Health Savings Accounts and Life-Cycle Saving: Implications for Retirement Preparedness*

Leora Friedberg
University of Virginia

Jaeki Jang
Korea Institute
for Industrial Economics and Trade

Adam Leive
University of California, Berkeley

Eric R. Young
University of Virginia
and Federal Reserve Bank of Cleveland

October 2023

Abstract

This paper analyzes how Health Savings Accounts (HSAs) can be used in conjunction with other savings vehicles to improve retirement preparedness. We build a life-cycle model that incorporates an HSA, a tax-advantaged defined contribution (DC) account, and a liquid taxable account. Individuals face multiple shocks and are liquidity constrained. They face a medical spending shock that can be satisfied by the HSA but not the DC account, and both impose contribution limits. The complexity of our economic model requires the use of machine-learning methods to overcome the curse of dimensionality due to the presence of numerous state variables. We find that, at low levels of DC saving, HSAs are complementary to DC accounts, but at high levels of DC saving, HSAs are substitutes. Complementarity arises because HSAs provide liquidity to finance both health care and other spending shocks, since out-of-pocket health care costs can be paid for out of the HSA retroactively, which allows people to lock up more saving in illiquid DC accounts. Compared to a system with only tax-preferred illiquid accounts and taxable liquid saving, adding HSAs raises the optimal amount of total contributions to tax-preferred accounts.

1 Introduction

Health Savings Accounts (HSAs) are now commonplace, as employers that offer health insurance increasingly embrace high-deductible health plans (HDHPs). In 2020, over half of employers with at least 200 employees offered an HDHP with an HSA, and nearly a quarter of employees across all firms were enrolled in one (Claxton et al. 2021). While often presented as a way to finance current health care expenses, HSAs actually resemble a retirement savings vehicle (Aaron, Healy and Khitatrakun 2008, Leive 2022) that offers key advantages over traditional tax-deferred (or Roth) retirement accounts. As with a defined contribution (DC) plan like 401(k) or 403(b) account, HSA contributions are income tax-deductible and interest grows tax-deferred; yet, contributions are also exempt from FICA taxes, and assets are not subject to required minimum distributions at older ages until they are rolled into a DC account. Moreover, HSA funds remain accessible on a pre-tax basis for health care expenses incurred at not just the current but also at earlier times, providing unparalleled flexibility.[1]

Despite these superior features, HSA balances remain small, with an average of just \$3,600, and are almost entirely held in cash (Fronstin and Spiegel 2021). Many employees go so far as to shun HDHP/HSA plans, even when they (commonly) dominate other health insurance choices on a current basis, *ignoring* their life-cycle saving features (Leive, Friedberg and Davis 2022, Liu and Sydnor 2022). With the reasons for this underuse remaining unclear, the goal of this research is to study the potential gains from their use as a savings vehicle.[2] To do so, we develop methods to solve an optimal life cycle model with multiple heterogeneous savings vehicles, which has broad applications in the setting of retirement planning policy. We demonstrate how individuals can improve retirement preparedness by developing strategies for use of HSAs in combination with both DC retirement accounts and non-employer accounts that are taxable and liquid.

By unifying the analysis of both health insurance and retirement plan choices—often viewed as unrelated—we demonstrate the dual life-cycle savings possibilities and liquidity insurance features of HSA plans, and we investigate complementarities between HSA and DC plan saving. This builds on research demonstrating the correlation of puzzling choices across both health insurance and retirement saving domains. These choices do not take the form of scrimping on cash outlays; rather, with surprising frequency, individuals who fail to take advantage of employer matching funds *overpay* for health insurance by avoiding

---

[1]This flexibility contrasts drastically with a Flexible Spending Account (FSA), which is in fact much less flexible. Only \$500 of unspent assets might be carried over (depending on the employer) from one year to the next, and FSA funds cannot be invested.

[2]Some hypothesized reasons why many consumers avoid HDHPs include information frictions and perceived hassle costs (Handel and Kolstad 2015), inertia (Handel 2013), liquidity constraints (Ericson and Sydnor 2018, Davis, Leive and Gellert 2022), and financial literacy (Davis, Leive and Gellert 2022).

the HDHP/HSA plan (Leive, Friedberg and Davis 2022), suggesting welfare gains that are possible from considering both decisions jointly and that would not require saving more today. Moreover, the employer setting for both sets of decisions, in an era of increasing employer attention to financial wellness, provide opportunities to offer simple strategies or defaults that work in concert.[3] This research also outlines an analytical framework to guide the recent policy focus on emergency savings accounts (Beshears et al. 2019), which may particularly suit individuals who experience both difficulty committing to a savings plan and also genuine liquidity needs.

We build a life-cycle model of saving decisions that incorporates three accounts with different tax and liquidity characteristics: an HSA, a tax-deferred DC account, and a liquid after-tax account. We assume exponential discounting, constant relative risk aversion utility, and exogenous retirement. We incorporate medical spending shocks that depend on health status, gender, and age and track the stock of out-of-pocket medical expenses, which may be reimbursed out of the HSA at any time. The model includes a second shock to consumption that generates a motive for liquidity to finance risks unrelated to health care.

In the absence of HSAs, liquidity constraints resulting from not only medical spending shocks but also uninsurable consumption shocks introduce a critical trade-off between saving in tax-preferred illiquid DC accounts versus after-tax liquid accounts. The role of an HSA as an "in–between" option that increases the value of DC saving can be demonstrated with a simple example. Consider a person who over 10 years has incurred $10,000 in out-of-pocket health care costs that they did not pay immediately out of their HSA. If the person must suddenly pay $10,000 for an expense not related to health care, they can reimburse themselves out of their HSA to get the needed cash (and in fact can bear an even larger expense if they have invested their HSA funds and accumulated interest tax-free). In comparison, accessing funds from DC accounts is difficult and potentially costly. For example, many plans allow loans, which must be paid back or else are subject to both taxes and a penalty, and more commonly, individuals withdraw assets from IRAs or 401(k) accounts from previous employers, again incurring taxes and a penalty.[4] HSAs can therefore can serve as a retirement savings vehicle, while also providing an option value for liquidity shocks that arise while working.

The complexity of the model set-up necessitates the use of machine-learning methods

---

[3]Of course, an important subset of employees lack access to one or both types of employer benefits. Our analysis applies similarly to Individual Retirement Accounts (IRAs) and to the increasing range of choices that individuals face even in government provided health care, since the Affordable Care Act exchanges, Medicare, and even in some cases Medicaid now offer plan choice and managed care options.

[4]The loan repayment terms on employer DC accounts may be quite favorable, yet anecdotal evidence indicates that few employees take up this option.

for dynamic programming solutions. Including multiple accounts and tracking out-of-pocket medical spending as well gives rise to a large number of state variables. And further, with occasionally-binding constraints resulting from differing contribution limits and withdrawal provisions, decision rules in conventional numerical solution methods are highly nonlinear, necessitating an increasingly large number of points and a flexible approximation method to capture them accurately. Both features make conventional techniques infeasible in this case. In contrast, machine learning methods use a very flexible approximation form (a neural network) and a scattered set of training points to approximate the decision rules, and so do not suffer (at least not meaningfully) from the curse of dimensionality (Hornik, Stinchcombe and White 1989, Fernández-Villaverde, Hurtado and Nuño 2019, Maliar, Maliar and Winant 2021, Duarte et al. 2022, Azinovic, Gaegauf and Scheidegger forthcoming).

Our model determines how the optimal net saving rate in the taxable liquid account and the optimal withdrawal strategy from the HSA each year vary as we alter the contribution rates to the HSA and illiquid account. We then search for which combination of contribution rates to the HSA and illiquid retirement account yields the highest lifetime utility. In considering fixed contribution rates to the tax-advantaged accounts, our approach suggests simple strategies for HSA saving in combination with illiquid retirement accounts. It is well known that many individuals apply rules of thumb, seek guidance from multiple sources, and respond to features that should not matter, like defaults (Thaler and Sunstein 2008), while failing to respond to features that should, like employer contributions (Friedberg, Leive and Cai 2020) and employer matches (Bubb and Warren 2020).

