear. Demonstrating that these programs are effective would require detailed longitudinal data and a natural or controlled experiment, data that do not exist in the United States. Researchers have responded to this dearth of information by using a variety of methods to back out the relationship between the incentives and overall saving.[18] Some researchers have consistently found that most or all the saving in these tax-deferred plans has been substituted from other kinds of accounts.[19] Others have concluded that a substantial fraction of the saving is new.[20] At times researchers referencing the same data have come to opposite conclusions. But the bottom line is that, in the United States, the policy experiment in tax-based retirement saving incentives is now four decades old, and we still do not have enough information to evaluate its outcomes.

The most recent contribution to the debate is the Chetty et al. (2014) study of Danish saving behavior discussed above. In addition to providing insights into savings mandates, the study also offers clear (but limited) insights into the impact of tax incentives. The authors show that reducing the deduction available to high-income households for their contributions to one type of tax-preferred retirement account reduced contributions to that account, though 57 percent of that reduction was offset by increased contributions to another tax deferred account. The authors also estimate that almost all (99 percent) of this population's contributions to tax-deferred accounts would otherwise have been saved in

. . .

16. Chetty et al. (2014)

17. Critics suggest that there may be a higher proportion of active savers in the United States because its aggregate savings rates are higher than those of Denmark. This would make a mandate less effective in the U.S. than in Denmark because there would be more offsetting behavior. But the difference in savings rates between the two nations is almost entirely explained by differences in dissaving rates among retired people, rather than by differences in savings rates among working age people (the focus of retirement saving initiatives). In fact, other studies conducted in the U.S. context suggest very similar rates of passive saving to those implied by the Danish experiment; see Banerjee and Adams (2013) and the reply from Chetty et al. (2013). For passive saving comparisons, see Madrian and Shea (2001).

18. For a review of the literature and its methodological pitfalls, see Bernheim (2002).

19. Benjamin (2003), Engen and Gale (2000), Engen, Gale, and Scholz (1994, 1996).

20. Gelber (2011), Poterba, Venti, and Wise (1995, 1996), Venti and Wise (1986).

taxable accounts. These findings confirm the idea that high-income households are able to substitute relatively freely between tax-preferred saving and other wealth accumulation, but it is unclear how widely applicable those findings are to other households.

C. Financial Literacy

Improving people's financial literacy is often seen as a way to boost saving and wealth accumulation. But the effect of financial literacy interventions on saving remains unclear.[21] Though some non-experimental survey designs show a positive relationship between the two, these results are weakened by sample selection bias and endogeneity. Some experimental studies have shown that certain kinds of interventions have a positive effect on financial decisions, though the effect tends to be more muted than those touted in observational studies.[22] In addition, the generalizability (external validity) of the results is uncertain.[23] For example, one meta-analysis concluded that financial education improves financial literacy and financial behavior in general, but works less well for the poor, and works better during "teachable moments," as opposed to through education mandates.[24] In contrast, another meta-analysis found that financial education has an almost negligible effect on financial behavior and only a small and diminishing effect on financial literacy.[25] More randomized experiments are required to further test and generalize these findings.

D. Design Architecture

Researchers have made significant contributions to understanding the effects of design architecture on saving behavior. But the theoretical underpinnings of these empirical results remain opaque.

The central results in this literature show that "defaults" affect saving behavior. First, automatic enrollment raises participation rates. Employees enroll in employer-sponsored retirement plans at higher rates when they are automatically enrolled (and must actively choose to opt out if they don't want to participate) than when they must opt in.[26] Second, automatic escalation of contributions raises contributions. In one experiment, 78 percent of workers accepted an offer to automatically divert a portion of their future raises to their retirement savings plan (while the theoretical baseline suggests that no one would find this

. . .

21. Lusardi and Mitchell (2007).

22. One randomized study showed that a workplace seminar improved household saving behavior, while another showed that informational videos about the features of different retirement savings accounts substantially improved the decisions of participants confronted with a variety of saving scenarios. See Boyer, d'Astous, and Michaud (2019) and Duflo and Saez (2003).

