Optimal savings with taxable and tax-deferred accounts

Francisco Gomes a,b,1, Alexander Michaelides b,c,d,2, Valery Polkovnichenko e,*

a London Business School, London, UK
b CEPR, London, UK
c London School of Economics, London, UK
d FMG, UK
e UT at Dallas, Richardson, TX, USA

1. Introduction

Individual tax-deferred retirement accounts in the U.S. pension system have grown considerably during the last two decades. According to the Investment Companies Institute 2007 Fact Book (http://www.icifactbook.org), as of December 2006, the retirement plans with individually-directed asset allocation decisions accounted for $8.3 trillion (that is, more than 50% of all retirement assets of $16.4 trillion).3 As more households rely on defined contribution plans to finance their retirement expenditures, understanding the influence of such tax-deferred accounts (TDAs) on their individual decisions, and on aggregate economic activity becomes increasingly important.

In the last two decades there has been substantial work devoted to the understanding of the role of tax-deferred accounts in individual saving over the life cycle. Tax-deferred accounts provide investors with a higher effective after-tax return than (standard) taxable accounts (TAs) since taxes are only paid upon withdrawal of the funds. More precisely, the initial investment is exempt from labor income taxes, and the investments accumulate tax free. Consequently, for a given consumption (and savings) path, households will accumulate more wealth by investing in TDAs rather than TAs. As well
understood by economists, this creates two opposing incentives for households: a substitution and an income effect. Therefore, the impact on the optimal savings decision is ambiguous. Empirically, the impact has been hard to measure, due to the endogeneity of the TDA participation decision (TDA-holders may have stronger saving motives). For instance, eligibility for a TDA, which is often used as an instrument, can also be endogenously related to savings motives since workers with higher savings motives may opt into the jobs with 401(k) retirement plans (see, for example, Bernheim, 2002 for a discussion).

As a complement to this approach, here instead we estimate a structural life cycle model to shed light on this very important question. Our motivation is that, while ambiguous in general, the effect of TDAs on savings may become more precise in a relatively rich model parameterized consistently with empirical observations. We therefore build a structural life cycle model with both taxable and tax-deferred accounts, designed to capture some key features of household-level data by integrating two main motives previously identified as quantitatively important in explaining individual and aggregate savings: a precautionary and a retirement motive. These motives vary in importance among two distinct household groups: indirect and direct stockholders. Indirect stockholders only own stocks in their tax-deferred accounts, while direct stockholders hold equities in their taxable accounts (and may also own stocks in tax-deferred accounts). This categorization is motivated by empirical evidence from the Survey of Consumer Finances (SCF) which shows that these two groups have substantially different wealth accumulation profiles. Specifically, indirect stockholders accumulate less wealth, especially in their taxable accounts.

To generate the empirical differences in wealth accumulation between direct and indirect stockholders, we allow them to have different preference parameters in the estimation, more precisely, we allow for heterogeneity in the elasticity of intertemporal substitution, and in the discount factor. The estimation results confirm that there is indeed a significant difference in these parameters across the two groups. Moreover, we assume that there is a small fixed cost of becoming a direct stockholder, and making the decision of direct stock market participation endogenous. We estimate the two preference parameters for the two different groups to match their respective median TA and TDA wealth accumulation over the working lifecycle. With a fixed cost of around 5% of mean annual labor income, the model can replicate limited direct stock market participation and reasonable wealth accumulation profiles over the life cycle.

The estimated structural model is then used as a benchmark to study the impact of TDAs along a number of different dimensions. We find that, in the presence of tax-deferred accounts, wealth accumulation increases but household net savings (total income minus consumption) are only marginally affected. The income effect from the tax shelter offsets the substitution effect for a large range of preference parameters based on the previously estimated values. The tax savings from the TDA generate higher wealth accumulation for a very similar savings rate and therefore households can enjoy a higher wealth-to-earnings ratio at retirement without having to decrease their working-life consumption. As a result, early in life the consumption patterns are essentially the same with and without the TDA, and at mid-life households already start to increase their consumption. Therefore, the TDA wealth accumulation comes mostly from the combination of a crowding-out effect in the TAs, and the income effect from the tax savings. This conclusion is very important since, if TDAs generate higher total wealth accumulation mostly due to the tax savings effect, then this might not carry over in general equilibrium, where those tax deductions have to be financed. The analysis in the paper is partial equilibrium in nature, and therefore we do not make claims about overall increases in wealth or savings.