Our model delivers several findings on the links between HSAs and DC accounts. First, HSAs raise optimal tax-preferred saving, compared to a system with only tax-preferred illiquid accounts and taxable liquid saving. Second, at low levels of contribution rates, which are what is commonly observed in practice, HSAs and illiquid accounts are complements rather than substitutes. Third, high earners benefit more than low earners from HSAs, with the former benefiting primarily from tax preferences and the latter primarily from liquidity insurance. By unifying the analysis of both health insurance and retirement plan choices—often viewed as unrelated—we demonstrate the dual life-cycle savings possibilities and liquidity insurance features of HSAs.

The paper proceeds as follows. Section 2 describes the key features of HSAs relative to other savings vehicles. Section 3 presents the model and section 4 details the solution methods based on machine learning techniques. We present our key results in Section 5 and briefly conclude in Section 6.

2   Institutional Background

Among employers that offer pensions and health insurance, defined contribution (DC) plans and high deductible health plans (HDHPs) have become the norm. DC plans have dominated retirement plan offerings for the last twenty or more years.[5] HDHPs are more recent, however; the inclusion of HSAs in HDHPs was first enabled in 2003 by the Medicare Modernization Act, and now over 1 in 5 workers are enrolled in these plans (Claxton et al. 2021). Their obvious similarity as retirement assets, detailed below, means that DC plans and HSAs place a burden on individuals to make optimal choices regarding contribution amounts, portfolio allocations, and withdrawals relative to prior versions of retirement and health insurance plans.

2.1   Defined Contribution Retirement Plans

DC plans like 401(k) accounts face rules affecting contributions, accumulations, and withdrawals. While our model focuses primarily on the optimal contribution decision, withdrawal constraints are the key feature that makes a 401(k) account a retirement asset. 401(k) assets only become liquid after age $59\frac{1}{2}$, and in many cases only after leaving one's employer as well. At that point, assets may be withdrawn without penalty, and after age 72, assets must be withdrawn gradually in the form of Required Minimum Distributions. Before then, assets are only accessible through a loan against one's asset balances (if the employer allows such loans), or after leaving an employer, in which case withdrawals incur a 10% penalty.[6] In all cases, whether before or after age $59\frac{1}{2}$, withdrawals are fully income-taxable. Since the Pension Protection Act of 2006, some employers have begun to offer Roth 401(k) plans, in which post-tax assets are placed in the plan and accumulate tax-free afterward. Contribution limits apply to the sum of tax-deferred and Roth plans, and for our purposes, these are not different enough to model as distinct.

The contribution and accumulation subsidies for 401(k) accounts that are written into the tax code are designed as the "carrot" that makes plans attractive in spite of their illiquidity. Individuals can make contributions to 401(k) plans up to an annual limit, which reached $20,500 in 2022.[7] Many employers also make contributions into 401(k) accounts,

---

[5]Since the decline of defined benefit (DB) plans (Friedberg and Owyang 2002), the value of assets in private sector DC plans increased from $74 billion in 1975 to almost $10 trillion in 2021 (Holden, Schrass and Barone Chism 2022).

[6]One exception is 457 accounts, which do not permit loans while working but allow employees to withdraw funds after separation from their employer. These accounts are only available for non-profits or state employees.

[7]Further provisions, such as catch-up contributions after age 50, apply less widely and are not incorporated into our model.

often in matching form.

Tax preferences take one of two forms. Conventional 401(k) accounts are income tax-deferred; contributions are tax-deductible and returns on assets accumulate tax-free. All withdrawals are taxed as personal income. On the other hand, contributions to Roth 401(k) accounts, enabled by the IRS in 2006, are taxable upfront, while all returns on assets and later withdrawals are non-taxable.[8] In our life-cycle model, we incorporate a tax-deferred DC account.

Notably, the same considerations apply to Individual Retirement Accounts (IRAs), with both tax-deferred and Roth accounts available. In some ways, IRAs are less constrained than are DC accounts, since individuals do not need to rely on an employer choosing to offer one. In other ways, they are more limited. Contributions to IRAs are subject to stricter limits than to 401(k)s, while household income is also subject to a restriction (which is tighter for tax-deferred than for Roth accounts) in order for contributions to be tax-preferred. Nevertheless, tax deferral of returns is available even for contributions that do not qualify for upfront deductibility.

2.2   Health Savings Accounts

HSAs offer powerful tax advantages to finance health care expenses. Contributions are deductible from all taxable earnings (including not just income taxes but also FICA taxes), interest grows tax free, and withdrawals for qualified health expenses are tax-free. Qualified expenses, which must be incurred after the HSA has been established, include out-of-pocket payments associated with a HDHP while working, Medicare premiums and out-of-pocket payments, long term care insurance, and long term care costs. One important feature of HSAs is that balances can also be used to reimburse oneself for past qualified expenses, no matter how long ago they occurred.

If not used for health expenses, HSAs offer tax benefits similar to tax-deferred retirement accounts. Withdrawals for non-medical consumption are subject to income taxation and a 20% penalty if before age 65 (which is stricter than the 10% penalty on pre-age $59\frac{1}{2}$ withdrawals). We will assume that HSAs are only used to finance health expenses, rather than other consumption. All HSA contributions roll over each year. This feature distinguishes them from Flexible Spending Accounts (FSAs), in which at most $500 can be

---

[8]Tax-deferred and Roth accounts yield identical returns if an individual faces the same marginal tax rate when contributions and withdrawals are made. The accumulation of assets on a tax-deferred basis confers an advantage for tax-preferred accounts compared to taxable savings accounts, even if an individual faces the same marginal tax rate in retirement as earlier. Tax-deferred DC assets may also benefit by being taxed at a lower rate after individuals retire and no longer have earned income, though the anticipation of future tax rate increases the advantage of a Roth account.

carried over to the following year.

3 Model

We build a life-cycle model of savings decisions that incorporates key features of the employee benefit choice setting and determines optimal contribution rates across accounts. Our model incorporates three accounts with different tax and liquidity characteristics: an HSA, a tax-deferred DC account which becomes liquid in retirement, and a liquid after-tax account. We incorporate medical spending risk, which depends on health status, gender, and age, and we track the stock of out-of-pocket medical expenses, which may be reimbursed out of the HSA at any time. We also include a consumption shock that is unrelated to health care. Both shocks generate a need for liquidity which reduces the value of illiquid retirement saving.

This section describes the specification of the model. In the next section, we discuss implementation and solution methods. With multiple state variables and shocks (three accounts, the stock of unreimbursed medical expenses, health status, the health shock, and the consumption shock), we solve the model using neural networks rather than standard dynamic programming techniques.

3.1 Preferences and budget constraints

We assume individuals have the per-period utility function $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$, where c denotes consumption and σ represents the coefficient of relative risk aversion. Utility across periods is separable, and individuals have a time preference rate of δ.

Individuals choose how much to save in each of three types of accounts. Each savings account involves a different set of constraints. Limits on contributions to the DC account and HSA are based on age. Differential tax treatment of contributions and returns will be incorporated into the full budget constraint that we introduce after we describe each account.

3.1.1 Taxable liquid account

Assets a_1 in the conventional savings account grow at the gross rate R_1 and evolve according to:

$$a_1' = R_1 \left(a_1 + s_1 - w_1 \right) \tag{1}$$

where s_1 are taxable account contributions, w_1 are withdrawals, and a_1' are next period's liquid assets. Borrowing is not allowed, while withdrawals are completely flexible in any

period up to the amount of available assets:

$$w_1 \leq a_1 \tag{2}$$

Since both s_1 and w_1 will not be positive at the same time, the above constraint is equivalent to imposing $a_1' \geq 0$. We will consider the impact of varying initial liquid wealth $a_{1,t=0}$.

3.1.2 Retirement accounts

Assets a_2 in the DC retirement account grow at the gross rate R_2 and evolve according to:

$$a_2' = R_2 (a_2 + s_2) \tag{3}$$

where s_2 denotes retirement contributions. This account is totally illiquid until age $59\frac{1}{2}$ but has tax advantages. We model these as traditional tax-deferral—contributions are excluded from income taxation, interest is tax-deferred, and withdrawals are taxable. We assume that people annuitize the value of this account upon retirement, so we do not currently model withdrawal choices from this account during retirement.[9]

3.1.3 Health Savings Accounts

Assets a_3 in the HSA grow at the gross rate R_3 and evolve according to:

$$a_3' = R_3 (a_3 + s_3 - w_3) \tag{4}$$

where HSA contributions s_3 may include both employee and employer contributions. Withdrawals w_3 from HSAs can be made up to the sum of current health care expenses m and accumulated past expenses z that have yet to be reimbursed from the HSA:

$$w_3 \leq m + z \tag{5}$$

where the amount of unreimbursed past health expenses evolves according to:

$$z' = m + z - w_3 \tag{6}$$

---

[9]Since many employees do not have access to actuarially fair annuities, eliminating annuitization from the model will increase the value of retirement saving relative to maintaining liquidity while working.