23. Gale, Harris, and Levine (2012).

24. Kaiser and Menkhoff (2017).

25. Fernandes, Lynch, and Netemeyer (2014).

26. Beshears et al. (2009), Choi et al. (2001), Choi, Laibson, and Madrian (2004), Karamcheva and Butrica (2012), Madrian and Shea (2001).

option attractive), and the program raised average contribution rates among participants by more than 10 percentage points.[27] Third, the higher contributions that come from automatic enrollment and escalation appear to raise overall wealth accumulation.[28]

But not all of the news is good. Some research has shown that implementing automatic enrollment makes plan participants more likely to accrue more debt.[29] Other research shows that pension participants tend to cluster at the default contribution rate, or at the maximum employer match, potentially inducing less optimal contribution behavior among participants.[30]

But these are all "reduced form" results: they do not provide insight into underlying mechanisms. There is a joke that an economist is someone who sees something work in the real world and wonders if it works in theory. That is, indeed, the situation with design architecture. The empirical literature has compellingly demonstrated that automatic enrollment raises enrollment, automatic escalation of contributions raises contributions, and the increased contributions from automatic enrollment turn into higher overall wealth. But we don't know why. The results could be due to obliviousness, myopia, procrastination, or other factors.[31] Further research, combining more detailed data collection, careful theoretical specification, and randomized controlled experiments, could help policymakers and researchers understand why automatic policies work well and how to extend that insight into other areas.

E. Private Insurance Products

Social Security and Medicare are the foundation of most people's retirement-related insurance, but private products—including long-term care insurance, annuities, and reverse mortgages—can play an important supplementary role.

These products share two features. First, in most cases, they receive public subsidies. Those who purchase long-term care insurance can claim premiums as an itemized deduction; annuities receive preferential tax treatment on the "inside build-up" of investments; and the bulk of reverse mortgages are issued through the Department of Housing and Urban Development (which has provided an implicit subsidy in recent years by insuring private loans made on behalf of lenders). Second, research suggests that take-up of each of these products has many benefits and should be higher than is observed in the real world. The difference between research findings and observed behavior suggests a gap in the knowledge policymakers need to make informed choices.

For example, although many people will need long-term care, take-up rates for long-term care insurance are low. In 2010, an estimated 12 million Americans needed long-term support and services (a number expected to rise to 27 million by 2050), yet the number of private long-term care insurance policies purchased each year plummeted from 750,000

. . .

27. Thaler and Benartzi (2004).

28. Chetty et al. (2014).

29. Beshears et al. (2017).

30. Madrian (2012), Madrian and Shea (2001).

31. Bernheim, Fradkin, and Popov (2015).

to 89,000 between 2002 and 2016.[32] Several studies have attributed this lack of demand to the Medicaid backstop, under which retirees who sufficiently deplete their income and assets can receive long-term services and supports.[33] This finding led the Commission on Long-Term Care to recommend that policymakers "minimize Medicaid crowd-out" and "strengthen Medicaid eligibility requirements for middle-income Americans."[34]

Economists have reached similar conclusions about annuities. A large literature surrounds the "annuity puzzle"—or the gap between the theoretical and observed demand for annuities. Many papers have shown that older Americans can increase their welfare by partially annuitizing a fraction of their financial assets.[35] Despite these theoretical findings, annuity ownership rates remain low. Given the disconnect between theory and evidence, policymakers have been slow to implement changes. Perhaps the most significant recent change has been a 2014 Treasury guideline establishing "Qualifying Longevity Annuity Contracts," which allow retirement savers to direct a share of their tax-preferred savings to a deferred annuity without violating minimum distribution rules. The *SECURE Act,* signed into law in late December 2019, reduces barriers for plan sponsors seeking to offer annuities to their participants, which may marginally increase demand for annuities.[36]

Reverse mortgages are a third form of private-sector "insurance" with a substantial gap between academic evidence and policy action. Although they are non-recourse loans backed by a homeowner's primary residence, reverse mortgages are a form of insurance since they protect older homeowners against falling home prices and extended longevity.[37] Like annuities, the predicted demand for reverse mortgages is far above that observed in the actual market.[38] While the literature on reverse mortgages pales in comparison to that of annuities, a handful of studies document that demand is suppressed by high fees (which

. . .

32. But at the same time, the market for hybrid products that combine life and long-term care insurance is growing.

33. For example, Brown and Finkelstein (2008) find that for a median-wealth man, 60 percent of the benefits of a long-term care policy are redundant with Medicaid benefits; this share rises to 75 percent for a similarly aged woman.