Our main focus is instead on the differential impact of these accounts on the individual life-cycle profiles, as all households will face the same potential changes in rates of return and wages (or even tax rates) that might arise in a general equilibrium analysis. First, we find that the increase in consumption occurs mostly during retirement. This is important because it shows that TDAs are effective in promoting a transfer of resources from working life to the retirement period, which is what they are designed to do. So, even though they do not promote significant additional household savings, the tax benefits are used to finance extra retirement consumption. However, we also find that the households that are most responsive to these incentives are those that would already save more in their absence. More precisely, while the households with the highest savings motive (thus the highest saving rates in the absence of a TDA), increase their net wealth accumulation when offered the opportunity to invest in these accounts, those with the weaker savings incentives keep their consumption levels unchanged and simply transfer (part of) their previous savings to the new account. Therefore, TDAs will have a smaller impact on the households that save less for retirement in the first place. In addition, for these households, the higher consumption is almost exclusively concentrated in the years just before retirement and in the first 15 years of retirement, so that they are still left with very little wealth to finance their old age expenditures.

The contribution of this paper complements the findings of the empirical literature studying the role of tax-deferred accounts in individual savings. Some of the earlier papers highlighted the difficulties in identifying empirically the impact of TDAs and reached conflicting conclusions. For example, Poterba et al. (1995, 1996) argue that 401(k) participants have increased savings, while Engen et al. (1996) conclude that no new savings were created. Bernheim (2002) provides an excellent survey of the literature on this and related topics. More recently, Engen and Gale (2000) find that for higher income 401(k)-plan participants, TDA savings represent a substitution of other savings. On the other hand, lower income participants appear to have increased their savings after becoming eligible to participate in 401(k) plans. Gale (2005) surveys the literature and argues that there is a mismatch between those who take advantage of the pension subsidies and those who need to save more for retirement. This is consistent with our results.

---

4 We acknowledge that there is still potential for some differences, for example due to the progressivity of personal income tax rates.
Our paper is also related to a growing literature on individual investment decisions with taxable and tax-deferred accounts. The papers by Dammon et al. (2004) and Garlappi and Huang (2006) are primarily concerned with normative aspects of asset location and allocation and tax efficiency. Our paper instead attempts to explain empirical patterns in wealth accumulation, portfolio choice and direct stock market participation, and focuses on the impact of TDAs on savings decisions.

The paper is organized as follows. Section 2 describes the empirical evidence on wealth accumulation and asset allocation. We present the model in Section 3. Section 4 discusses the estimation and compares the model with the data. Section 5 discusses the impact of TDAs on wealth accumulation and welfare. Section 6 concludes. In Appendix A we describe the construction of variables from SCF data, in Appendix B we outline the numerical procedure used to solve the model, in Appendix C we provide some details with regards to the estimation procedure and in Appendix D we outline a three period model.

2. Wealth accumulation by direct and indirect stockholders over the life cycle

We investigate stock ownership and TDA participation among the U.S. population using the Survey of Consumer Finances database for 2004. We divide households into two groups based on stock ownership: (i) direct stockholders (DS) who own taxable stocks/equity funds (note that these households also may own equities in the TDA); (ii) indirect stockholders (IS) who have equity only in the TDA (employer stock and/or funds with equity investments). We define TDAs as account-type pension plans, i.e. defined contribution plans where participants accumulate balances.5

The distinction between direct and indirect stockholders is important for our subsequent theoretical analysis and is motivated by the following considerations. We present empirical evidence that stock market participation through the TAs and TDAs is undertaken by different types of households. Buying stocks and mutual funds “on your own” in the TA requires a certain degree of financial sophistication as well as sufficient funds to justify paying certain transaction costs. This intuition is usually captured in the limited stock market participation literature through a fixed cost that represents a combination of explicit and implicit hurdles such as brokerage fees, information acquisition about various type of accounts and investment opportunities, more complicated tax filing and time spent on setting up, rebalancing, and monitoring the investment. In the context of the TDAs it is hard to justify such frictions, especially in the case of employer-provided retirement plans. TDAs simplify access to capital markets by providing a uniform, simple and virtually costless vehicle in which to invest one’s savings for retirement. Thus, we would expect households that participate in the stock market only through the TDAs to have different financial characteristics from those that hold stocks directly.