3.1.4 Income and Budget Constraints

While working, consumption and saving satisfy:

$$c + s_1 + s_2 + s_3 + m + e \quad \leq \quad y + R_1 a_1 - T\left(y - s_2 - s_3, (R_1 - 1)\, a_1\right) + w_3 \qquad (7)$$

where e is a consumption shock and $T(\cdot)$ is a progressive tax function that delivers income taxes paid as a function of labor earnings (after deducting HSA and retirement account contributions) and returns on the taxable (but not DC or HSA) account.[10] We interpret health expenses as spending shocks that must be financed, rather than part of consumption c that enters utility. Individuals are always employed until they exogenously retire at 65, and labor earnings grow deterministically, with $y' = gy$. We will consider the impact of varying y_0.

When retired, labor earnings are replaced with lower Social Security benefits b, and contributions to the DC account or HSA are no longer permitted. Defining the annuity payment from the illiquid account as p, the budget constraint is then:

$$c + s_1 + m + e \quad \leq \quad b + p + R_1 a_1 - T_{old}\left(b, (R_1 - 1)\, a_1, p\right) + w_3 \qquad (8)$$

where the tax function while old $T_{old}(\cdot)$ may differ from that while working to account for the absence of payroll taxation when retired and the partial taxation of Social Security benefits.

3.2 Sources of Uncertainty

The model currently allows for uncertainty in health spending which in turn depend on uncertain health status (which does not itself affect anything else, such as work capacity, in the model), and an uninsurable consumption shock that is unrelated to health. This not only generates a self-insurance motive for saving against liquidity shocks, but it specifically advantages the HSA as a form of liquidity insurance. We treat health spending shocks as related to health status but not influenced by resources on hand. Details of the calibration of health spending by age, health status, and the consumption shock are presented in Appendix A.

3.2.1 Health status uncertainty

Health status h helps us calibrate health spending shocks and is classified as either "healthy" $(h = 1)$ or "sick" $(h = 0)$. The probability of being healthy in year t depends on age j, gender

---

[10]The current formulation does not include the extra deductibility of HSA contributions from FICA taxation.

g, and health status in year $t-1$:

$$\pi_{jgt} = Pr(h_t = 1 | h_{t-1}, j, g) \tag{9}$$

We use self-reported health from the Medical Expenditure Panel Survey (MEPS) and Health & Retirement Survey (HRS) to calculate these transition probabilities. Respondents who report self-assessed health as "Excellent," "Very good," or "Good" are classified as healthy. Those reporting "Fair" or "Poor" health status are classified as sick. We smooth the empirical probabilities calculated from the MEPS and HRS by fitting a probit with a cubic function of age and an indicator for the HRS.

3.2.2 Health spending shocks

We split health spending into the periods before and after retirement. To obtain moments of health care spending shocks, we use both out-of-pocket and total health spending the Medical Expenditure Panel Survey (MEPS) pre-retirement and the Health and Retirement Study (HRS) post-retirement. The HRS has two advantages relative to the MEPS for measuring spending at older ages. First, it includes Medicare premiums, which are qualified expenses for HSAs, and are only payable beginning at age 65. Second, it surveys residents of nursing homes (who are excluded from the MEPS), and we include nursing home care costs as a form of health spending, since it is exceedingly costly at older ages (De Nardi, Pashchenko and Porapakkarm (2017)) and reimbursable out of HSAs.

Our approach for dealing with skewness in the health spending distribution is a modified procedure of Pashchenko and Porapakkarm (2017) and De Nardi, Pashchenko and Porapakkarm (2017). For each age and health status, we approximate the spending shock by assuming it takes one of three values ("low", "medium", or "high"). We classify low spending as below median spending for that age and health status. Medium spending is between the 50th and 95th percentiles, and high spending as above the 95th percentile. After classifying survey respondents into these three groups, we fit a 4th-order polynomial in age to smooth the empirical distribution.

After retirement, we assume that individuals switch from a high-deductible health plan to Medicare coverage (as almost everyone who has left employment does). We take the sum of out-of-pocket spending and Medicare premiums as our measure of spending that can be financed from the HSA.

Prior to retirement, we construct out-of-pocket spending in two steps. We first predict total health spending (including both out-of-pocket and insurer payments) by age, health status, and spending group as described above. We then apply a non-linear function that

maps total health spending to out-of-pocket payments that can be reimbursed from the HSA. This function captures the cost-sharing features of the high-deductible health plan. We assume that each year while working, the plan has a $2,500 deductible, 20% coinsurance rate, and $5,000 out-of-pocket maximum.[11]

3.2.3 Consumption shock

We incorporate an uninsurable consumption shock e that is unrelated to the health care shocks, and unlike the health care shock which simply reduces assets, this shock affects utility from consumption. Our current formulation is highly stylized and intended to capture typical moderately-sized shocks, such as a car repair or house repair. We simply assume that individuals incur a $2,000 loss with a fixed probability each year that does not depend on age or income. We use data on house and vehicle repairs from the Panel Study of Income Dynamics to parameterize the shock's probability, which indicates a 40 percent probability of incurring a loss of this size in any given year. We chose to focus on a random $2,000 shock because it is a similar magnitude to health care costs while working, and prior survey research in household finance has focused on emergency expenses of this size to classify households as financially "fragile" (Lusardi et al. 2011, Clark, Lusardi and Mitchell 2021).

3.3 Recursive formulation

As is typical in life-cycle savings problems, we represent the dynamic programming problem recursively. We split the problem into the years when the person is working versus retired. The decision problem while working is given by:

$$
v_j\left(a_1, a_2, a_3, m, z, h, e\right) = \max_{c, w_3, s_1, s_2, s_3}\left\{\frac{c^{1-\sigma}}{1-\sigma} + E\left[v_{j+1}\left(a_1', a_2', a_3', m', z', , h', e'\right)\right]\right\}
$$

subject to constraints (1)–(7) and the non-negativity constraints $a_2' \geq 0, a_3' \geq 0, z' \geq 0$. Individuals must choose how much to contribute to each of the three accounts, allowing negative HSA contributions (that is, withdrawals) to pay for health spending in case liquidity is of value.

Once individuals are retired, they can still save in the liquid account but no longer contribute to the retirement account or HSA. They choose how much to withdraw from the liquid account and HSA. The decision problem when retired is given by:

---

[11]These parameters are close to the averages of HDHPs offered by US employers according to the Kaiser Family Foundation's Employer Benefit Survey (Claxton et al. 2021).)

$$v_j\left(a_1, a_2, a_3, m, z, h, e\right) = \max_{c, w_3, s_1}\left\{\frac{c^{1-\sigma}}{1-\sigma} + E\left[v_{j+1}\left(a_1', a_2', a_3', m', z', h', e'\right)\right]\right\} + V_T$$

subject to constraints (1)–(6), (8), and the same non-negativity constraints. V^T denotes the scrap value of assets (as in a bequest function) in the terminal period.

4   Implementation and Solution Method

The reason that we use machine-learning solution methods owes to the complexity of the model set-up. With occasionally-binding constraints, decision rules in conventional numerical solution methods are highly nonlinear, necessitating an increasingly large number of points and a flexible approximation method to capture them accurately. We also have many state variables due to the multiple accounts and the shocks. The curse of dimensionality implies that the number of points in a rectangular grid rises exponentially in the dimension of the state space, implying that with n points in each direction we would have n^7 points at which we would need to solve the model, which is infeasible for any value of n sufficient to accurately capture the behavior of agents.

Machine learning methods use a very flexible approximation form (a neural network) and a scattered set of training points to approximate the decision rules, and so do not suffer (at least not meaningfully) from the curse of dimensionality (Hornik, Stinchcombe and White 1989). In particular, neural nets predict outputs by applying nonlinear transformations to linear combinations of inputs. The weights of the linear combination are chosen by an algorithm that minimizes the difference between the outputs and the predicted outputs based on the weights. Recently, the economics literature has begun to utilize these methods to solve complex dynamic economic models (Fernández-Villaverde, Hurtado and Nuño 2019, Maliar, Maliar and Winant 2021, Duarte et al. 2022, Azinovic, Gaegauf and Scheidegger forthcoming).