34. Atkins et al. (2013)

35. For example, Gong and Webb (2010) find that retirees can improve their wellbeing by about 5–10 percent through purchasing deferred income annuities. Horneff, Maurer, and Mitchell (2019) find that defaulting 10 percent of retirement saving into a deferred income annuity boosts average consumption by $700 at age 85 and $2,600 by age 95. Income annuities refer to annuity contracts that pay annuitants a prescribed amount of income for a set period of time, often for life. This differ in practice from variable annuities, which are often investment vehicles for investors in high marginal tax rates that may never be converted into an annuity. For greater exposition of the annuitization puzzle, see Benartzi, Previtero, and Thaler (2011), Brown (2007), Davidoff, Brown, and Diamond (2005), Gong and Webb (2007), Mitchell et al. (1999), and Yogo (2009).

36. See Baily, Harris, and Iwry (2019) for additional discussion.

37. The term "non-recourse" refers to the characteristic of the loan that allows the balance of the amount borrowed, plus interest, to exceed the value of the home.

38. Merrill, Finkel, and Kutty (1994) find that about 9 percent of older homeowners would benefit from the reverse mortgages. Morgan, Megbolugbe, and Rasmussen (1996) estimate that nearly 2 million low-income older women could boost their income with reverse mortgages.

compensate lenders for the riskiness of the product) and supply is suppressed by high rates of foreclosures (which happens when reverse mortgage borrowers don't pay property taxes or insurance premiums, or move out of their homes). Proposed reforms offer ways to reduce costs or foreclosures.[39] These proposals could help raise the number of loans in the market, but they have not been a political priority.

F. Retirement Timing

In contrast to the areas above, many aspects of people's choices about when to retire have been fairly well-studied. Key determinants of the decision include the rules of Social Security, Disability Insurance, and Medicare; the incentives in private pensions; the status of labor markets; the business cycle; job quality; health shocks; and the retirement choices of one's spouse.[40]

All told, the decision about retirement timing is likely one of the more-studied aspects of the public finance and labor literatures, primarily because the data available are sufficient to answer the questions posed. First, several large surveys are well-designed to address the retirement decision. The longitudinal nature of the Health and Retirement Study, coupled with its detailed labor market, health, demographic, and financial data, allows researchers to study decisions around the labor supply of older workers. Second, several generous sources of funding have allowed scholars to undertake this research. This includes public sources, such as the Social Security Administration's annual research budget of roughly $100 million, and funding from private foundations. Why people decide to retire when they do is a question well-understood by researchers, the result of significant public investment in data collection coupled with dedicated resources from public and private sources.

III. New Directions for Evidence-Based Policy

Researchers and policymakers have an uneven and inconsistent understanding of the retirement landscape. They have gained precise and actionable insights about retirement timing but have much less clarity on other issues—such as the efficacy of financial education programs, the reasons for the limited demand for private insurance products, and the impact of saving mandates and incentives.

Making effective evidence-based policy requires both a comprehensive understanding of the policy challenge and attendant remedies, as well as the will and ability to implement the necessary reforms. In this section, we suggest several reforms that can advance evidence-based policymaking in the retirement sphere.

. . .

39. Davidoff (2019), Goodman (2019), and Moulton and Haurin (2019).

40. Angrisani et al. (2013), Coile and Levine (2011), Conley and Thompson (2013), Johnson and Favreault (2001), Song and Manchester (2007), Stock and Wise (1990).

From the outset, we stipulate a widespread and pressing need to reform many aspects of the saving and retirement system. Saving must be made easier—such as through automatic enrollment and investment strategies—and the returns to saving must be made more equitable—for example, by expanding the Savers' Credit or equalizing the tax benefits of defined contribution plans across contributors. But more broadly, there is a need to understand and change the framework in which people operate and markets and public policies function. The following policies are designed to address these issues.

Devote 1 Percent of Program Funds to Evaluation

The bulk of federal spending programs often escape rigorous evaluation. A study by the Government Accountability Office found that less than 40 percent of federal program managers reported that their programs had been evaluated in the prior five years.[41] This lack of evaluation is due, in part, to lack of funds for this purpose.

We support the guiding notion—long held by Results for America—that 1 percent of discretionary program funds should be earmarked for rigorous, independent evaluations.[42] We also believe that for many mandatory programs, earmarking 1 percent of mandatory funds for evaluations would markedly improve programs' impact. At the same time, grant recipients should be authorized (but not required) to devote grant receipts to various evaluation activities, as well. In all cases, evaluation should focus on measuring the impact of a given program, in addition to identifying new ways to better deliver public services to the target population.