We construct age profiles for wealth and portfolio allocations according to the following age groups: (i) 35 and under; (ii) 36–50; and (iii) 51–65. We ignore the over 65 category because a more detailed model about the retirement period will be needed as, for instance, in De Nardi et al. (2006). Nevertheless, we think that the additional explanatory power with regards to working life saving decisions will not compensate for the additional complexity to be introduced and we leave this to future research. Another complication involves the linear relationship between time, cohort and age effects. TDAs have been gradually expanding over time and, as a result, several 40-year old households now have a TDA, while most 60-year old households did not have one when they themselves were 40. In the model we will later include a distribution of “years since access to a TDA” and, therefore, both the SCF data and the data simulated from the model will include cohort effects. Conditioning on year 2001 instead of year 2004 does not significantly affect the wealth to earnings ratios in the data. We view this as suggesting that the time effects with regards to wealth variables are very small and that therefore we can interpret the reported wealth accumulation as a life-cycle one.

Table 1 shows separately for the TAs and TDAs the cross-sectional medians of wealth-earnings ratios.6 For the earnings variable we use all non-financial income which includes wages and salaries, proprietor’s income, and various sources of government aid. Appendix A provides further details. Panel A of Table 1 shows wealth-earnings ratios for each age group during working life. Wealth-earnings ratios increase with age. Throughout the life cycle, DS have considerably higher taxable wealth compared to IS: taxable wealth differs by at least an order of magnitude across the two types in various age groups. In contrast, in the TDAs, the wealth-earnings ratios are closer, probably due to caps on contributions to TDAs. Further, the historical access to TDAs has not been uniform across the age cohorts. Tax-deferred retirement plans have only become widespread during the 1980s. For many older households, therefore, these plans were not available until they were close to retirement. This consideration becomes important when matching simulated to actual data. Panel B reports the 25th and 75th percentiles for the same variables and illustrates the substantial dispersion that exists in wealth to earnings. One message from the panel is that the dispersion of taxable wealth for the (richer) DS is much higher than the dispersion of outcomes for the TA of the (poorer) IS. The dispersions in the TDAs, on the other hand, are much more closely aligned (even though the DS still show a greater dispersion by the third phase of the life cycle).

We conclude that households participating in the stock market exclusively through TDAs (IS) are on average poorer than those holding stocks in their TAs (DS): DS have higher wealth to earnings ratios than IS at any point in the life cycle.

5 We explored the data for all available SCF waves since 1989, these are available upon request. With respect to the main differences reported about direct and indirect stockholders our findings are robust across the time periods and we only include the results for 2004 SCF. In computing the reported statistics we use all imputations with corresponding population weights.

6 We report median rather than mean wealth-income ratios because the wealth distribution is heavily skewed.
transitory component. We assume that ln Investors must pay a fixed entry cost before investing in stocks in the taxable account for the first time. In the TDA stock

3.3. Financial assets and taxation

Furthermore, the financial wealth of these two groups is split differently across TAs and TDAs. While the TDA wealth to earnings ratios are similar across the two types throughout working life, the DS have more significant savings in the TAs while IS concentrate their savings mostly in the TDAs.