We currently do not model each year as its own period due to computational costs, but we maintain roughly the same fraction of working and retirement periods as we observe for most individuals. Thus, the model currently includes 12 periods that correspond to roughly five years each, with 8 working periods and 4 retirement periods. We expect qualitatively similar results as if we had a larger number of periods that corresponded to one year each. We adjust the discount factor and growth rates accordingly so they correspond to realistic annualized rates.

At this point, we reduce the number of variables over which we optimize. The model

fixes the contribution rates to the illiquid account and HSA over the working years, and computes optimal net saving in the taxable liquid account each period and the optimal withdrawal strategy from the HSA each period. Fixing the contribution rate across all years while working also represents a simple decision rule, and then searching for which combination of these rates yields the highest utility can be seen as a "rule of thumb" for how much to save.

Our model considers the illiquid retirement plan as a tax-deferred account.[12] We assume that the illiquid account and the HSA have the same (certain) return equal to 5%, which is justifiable since HSA assets can be invested. Since the liquid account is more easily accessible, we assume it earns a slightly lower return of 4% return each period.

We assume real wage growth of 3 percent each period while working. Wages and withdrawals from the taxable account are subject to taxation. We specify the tax function as $T(y) = y - \lambda y^{1-\tau}$, where λ governs the level of taxation and τ determines its progressivity. Following Heathcote, Storesletten and Violante (2017), we assume $\tau = 0.18$ and $\lambda = 5.5$. To determine Social Security benefits in our model, we apply the standard primary insurance amount formula to average indexed monthly earnings. We incorporate the rules involving taxation of Social Security benefits based on the sum of the benefit and the annuity payment.[13] Table 1 lists the parameter values and other model inputs.

We build 11 neural networks—one for each of the non-terminal periods—that output the decision rules that period based on the state variables. The rules outputted from the network are those that minimize the sum of the squared errors from the Euler equations. We use ADAM (Adaptive Moment Estimation) as the algorithm for each network to learn the solution. ADAM is similar to stochastic gradient descent but is made faster by also using a function of prior squared gradients, capturing momentum; the momentum term keeps the optimizer from being diverted into narrow valleys that are only locally improving. Each network includes two hidden layers with 32 nodes each that are activated using the rectified linear activation unit (RELU). The networks use 150 training points. We experimented with different numbers of training points, activation functions, nodes, layers, and epochs for each network to assess sensitivity.

Since calculating lifetime utility for every possible combination of contribution rates is computationally intensive, we solve the model for 50 different combinations of contribution rates for each account and then fit a surface using the model's output at these points.

---

[12] Results would not differ qualitatively with a Roth account, which is less commonly used by employees

[13] The percentage of Social Security benefits that are taxed depends on "combined income", which is half of the Social Security benefit plus income from other sources. Social Security benefits are excluded from taxable income if combined income is below $25,000, half of the benefit is taxable if combined income is between $25,000 and $34,000, and 85 percent is taxable if combined income exceeds $34,000.

Table 1: Parameter Values

Parameter	Value
σ, Coefficient of relative risk aversion	2
δ, Discount factor	0.9
R_1, Gross return on liquid account	1.04
R_2, Gross return on illliquid account	1.05
R_3, Gross return on HSA	1.05
g, Wage growth	1.03
τ, Progressivity of tax function	0.18
λ, Level of tax function	5.5
π_{jgt}, Probability of being healthy	Appendix
Probability of low health expenditure shock	0.5
Probability of medium health expenditure shock	0.45
Probability of high health expenditure shock	0.05
Size of low health expenditure shock	Appendix
Size of low medium expenditure shock	Appendix
Size of low high expenditure shock	Appendix
Probability of consumption shock	0.4
Size of consumption shock	$2,000

5 Results

5.1 Baseline approach

We illustrate the findings from the model by focusing on a small set of individuals who are typical of the range of initial salary and endowment levels in an employer setting:

- Two initial salary levels: $65,000 or $105,000.[14]

- Two initial asset amounts in liquid account: $0 or $10,000.

We also split results separately by men and women because health spending risk varies by gender.

Using the model results, we calculate how different combinations of contribution rates to the illiquid account and HSA affect lifetime utility. We compare the impact of only having access to one or the other type of tax-preferred account, standardizing utility by calculating the percentage increase compared to the case when saving is only possible in the taxable

---

[14]The lower salary of $65,000 corresponds to approximately the median (and below the average) from the large University setting studied in Leive, Friedberg and Davis (2022), while the higher salary of $105,000 corresponds to the 75th percentile.

liquid account (and people do so optimally).[15] As described earlier, the model fixes the illiquid and HSA contribution rates over the working years, and computes the optimal net saving in the taxable liquid account and withdrawal strategy from the HSA each period. We focus on constant contribution rates to employer accounts across time because this behavior is typical for many individuals and yields simple strategies that raise lifetime utility and may apply in many situations. We consider contributions in the HSA up to 8% of salary for higher-salaried employees, since 8% is slightly above the contribution limit for high earners with family coverage, and 10% for lower-salaried employees. We consider higher contributions in the illiquid account since the limits are greater.

We summarize the model results graphically by plotting the gain (or loss) in utility for each combination of illiquid and HSA contribution rate that we consider, relative to saving only in the liquid account. The vertical axis denotes the change in lifetime utility for various combinations of contribution rates to the illiquid account and HSA. Each blue dot in represents the model's calculation of lifetime utility (relative to only saving in liquid taxable accounts) for that particular combination of contribution rates while working. We run the model for 50 different combinations of contribution rates to the illiquid retirement account and the HSA, and then use the collection of points to fit the three-dimensional surface using a second-order polynomial. For reference, the light-blue shaded plane denotes no change in lifetime utility relative to liquid saving only.

The main results appear in Figure 1, with separate plots by gender and salary level for employees with no initial assets. The plots are qualitatively similar for employees endowed with $10,000 in the taxable liquid account (Appendix Figure B.1).

5.2   Results with one tax-deferred account

Before discussing both accounts together, it is useful first to consider the results when the HSA contribution is set to zero and the contribution to the illiquid retirement account varies. In Figure 1, this can be seen by reading along the leftward axis of the three-dimensional plots. As expected, utility is a concave function of illiquid account contributions. For higher earners, increasing contributions to the illiquid account raises utility up until around 10 percent of salary, and then the gains begin to decline. For lower earners, the gains top out at much lower contribution rates. For women, the optimal illiquid contribution would be 2 percent if the only other option were taxable liquid saving. For men at the lower salary, requiring any contribution rate to the illiquid account when an HSA is not available lowers utility in our model, because the tax benefits are outweighed by reduced liquidity and delayed

---

[15]The level of the value function in the first period represents lifetime utility.
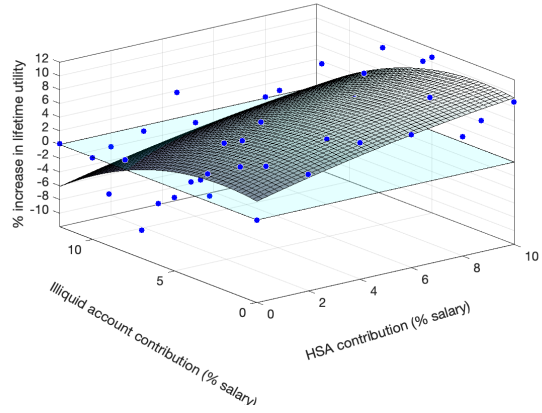
consumption.[16]

If there were no illiquid retirement account (reading along the rightward axis of the plots), then increasing HSA contributions raises utility up to the limits we set (10% for lower earners and 8% for higher earners).