There are several examples of progress in this area. For example, appropriations laws passed since FY2016 have permitted the Secretary of Labor to devote up to 0.75 percent of program funds for evaluation, and the *Every Student Succeeds Act* allows states to devote up to 0.50 percent of funds for evaluation.[43] While these steps incrementally move the federal government towards a more evidence-based framework, more must be done in program areas related to retirement.

A natural place to increase funding for evaluation is the Social Security Administration's evaluation budget. This comprises only a tiny share of overall expenditures, about 0.01 percent of the annual $1.1 trillion devoted to the program on the whole. Still, expenditures for research and demonstration projects are substantial: in recent years, Congress has appropriated $101 million to various research, outreach, and demonstration projects—including those devoted to either better provision of data (such as funding a supplement to the Health and Retirement Survey) or program evaluation (such as the RETAIN demonstration to test labor market interventions for workers who suffered a recent injury or disability).[44] While these expenditures are sizable, not all of this budget is used for program evaluation, and even if it was, the FY2020 budget request would comprise less than 0.8

. . .

41. Government Accountability Office (2013).

42. Results for America (2017).

43. Results for America (2019a).

44. Social Security Administration (2019).

percent of Social Security's total administrative funding; raising this to the 1 percent threshold would mean tens of millions more for program evaluation—leading to a better understanding of impacts of the more than $1 trillion spent annually through Social Security's various programs.[45]

Evaluate Retirement Saving Tax Expenditures

From homeownership to employer-provided health insurance, the tax code is the source of substantial social and economic policy. The tax provisions aimed at changing various behaviors are collectively referred to as "tax expenditures." These expenditures amount to roughly $1.5 trillion annually—approximately the same size as the annual budget for all discretionary spending. As noted above, the bulk of retirement saving policy is delivered through the tax code with more than $200 billion in annual incentives for retirement security, including saving provisions and tax breaks for employee pensions.

Despite the massive size of tax expenditures, these government programs are rarely evaluated for effectiveness. There is no annual review of how Congress spends through the tax code, no program staff dedicated to tax expenditures' administration, no inspector general for tax expenditures, and no rigorous government evaluation of the effectiveness of most tax breaks.[46] This dearth of systematic evaluation is in stark contrast to direct spending programs, including both discretionary and mandatory programs, which are typically subject to Congressional oversight and several layers of regular programmatic evaluation.

A natural step to improve retirement policy is to subject tax expenditures in general, and retirement incentives in particular, to a level of evaluation which approximates that for direct spending programs. Given recent economic research calling into question whether retirement saving incentives materially change net saving at all, investing in evaluation could potentially shift hundreds of billions to more effective strategies for increasing retirement saving. Harris, Steuerle, and Quakenbush (2018) lay out several options for increasing oversight and evaluation of tax expenditures. These include funding regular tax expenditure evaluations by executive branch offices, such as the Treasury Department and the Office of Management and Budget; expanding capacity at nonpartisan Congressional agencies—such as the Congressional Budget Office, the Joint Committee on Taxation, the Government Accountability Office, or the Congressional Research Service—to take on formal evaluation authority; funding evaluations by independent, non-governmental entities; and undertaking structural reform of budget processes—including requiring periodic reauthorization of expenditures or establishing a joint commission to study tax expenditures.

. . .

45. Social Security Administration (2019).

46. Harris, Steuerle, and Quakenbush (2018) further explain that "Although the Government Performance Results Act (GPRA) of 1993, as modified by the GPRA Modernization Act of 2010, aims to provide a performance planning and reporting framework, its implementation has been limited regarding tax expenditures. GAO (2017) recently found little progress in the executive branch toward systematically evaluating tax expenditures. In response, OMB (2018) provided a mere two pages describing a broad framework for evaluating tax expenditures in its Analytical Perspectives accompanying the fiscal year 2019 budget, identifying few concrete steps toward meaningful evaluation."

Collect Comprehensive, Linked Data

The federal government should commission a large-scale panel survey that, coupled with linked administrative data, provides detailed information on household wealth, labor market, health and demographic data. No current data set provides such information. The Panel Study of Income Dynamics (PSID) provides longitudinal data on about 4,800 households and their descendants since 1968, but has limited health and financial information. The Health and Retirement Study collects longitudinal data on a series of cohorts every two years, including more than 43,000 older Americans since its inception in 1992, but surveys only participants near or in retirement. The Survey of Consumer Finances is conducted once every three years by the Federal Reserve and includes a cross-section of about 6,000 observations with rich data on wealth and personal finance, but does not track individual households over time. The Survey of Income and Program Participation (SIPP) is a large-scale survey with information on households' participation in public programs, but it faces significant under-reporting and mis-reporting.[47]