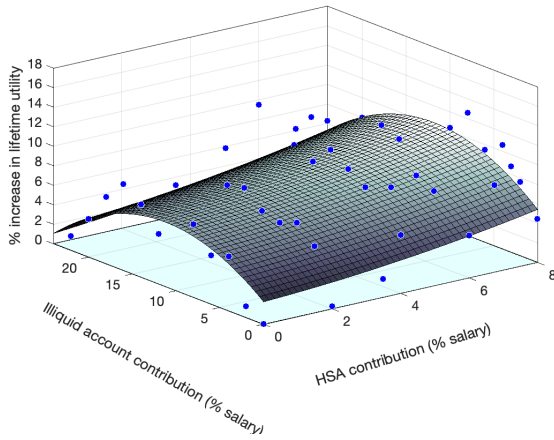Figure 1: Increase in Lifetime Utility Relative to Liquid Saving Only
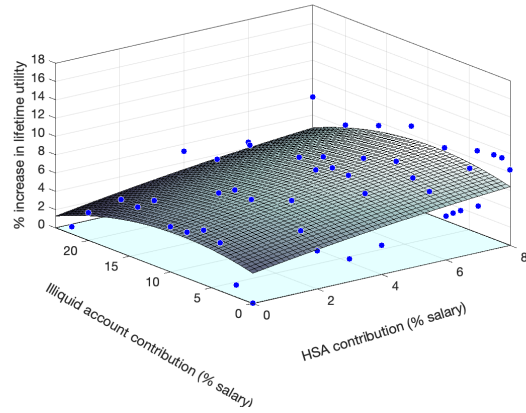
(a) $65,000 salary, men
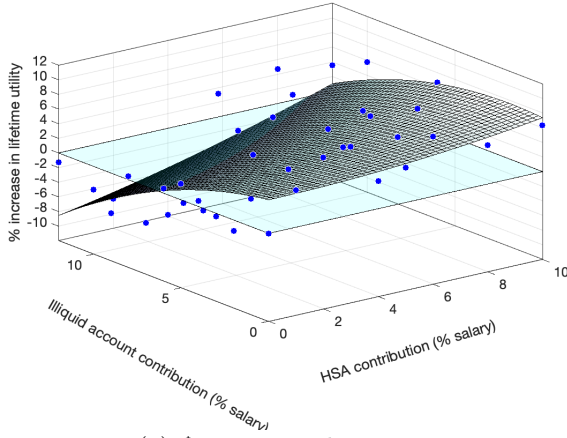
(b) $65,000 salary, women

(c) $105,000 salary, men

(d) $105,000 salary, women

Notes: Figures plot the percentage change in lifetime utility for different combinations of HSA and illiquid retirement account contribution rates as a fraction of annual salary. The increase is measured relative to optimally saving in the taxable liquid account only. Each panel plots the results for a person with different salary profiles and gender. Initial endowments of assets in the taxable account are set to zero. The surface plot is fitted using a 3rd-order polynomial calculated using the output of the model from 50 different combinations of contribution rates. Shaded plane denotes zero lifetime utility relative to liquid saving only.

---

[16]In our model, differences by gender in health spending over the life-cycle is the only factor leading to this difference.

5.3   Results with both tax-deferred accounts

Now we consider simultaneous increases in both illiquid account and HSA contributions. Whether HSAs and illiquid retirement accounts are complements or substitutes can be assessed by signing the gradient of these surfaces at various combinations of saving—they are complementary if utility increases as contributions to both accounts increase by a small amount. Interestingly, the surfaces demonstrate that the accounts are complements (the gradient is positive) for employees at low HSA contribution levels (which mirror what is often observed in practice) and substitutes (the gradient is negative) at high contribution levels. Starting from low saving rates, utility increases as contributions to both accounts simultaneously increase. Beyond a point, however, utility decreases as contributions to both accounts rise. So, if employees are hesitant to commit substantial funds to their HSAs (consistent with current evidence), then contributing to the HSA at least a little increases the value of contributing to their DC account.

What is the intuition for complementarity between the HSA and illiquid accounts? HSAs allow people to lock up more saving in illiquid form because HSAs provide liquidity to finance health spending. Without the HSA, more saving would have gone to the liquid account instead. This liquidity is quite valuable for those earning lower salaries, and notably it is only available if employees defer use of their HSA to pay for out-of-pocket medical costs—an action that very few currently take (Leive 2022). More broadly, these results point to the potential value of a more formal system of emergency savings (Beshears et al. 2019).

The combination of contribution rates that may be optimal in the current set-up of our model varies by income. It is evident in Figure 1 that lower earners begin to lose value from illiquid saving as their illiquid saving rates increase beyond a low amount, while high earners—especially men—continue to gain over a considerable range. Table 2 presents the combination of illiquid account and HSA saving that maximizes utility for each of the four groups.[17] To gauge how the HSA changes decisions relative to the status quo, we also present the resulting contribution rates in the illiquid account if HSAs are unavailable.[18] For each group, preferred illiquid saving stays the same or rises with the addition of HSAs. The result is that total tax-advantaged saving increases substantially with the addition of HSAs.

We can assess the benefits of HSAs by comparing the relative gains from the utility-maximizing contribution rule with HSAs (in the final column) to the

---

[17]Appendix Table B.1 presents the corresponding table when individuals are endowed with $10,000 in initial assets, which is similar.

[18]This calculation assumes that health care expenses do not change without the HSA—amounting to a counterfactual in which the person still has an HDHP but not an HSA, which is common in the individual market but less so in the employer market.

utility-maximizing contribution rule when only taxable and illiquid saving is available (in the fourth column). In our model, an illiquid account raises utility by small amounts, if at all, for low earners. Allowing people to also save in an HSA raises lifetime utility and increases tax-preferred saving rates overall. For low earners, the total contribution rates to tax-deferred accounts increase from zero or small amounts to about 12% of salary. The increases in utility are sizable, particularly for women due to higher expected health spending. For high earners, adding the HSA nearly doubles the total amount of tax-preferred saving. The relative increase in lifetime utility from adding the HSA is larger than for low earners, partly due to the tax benefits.

Table 2: Optimal Fixed Contribution Rates and Lifetime Utility

Salary level (initial)	Gender	without HSA		with HSA		
		optimal illiquid saving	% utility increase vs. liquid saving only	optimal illiquid saving	optimal HSA saving	% utility increase vs. liquid saving only
$65,000	Men	0	0.0	2	10	3.0
$65,000	Women	2	0.4	2	10	7.0
$105,000	Men	11	4.3	12	8	8.0
$105,000	Women	9	1.1	9	8	4.3

Note: Table presents the set of fixed contribution rates that maximize lifetime utility from the model. The points represent the maxima from the surfaces shown in Figure 1. For comparison, the table shows the rates that maximize utility when saving can only occur in the illiquid and liquid accounts but not the HSA, in addition to the case when all three accounts are available.

Figure 2 simulates the path of total assets (solid line) and HSA assets (dashed line) over time for each of the four sets of fixed contribution rates presented in Table 2.[19] HSA assets peak at the start of retirement for each group, at the same time that the other assets—which are larger by comparison—also peak. HSAs comprise a substantially larger share of total savings for lower earners than higher earners, since absolute amounts of HSA balances are similar between groups.

---

[19]To construct these figures, we simulate 10,000 paths for each group, and average saving and withdrawal rates over the optimal saving and withdrawal strategies of these paths.

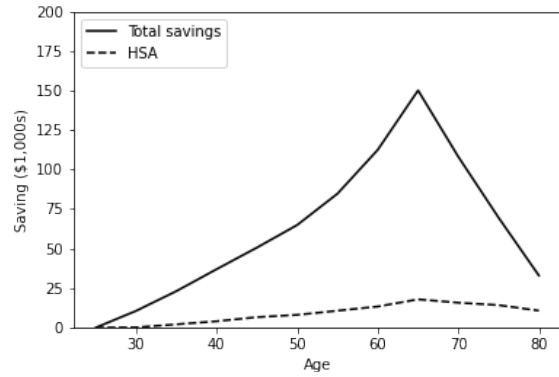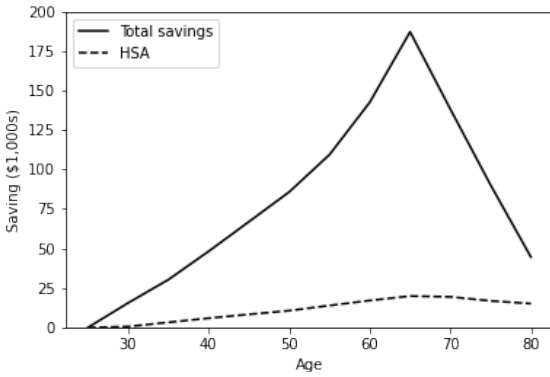Figure 2: Illustrative Paths of Savings