What the United States needs is a data set that closely mirrors the one used by Chetty et al. in their study of Danish saving behavior: a large, long-term, detailed data set linked to administrative records.[48] This data set should mirror the PSID in its longevity and panel structure, but include more health and financial data, and should link data to key administrative data sets, like tax returns, Social Security records, and Census data. This will allow the information collected to be corrected for some of the inaccuracies inherent in self-reporting, while also providing context to augment and enrich what administrative data alone report.[49]

This effort will require commitment to overcoming several obstacles. First, the process of data-linking must honor and protect the privacy of those whose data the platform relies upon. Defending data against de-anonymization, unauthorized access, and pernicious uses will require not just relying on current federal privacy law, but committing to secure data storage, management, and dissemination. Second, the survey must use technology capable of integrating data across platforms. Third, the survey should be subject to data management and ethics evaluations, to ensure that this boon for researchers does not become a burden those they study.

These guardrails are essential but are not reasons to avoid investing in a large-scale panel survey linked to administrative data. The costs of broader data collection will be more than balanced by the concrete, essential insights into retirement policy that such a study would provide.

. . .

47. Meyer, Mok, and Sullivan (2015).

48. For additional details on the benefits of linking survey and administrative data, see Doar and Gibbs (2017).

49. Meyer, Mok, and Sullivan (2015).

REFERENCES

Angrisani, Marco, Michael D. Hurd, Erik Meijer, Andrew M. Parker, and Susann Rohwedder. 2013. "Labor Force Transitions at Older Ages: The Roles of Work Environment and Personality." Brief 295. Michigan Retirement Research Center.

Atkins, G. Lawrence, Christine Bishop, Pamela Mazerski, Elizabeth Blair, and Dennis Kodner. 2013. "Commission on Long-Term Care: Report to Congress." Washington, D.C.: United States Congress.

Attanasio, Orazio, and Agar Brugiavini. 2003. "Social Security and Households' Saving." The Quarterly Journal of Economics 118 (3): 1075 - 1119.

Attanasio, Orazio, and Susann Rohwedder. 2003. "Pension Wealth and Household Saving: Evidence from Pension Reforms in the United Kingdom." American Economic Review 93 (5): 1499 - 1521.

Baily, Martin N., and Benjamin H. Harris, and J. Mark Iwry. 2019. "'RESA' 2019 Legislative Proposals to Improve Retirement Security and Saving." Washington, DC: The Brookings Institution.

Baily, Martin N., Benjamin H. Harris, and Ting Wang. 2019. "The Unfulfilled Promise of Reverse Mortgages: Can a Better Market Improve Retirement Security?" Washington, DC: The Brookings Institution.

Banerjee, Sudipto, Adams, Nevin. 2013. "Tax Preferences and Mandates: Is the Danish Savings Experience Applicable to the United States?" EBRI Notes.

Banks, James, Richard Blundell, and Sarah Tanner. 1998. "Is There a Retirement-Savings Puzzle?" The American Economic Review 88 (4): 769 - 788.

Barro, Robert J. 1978. "The Impact of Social Security on Private Savings: Evidence from the U.S. Time Series." Washington, DC: American Enterprise Institute.

Benartzi, Shlomo, Alessandro Previtero, and Richard H. Thaler. 2011. "Annuitization Puzzles." Journal of Economic Perspectives 25 (4): 143-164.

Benjamin, Daniel J. 2003. "Does 401(k) Eligibility Increase Saving? Evidence from Propensity Score Subclassification." Journal of Public Economics 87: 1259 - 1290.

Bernheim, B. Douglas. 2002. "Taxation and Saving." Handbook of Public Economics vol. 3 1173-1249. Elsevier.

Bernheim, B. Douglas, Andrey Fradkin, and Igor Popov. 2015. "The Welfare Economics of Default Options in 401(k) Plans." *American Economic Review* 105(9): 2798 – 2837.

Bernheim, B. Douglas, Jonathan Skinner, and Steven Weinberg. 1997. "What Accounts for the Variation in retirement Wealth Among U.S. Households?" Working Paper 6227. Cambridge, MA: National Bureau of Economic Research.

Beshears, John, James J. Choi, David Laibson, Brigitte C. Madrian, and William L. Skimmyhorn. 2017. "Borrowing to Save? The Impact of Automatic Enrollment on Debt." Cambridge, MA: Harvard University.

Beshears, John, James J. Choi, David Laibson, and Brigitte C. Madrian. 2009. "The Importance of Default Options for Retirement Saving Outcomes: Evidence from the United States." Social Security Policy in a Changing Environment. University of Chicago Press.

Biggs, Andrew. 2019. "The Phony Retirement Crisis." Wall Street Journal Opinion.

Boivie, Ilana. 2018. "Pensionomics 2018: Measuring the Economic Impact of DB Pension Expenditures." Washington, DC: National Institute on Retirement Security.

Boyer, M. Martin, Philippe d'Astous, Philippe, and Pierre-Carl Michaud. 2019. "Tax-Sheltered Retirement Accounts: Can Financial Education Improve Decisions?" Working Paper 26128. Cambridge, MA: National Bureau of Economic Research.

Brown, Jeffrey R., and Amy Finkelstein. 2008. "The Interaction of Public and Private Insurance: Medicaid and the Long-Term Care Insurance Market." American Economic Review 98 (3): 1083-1102.

Brown, Jeffrey R., Arie Kapteyn, and Olivia S. Mitchell. 2016. "Framing and Claiming: How Information-Framing Affects Expected Social Security Claiming Behavior." The Journal of Risk and Insurance 83 (1): 139 - 162.

Brown, Jeffrey R. 2007. "Rational and Behavioral Perspectives on the Role of Annuities in Retirement Planning." Working Paper 13537. Cambridge, MA: National Bureau of Economic Research.

Chetty, Raj, John N. Friedman, Søren Leth-Petersen, Torben H. Nielsen, and Tore Olsen. 2014. "Active vs. Passive Decisions and Crowd-Out in Retirement Savings Accounts: Evidence from Denmark." The Quarterly Journal of Economics 129 (3): 1141-1219.

Chetty, Raj, John N. Friedman, Søren Leth-Petersen, Torben H. Nielsen, and Tore Olsen. 2013. "Is the Danish Savings Experience Applicable to the U.S.? Response to Banerjee and Adams (2013) EBRI Report."

Choi, James J., David Laibson, David, and Brigitte C. Madrian. 2004. "Plan Design and 401(k) Savings Outcomes." Working Paper 10486. Cambridge, MA: National Bureau of Economic Research.

Choi, James J., David Laibson, David, Brigitte C. Madrian, and Andrew Metrick. 2001. "For Better or For Worse: Default Effects and 401(k) Savings Behavior." Working Paper 8651. Cambridge, MA: National Bureau of Economic Research.

Coile, Courtney C., and Phillip B. Levine. 2011. "Recessions, Retirement, and Social Security." American Economic Review 101 (3): 23-28.

Conley, Dalton, and Jason Thompson. 2013. "The Effects of Health and Wealth Shocks on Retirement Decisions." Federal Reserve Bank of Saint Louis Review 95 (5): 389 - 404.

Darby, Michael. 1978. "The Effects of Social Security on Income and Capital Stock." Washington, DC: American Enterprise Institute.

Davidoff, Thomas, Jeffrey R. Brown, and Peter A. Diamond. 2005. "Annuities and Individual Welfare." The American Economic Review 95 (5): 1573 - 1590.

Davidoff, Thomas. 2019. "Annuity-Enhanced Reverse Mortgage Loans." Washington, DC: The Brookings Institution.

Doar, Robert, and Linda Gibbs. 2017. "Unleashing the Power of Administrative Data: A guide for Federal, State, and Local Policymakers." Washington, DC: Results for America; The American Enterprise Institute.

Duflo, Esther, and Emmanuel Saez. 2003. "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment." The Quarterly Journal of Economics 118 (3): 815-842.

Engen, Eric M., William G. Gale, and John Karl Scholz. 1994. "Do Saving Incentives Work?" Brookings Papers on Economic Activity 25 (1): 85-180.

Engen, Eric M., William G. Gale, and John Karl Scholz. 1996. "The Effects of Tax-Based Saving Incentives on Saving and Wealth." Working Paper 5759. Cambridge, MA: National Bureau of Economic Research.

Engen, Eric M., and William G. Gale. 2000. "The Effects of 401(k) Plans on Household Wealth: Differences Across Earnings Groups." Working Paper 8032. Cambridge, MA: National Bureau of Economic Research.

Feldstein, Martin. 1979. "The Effect of Social Security on Private Saving: The Time Series Evidence." Working Paper 314. Cambridge, MA: National Bureau of Economic Research.