(a) $65,000 salary, men

(b) $65,000 salary, women
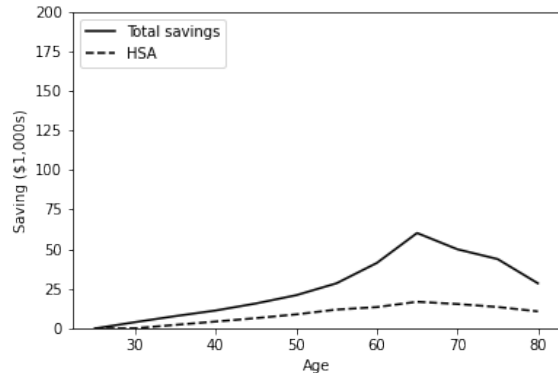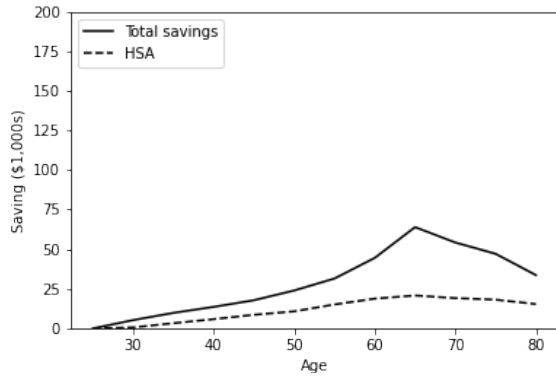
(c) $105,000 salary, men

(d) $105,000 salary, women

Notes: Figures plot the path of total savings and HSA savings over the life-cycle for the optimal contribution rates from Table 2. Total savings include balances in the taxable liquid account, illiquid account, and HSA. Graphs are split by gender and salary in first period. For all groups, wage growth is set at 3 percent between each 5-year period while working.

6   Discussion

We have developed a model of life-cycle saving with multiple accounts to explore optimal saving and withdrawal decisions from HSAs, tax-preferred defined contribution accounts, and taxable liquid accounts. Our focus has been on the level and mix of contributions in the HSA, which offers features of a retirement account but with access to some liquidity, versus the illiquid retirement account. In fact, HSAs would have a strong resemblance to emergency savings accounts—if employees defer use of their HSA to pay for out-of-pocket medical costs, an action that very few currently take (Leive 2022).

The results from our model indicate that the addition of HSAs to DC accounts should raise total tax-preferred saving *if people are optimizing in their saving decisions*. Consistent with empirical patterns observed in tax data (Helmchen et al. 2015), the benefits of HSAs in our model are larger for higher earners due to their higher marginal tax rate and greater levels of saving. HSAs nevertheless offer value to lower earners, providing liquidity insurance that is otherwise difficult to access in current retirement saving vehicles. For both high and low earners, low levels of HSA contributions (consistent with what is observed in practice) raise the value of illiquid DC plans and IRAs.

References

Aaron, Henry, Patrick Healy, and Surachai Khitatrakun. 2008. "What's in a Name? Are Health Savings Accounts Really Health Savings Accounts?" In *Using Taxes To Reform Health Insurance*. 92–118. Washington, DC:Brookings Institution.

Azinovic, Marlon, Luca Gaegauf, and Simon Scheidegger. forthcoming. "Deep Equilibrium Nets." *International Economic Review.*

Beshears, John, James Choi, Mark Iwry, David Laibson, and Brigitte Madrian. 2019. "Building Emergency Savings Through Employer-Sponsored Rainy-day Savings Accounts." *NBER Working Paper 26498.*

Bubb, Ryan, and Patrick Warren. 2020. "An Equilibrium Theory of Retirement Plan Design." *American Economic Journal: Economic Policy*, 12(2): 22–45.

Clark, Robert, Annamaria Lusardi, and Olivia Mitchell. 2021. "Financial Fragility during the COVID-19 Pandemic." *American Economic Association: Papers and Proceedings*, 111: 292–296.

Claxton, Gary, Matthew Rae, Gregory Young, Nisha Kurani, Heidi Whitmore, Jason Kerns, Jackie Cifuentes, Greg Shmavonian, and Anthony Damico. 2021. "2021 Employer Health Benefits Surveys." *Kaiser Family Foundation.*

Davis, Brent, Adam Leive, and Andrew Gellert. 2022. "How Do Employees Evaluate Workplace Benefits? Evidence from Health Savings Accounts."

De Nardi, Mariacristina, Svetlana Pashchenko, and Ponpoje Porapakkarm. 2017. "The Lifetime Cost of Bad Health." *NBER Working Paper 23963.*

Duarte, Victor, Julia Fonseca, Aaron Goodman, and Jonathan Parker. 2022. "Simple Allocation Rules and Optimal Portfolio Choice Over the Lifecycle." *NBER Working Paper 29559.*

Ericson, Keith, and Justin Sydnor. 2018. "Liquidity Constraints and the Value of Insurance." *NBER Working Paper 24993.*

Fernández-Villaverde, Jesús, Samuel Hurtado, and Galo Nuño. 2019. "Financial Frictions and the Wealth Distribution." *NBER Working Paper 26302.*

Friedberg, Leora, Adam Leive, and Wenqiang Cai. 2020. "Balancing Commitment and Liquidity: Empirical Evidence from Mandatory Retirement Saving."

Friedberg, Leora, and Michael Owyang. 2002. "The Rise of 401(k) and Other Defined Contribution Plans." *Federal Reserve Bank of St. Louis Review*, 84(1): 23–34.

Fronstin, Paul, and Jake Spiegel. 2021. "Trends in Trends in Health Savings Account Balances, Contributions, Distributions, and Investments and the Impact of COVID-19." Employee Benefits Research Institute 538, Washington, DC.

Handel, Ben, and Jonathan Kolstad. 2015. "Health Insurance for Humans: Information Frictions, Plan Choice, and Consumer Welfare." *American Economic Review*, 105(8): 2449–2500.

Handel, Benjamin. 2013. "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts." *American Economic Review*, 103(7): 2643–82.

Heathcote, J., K Storesletten, and G Violante. 2017. "Optimal tax progressivity: An analytical framework." *Quarterly Journal of Economics*, 132(4): 1693–1754.

Helmchen, Lorens, David Brown, Ithai Lurie, and Anthony Lo Sasso. 2015. "Health Savings Accounts: Growth Concentrated Among High-Income Households and Large Employers." *Health Affairs*, 34(9): 1594–1598.

Holden, Sarah, Daniel Schrass, and Elena Barone Chism. 2022. "Defined Contribution Plan Participants' Activities,2021." Investment Company Institute April.

Hornik, Kurt, Maxwell Stinchcombe, and Halbert White. 1989. "Multi-Layer Feedforward Networks Are Universal Approximators." *Neural Networks*, 5(2): 359–366.

Leive, Adam. 2022. "Health Insurance Design Meets Saving Incentives: Consumer Responses to Complex Contracts." *American Economic Journal: Applied Economics*, 14(2).

Leive, Adam, Leora Friedberg, and Brent Davis. 2022. "Overpaying and Undersaving? Correlated Mistakes in Health Insurance and Retirement Saving."

Liu, Chenyuan, and Justin Sydnor. 2022. "Dominated Options in Health Insurance Plans." *American Economic Journal: Economic Policy*, 14(1): 277–300.

Lusardi, Annamaria, Daniel Schneider, Peter Tufano, Adair Morse, and Karen M Pence. 2011. "Financially Fragile Households: Evidence and Implications." *Brookings Papers on Economic Activity*, 83–150.

Maliar, Lilia, Serguei Maliar, and Pablo Winant. 2021. "Deep Learning for Solving Dynamic Economic Models." *Journal of Monetary Economics*, 122(2): 76–101.

Pashchenko, Svetlana, and Ponpoje Porapakkarm. 2017. "Work Incentives of Medicaid Beneficiaries and the Role of Asset Testing." *International Economic Review*, 58(3-4): 689–716.

Thaler, Richard, and Cass Sunstein. 2008. *Nudge: Improving Decisions about Health, Wealth, and Happiness.* Yale University Press.