Fernandes, Daniel, John G. Lynch Jr., and Richard G. Netemeyer. 2014. "Financial Literacy, Financial Education, and Downstream Financial Behaviors." Management Science 60 (8): 1861-1883.

Gale, William G. 2019. Fiscal Therapy: Curing America's Debt Addiction and Investing in the Future. Oxford University Press.

Gale, William G., Benjamin H. Harris, and Ruth Levine. 2012. "Raising Household Saving: Does Financial Education Work?" Social Security Bulletin 72 (2): 39-48.

Gelber, Alexander M. 2011. "How Do 401(k)s Affect Saving? Evidence from Changes in 401(k) Eligibility." American Economic Journal: Economic Policy 3 (4): 103-122.

Gong, Guan, and Anthony Webb. 2007. "Evaluating the Advanced Life Deferred Annuity: An Annuity People Might Actually Buy." Discussion Paper 714. London: The Pensions Institute.

Gong, Guan, and Anthony Webb. 2010. "Evaluating the Advanced Life Deferred Annuity -- An Annuity People Might Actually Buy." Insurance: Mathematics and Economics 46 (1): 210-221.

Goodman, Laurie. 2019. "Protecting Seniors: A Review of the FHA's Home Equity Conversion Mortgage (HECM) Program." Subcommittee on Housing, Community Development, and Insurance, Committee on Financial Services. United States House of Representatives. Washington, DC: The Urban Institute

Government Accountability Office. 2013. "Program Evaluation: Strategies to Facilitate Agencies' Use of Evaluation in Program Management and Policy Making." GAO-13-570. Washington, D.C.

Government Accountability Office. 2017. "Managing for Results: Further Progress Made in Implementing GPRA Modernization Act, but Additional Actions Needed to Address Pressing Governance Challenges." GAO17-775. Washington, D.C.

Grinstein-Weiss, Michal, Michael Sherraden, William G. Gale, William M. Rohe, Mark Schreiner, and Clinton Key. 2013. "Long-Term Impacts of Individual Development Accounts on Homeownership Among Baseline Renters: Follow-Up Evidence from a Randomized Experiment." American Economic Journal: Economic Policy 5 (1): 122-145.

Harris, Benjamin H., Eugene Steuerle, and Caleb Quakenbush. 2018. "Evaluating Tax Expenditures: Introducing Oversight into Spending Through the Tax Code." Washington, D.C.: Tax Policy Center.

Horneff, Vanya, Raimond Maurer, and Olivia S. Mitchell. 2019. "How Will Persistent Low Expected Returns Shape Household Economic Behavior?" Journal of Pension Economics & Finance 18 (4): 612-622.

Johnson, Richard W., and Melissa M. Favreault. 2001. "Retiring Together or Working Alone: The Impact of Spousal Employment and Disability on Retirement Decisions." Working Paper 2001-01. Chestnut Hill, MA: Center for Retirement Research.

Joint Committee on Taxation. 2018. "Estimates of Federal Tax Expenditures for Fiscal years 2018 - 2022." JCX-81-18. Washington, DC: United States Congress.

Kaiser, Tim, and Lukas Menkhoff. 2017. "Does Financial Education Impact Financial Literacy and Financial Behavior, and If So, When?" Policy Research Working Paper 8161. Washington, DC: The World Bank.

Karamcheva, Nadia, and Barbara A. Butricia. 2012. "Automatic Enrollment, Employee Compensation, and Retirement Security." Working Paper 2012-25. Chestnut Hill, MA: Center for Retirement Research.

Lee, Ronald. 2014. "Macroeconomic Consequences of Population Aging in the United States: Overview of a National Academy Report." *American Economic Review: Papers and Proceedings* 104(5): 234-39.

Lusardi, Annamaria, and Olivia S. Mitchell. 2007. "Financial Literacy and Retirement Preparedness: Evidence and Implications for Financial Education." Business Economics 42 (1): 35-44.

Madrian, Brigitte C., and Dennis F. Shea. 2001. "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." The Quarterly Journal of Economics 116 (4): 1149-1187.

Madrian, Brigitte C. 2012. "Matching Contributions and Savings Outcomes: A Behavioral Economics Perspective." Working Paper 18220. Cambridge, MA: National Bureau of Economic Research.

Merrill, Sally R., Merly Finkel, and Nandinee K. Kutty. 1994. "Potential Beneficiaries from Reverse Mortgage Products for Elderly Homeowners: An Analysis of AHS Data." Journal of the Real Estate and Urban Economics Association 22 (3).

Meyer, Bruce D., Wallace K. C. Mok, and James X. Sullivan. 2015. "Household Surveys in Crisis." Journal of Economic Perspectives 29(4): 199 – 226.

Miller, Margaret, Julia Reichelstein, Christian Salas, and Bilal Zia. 2014. "Can You Help Someone Become Financially Capable? A Meta-Analysis of the Literature." Policy Research Working Paper 6745. Washington, DC: The World Bank.

Mills, Gregory, William G. Gale, Rhiannon Patterson, Gary V. Engelhardt, Michael D. Eriksen, and Emil Apostoloy. 2009. "Effects of Individual Development Accounts on Asset Purchases and Saving Behavior: Evidence from a Controlled Experiment." Journal of Public Economics 92 (5 - 6): 1509 - 1530.

Mitchell, Olivia S., James M. Poterba, Mark J. Warshawsky, and Jeffrey R. Brown. 1999. "New Evidence on the Money's Worth of Individual Annuities." The American Economic Review 89 (5): 1299 - 1318.

Morgan, Barbara A., Isaac F. Megbolugbe, and David W. Rasmussen. 1996. "Reverse Mortgages and the Economic Status of Elderly Women." The Gerontologist 36 (3): 400 - 405.

Moulton, Stephanie, and Donald Haurin. 2019. "Unlocking Housing Wealth for Older Americans: Strategies to Improve Reverse Mortgages." Washington, DC: The Brookings Institution

Munnell, Alicia. 1974. "The Effect of Social Security on Personal Saving." Cambridge, MA: Ballinger Publishing Company.

Munnell, Alicia, Wenliang Hou, and Geoffrey Sanzenbacher. 2018. "National Retirement Risk Index Shows Modest Improvement in 2016." Issue Brief 2018-1. Chestnut Hill, MA: Center for Retirement Research.

Office of Management and Budget. 2018. *Analytical Perspectives, Budget of the United States Government, Fiscal Year 2019*. Washington, D.C.

Poterba, James M., Steven F. Venti, and David A. Wise. 1995. "Do 401(k) Contributions Crowd Out Other Personal Saving?" Journal of Public Economics 58 (1): 1 – 32.

Poterba, James M., Steven F. Venti, and David A. Wise. 1996. "How Retirement Programs Increase Savings." Journal of Economic Perspectives 10 (4): 91 – 112.

Poterba, James M. 2014. "Retirement Security in an Aging Society." Working Paper 19930. Cambridge, MA: National Bureau of Economic Research.

Results for America. 2015. "Invest in What Works Fact Sheet: Pay for Success." Washington, DC.

Results for America. 2017. "Funding What Works: Incorporating Pay for Success Language in Federal Legislation." Washington, DC.

Results for America. 2019a. "9 Ways to Make Federal Legislation Evidence-Based: 2019 What Works Guide for Congress." Policy Brief. Washington, D.C.

Results for America. 2019b. "The Promise of the *Foundations for Evidence-Based Policymaking Act* and Proposed Next Steps." Evidence Act Brief, Washington, DC.

Rhee, Nari. 2013. "The Retirement Savings Crisis: Is It Worse Than We Think?" Washington, DC: National Institute on Retirement Security.

Scholz, John Karl, Ananth Seshadri, and Surachai Khitatrakun. 2006. "Are Americans Saving "Optimally" for Retirement?" Journal of Political Economy 114 (4): 607 - 643.

Social Security Administration. 2019. "FY 2020 Congressional Justification." Washington, D.C.

Song, Jae, and Joyce Manchester. 2007. "Have People Delayed Claiming Retirement Benefits? Responses to Changes in Social Security Rules." Social Security Bulletin 67 (2).

Stock, James H., and David A. Wise. 1990. "Pensions, the Option Value of Work, and Retirement." Econometrica 58 (5): 1151-1180.

Thaler, Richard H., and Shlomo Benartzi. 2004. "Save More Tomorrow™: Using Behavioral Economics to Increase Employee Saving." Journal of Political Economy 112 (S1): S164-S187.

Venti, Steven F., and David A. Wise. 1986. "IRAs and Saving." Working Paper 1879. Cambridge, MA: National Bureau of Economic Research

Yogo, Motohiro. 2009. "Portfolio Choice in Retirement: Health Risk and the Demand for Annuities, Housing, and Risky Assets." Working Paper 15307. Cambridge, MA: National Bureau of Economic Research.