Appendix A: Data and Calibration for Sources of Uncertainty

Health status: To calculate the transition probabilities between healthy and unhealthy states, we use self-reported health status reported in the Medical Expenditure Panel Survey (MEPS) and Health & Retirement Survey (HRS). We need to combine the datasets because each one alone provides an insufficient sample for our population of interest. The MEPS top-codes age at 85 and excludes residents of institutions (such as nursing homes). The HRS does not top-code age, but samples individuals over age 50 (the HRS includes individuals younger than 50 if they are the spouse of the respondent, but the sample size is small). We therefore pool the two surveys and adjust for fixed differences in health status across the surveys using data from ages that overlap. Respondents who report self-assessed health as "Excellent," "Very good," or "Good" are classified as healthy. Those reporting "Fair" or "Poor" health status are classified as sick. Probabilities are calculated from a probit regression with a cubic in age using the HRS, estimated separately by gender and current health status. Figure A.1 plots predicted probability of being healthy next year given age, gender, and current health status.

Figure A.1: Probability of Healthy Next Period given Age, Gender, and Current Health Status
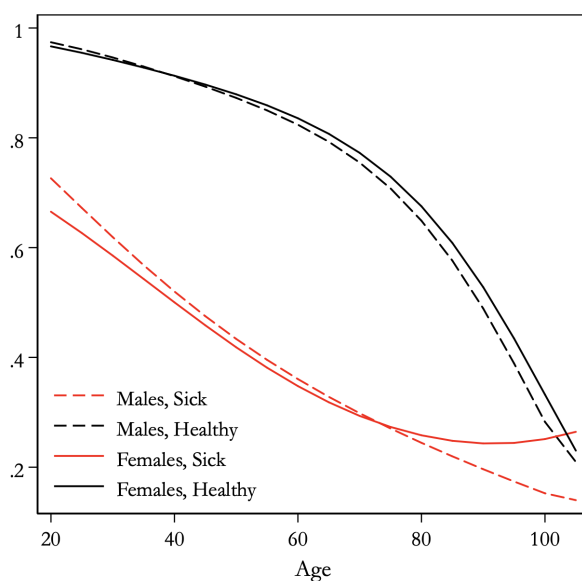
Health spending: We use the Medical Expenditure Panel Survey (MEPS) and Health & Retirement Survey (HRS) to calculate health spending. To increase sample size, we pool years 2000-2019 for the MEPS and years 1996-2016 of the HRS (waves 3-13), and adjust spending to 2019 dollars using the medical CPI. Figure A.2 plots out-of-pocket spending by age. The shock can take 1 of three states each year ("low", "medium", or "high"). We first predict total health spending with a 4th-order polynomial in age using the MEPS, and then

apply a $2,500 deductible, 20% coinsurance, and $5,000 out-of-pocket max to represent cost sharing in a high-deductible health plan while working. These parameters are close to the averages of HDHPs offered by US employers according to the Kaiser Family Foundation's Employer Benefit Survey. These parameters produce a 73% actuarial value (the amount of total care covered by the plan rather than the individual, excluding premiums), based on the calculator provided by the Center for Medicare and Medicaid Services' Center for Consumer Information and Insurance Oversight. This level is slightly more generous than a "Silver" plan on the Affordable Care Act's exchanges, which is the benchmark level for determining insurance subsidies. In retirement, we predict the sum of out-of-pocket payments and Medicare premiums from the HRS with a 4th-order polynomial in age. We again plot out-of-pocket spending by age, and allow the shock to take 1 of three states each year ("low", "medium", or "high").

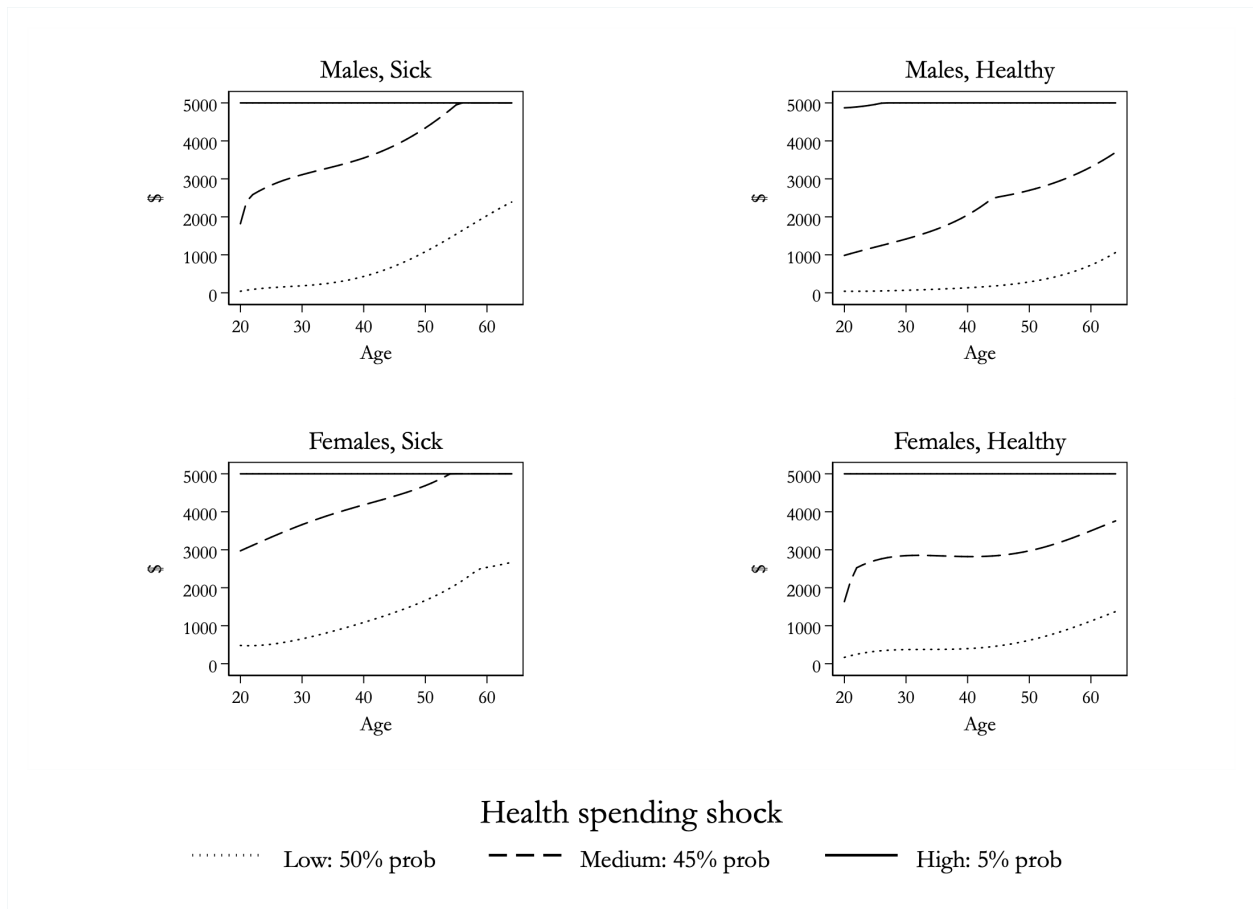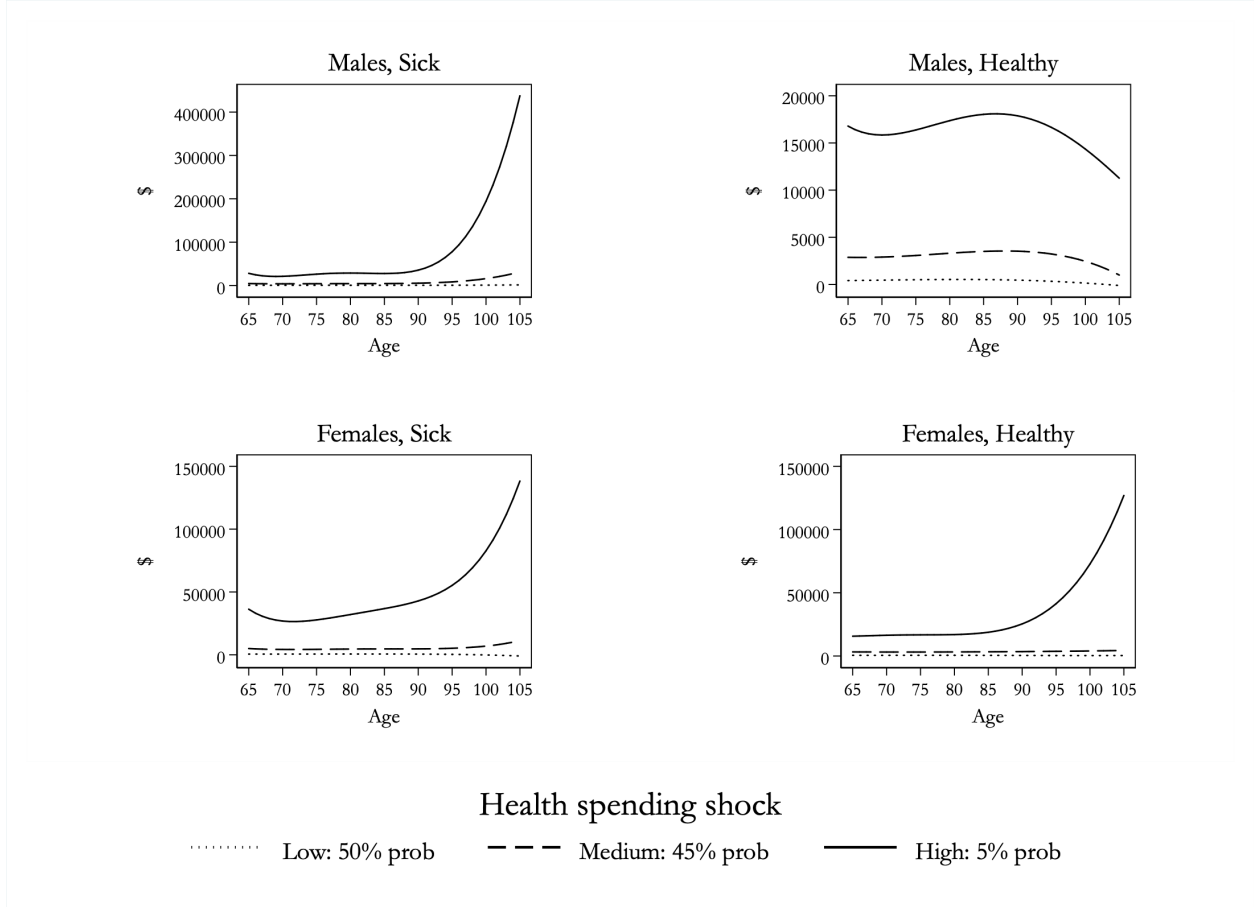Figure A.2: Distribution of Spending Shock During Working Years

Figure A.3: Distribution of Spending Shock During Retirement

Consumption shock: We use the Panel Study of Income Dynamics (PSID) to parametrize the consumption shock. We are interested in modeling the role of moderately-sized, typical shocks besides health care that might affect decisions about saving and withdrawals from tax-preferred accounts. We consider vehicle repairs and house repairs, which are each measured in the PSID. Using the surveys from 2011 to 2019, we run the following regression:

$$x_{it} = \gamma_0 + \gamma_1 y_{it} + \gamma_2 y_{it-1} + \gamma_3 n_{it} + \mu_i + \lambda_t + e_{it}$$

where x_{it} is the sum of expenses on repairs for individual i in year t, y_{it} denotes income, n_{it} denotes the Census needs standard based on household size and the age of household members, μ_i denotes individual fixed effects, λ_t denotes year effects, and e_{it} is an i.i.d. residual. We choose to fix the shock at \$2,000 and then use the regression to calculate the probability of an average shock of this size. Specifically, we calculate the fraction of residuals such that the average $\hat{e}_{it}$ equals \$2,000. We find this amounts to the top 40% of residuals and use this percentage as the likelihood of a consumption shock. We do not consider the possibility of positive shocks (a negative residual) and assume with 60% there is no shock.
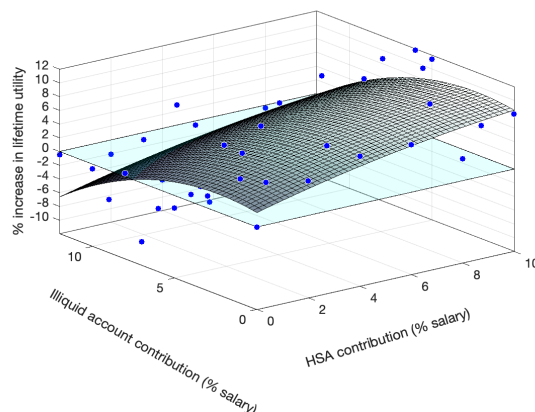
Appendix B: Additional Results

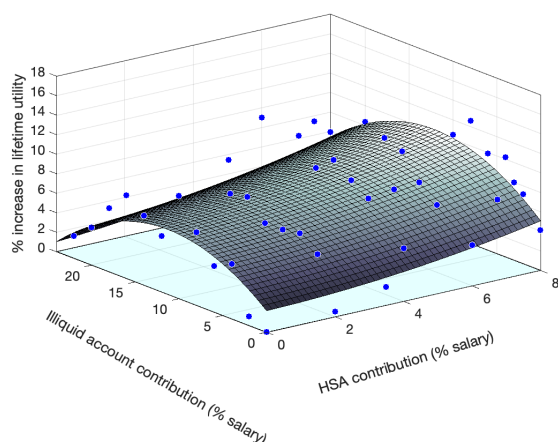Figure B.1: Increase in Lifetime Utility Relative to Liquid Saving Only, $10,000 initial assets
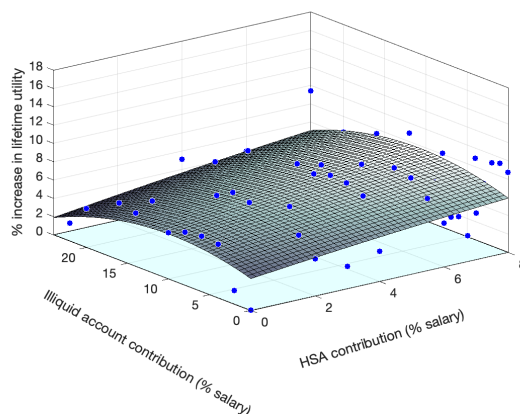
(a) $65,000 salary, men
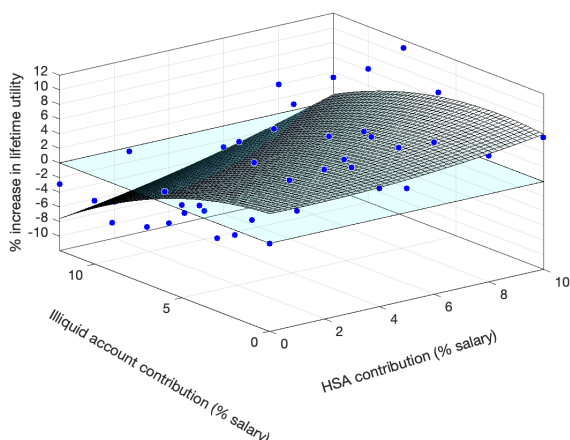
(b) $65,000 salary, women

(c) $105,000 salary, men

(d) $105,000 salary, women

Notes: Figures plot the percentage change in lifetime utility for different combinations of HSA and illiquid retirement account contribution rates as a fraction of annual salary. The increase is measured relative to optimally saving in the taxable liquid account only. Each panel plots the results for a person with different salary profiles and gender. Initial endowments of assets in the taxable account are set to $10,000. The surface plot is fitted using a 3rd-order polynomial calculated using the output of the model from 50 different combinations of contribution rates. Shaded plane denotes zero lifetime utility relative to liquid saving only.

Table B.1: Optimal Fixed Contribution Rates and Lifetime Utility

Salary level (initial)	Gender	without HSA		with HSA		
		optimal illiquid saving	% utility increase vs. liquid saving only	optimal illiquid saving	optimal HSA saving	% utility increase vs. liquid saving only
$65,000	Men	0	0.0	3	10	2.8
$65,000	Women	2	0.5	2	10	7.0
$105,000	Men	11	4.4	12	8	8.1
$105,000	Women	10	1.4	11	8	4.2

Note: Table presents the set of fixed contribution rates that maximize lifetime utility from the model. The points represent the maxima from the surfaces shown in Figure B.1. For comparison, the table shows the rates that maximize utility when saving can only occur in the illiquid and liquid accounts but not the HSA, in addition to the case when all three accounts are available.