3. Model

3.1. Preferences

Time is discrete and \( t \) denotes adult age which, following the typical convention in the life-cycle literature, corresponds to effective age minus 19. Each period corresponds to one year and agents live for a maximum of \( T = 81 \) periods (age 100). The probability that a consumer/investor is alive at time \( (t+1) \) conditional on being alive at time \( t \) is denoted by \( p_t \) \((p_0 = 1 \text{ and } p_T = 0)\). Households have Epstein–Zin–Weil utility functions (Epstein and Zin, 1989; Weil, 1990) defined over a single

3.2. Labor income process

The labor income process before retirement follows the standard specification in the life-cycle literature and is given by

\[
Y_t = P_t U_t, \quad (2)
\]

\[
P_t = \exp( f(t, Z_t))P_{t-1}N_t, \quad (3)
\]

where \( f(t, Z_t) \) is a deterministic function of age and household characteristics \( Z_t \), \( P_t \) is a permanent component, and \( U_t \) a transitory component. We assume that ln \( U_t \) and ln \( N_t \) are independent and identically distributed with variances \( \sigma^2_u \) and \( \sigma^2_n \) respectively. The log of \( P_t \) evolves as a random walk with a deterministic drift \( f(t, Z_t) \).

For simplicity, retirement is assumed to be exogenous and deterministic, with all households retiring in time period \( K \), corresponding to age 65 \((K = 46)\). In retirement \((t > K)\) the investor can start to withdraw wealth from their TDA and also

3.3. Financial assets and taxation

The investment opportunity set is constant with two financial assets: one riskless (bonds or cash) and one risky (stocks). Investors must pay a fixed entry cost before investing in stocks in the taxable account for the first time. In the TDA stock

Table 1

Wealth to earnings ratios (medians) and percentiles (25th/75th) by account, age and stockholding status. Included components of wealth and earnings are defined in the data appendix. Households are classified by stock ownership as follows: (i) direct stockholders own some equities/equity funds in TAs (also may have some equity in the TDA) – labeled “DS”, (ii) indirect stockholders own equities/equity funds only through TDA’s – labeled “IS”. We use the 2004 SCF with the sample weights to compute reported statistics. Bootstrapped standard errors are in parentheses.

<table>
<thead>
<tr>
<th>Age</th>
<th>DS Taxable</th>
<th>Tax-Def.</th>
<th>IS Taxable</th>
<th>Tax-Def.</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;36</td>
<td>0.29</td>
<td>0.11</td>
<td>0.03</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.021)</td>
<td>(0.004)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>36–50</td>
<td>0.41</td>
<td>0.47</td>
<td>0.04</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.021)</td>
<td>(0.004)</td>
<td>(0.022)</td>
</tr>
<tr>
<td>51–65</td>
<td>1.14</td>
<td>1.23</td>
<td>0.06</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>(0.042)</td>
<td>(0.061)</td>
<td>(0.010)</td>
<td>(0.032)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age</th>
<th>DS Taxable</th>
<th>Tax-Def.</th>
<th>IS Taxable</th>
<th>Tax-Def.</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;36</td>
<td>0.06/0.78</td>
<td>0.00/0.38</td>
<td>0.00/0.12</td>
<td>0.07/0.41</td>
</tr>
<tr>
<td>36–50</td>
<td>0.11/1.20</td>
<td>0.09/1.18</td>
<td>0.00/0.21</td>
<td>0.14/0.97</td>
</tr>
<tr>
<td>51–65</td>
<td>0.37/3.49</td>
<td>0.23/3.08</td>
<td>0.00/0.26</td>
<td>0.28/1.54</td>
</tr>
</tbody>
</table>

Panel A. Wealth to income ratios (medians)

Panel B. Wealth to income ratios percentiles (25/75)
market participation is costless. The fixed cost is expressed as a percentage of the household’s current permanent income: \( F \times P_t \). This entry fee represents both the explicit transaction cost from opening a brokerage account and the opportunity cost of acquiring information about the stock market. The fixed cost \( F \) is scaled by the current level of the permanent component of labor income \( P_t \) as this simplifies significantly the solution of the model. However, this specification is also motivated by the interpretation of the entry fee as the opportunity cost of time.

There are no additional transaction costs and we do not allow for short sales. Given that asset returns are taxed and taxes are paid on nominal returns, we assume a constant inflation rate \( \pi \). The riskless asset yields a constant real return \( r^b \).

The nominal return is taxed at a rate \( \tau_d \), which is also assumed to be the tax rate on both labor income and dividends. The after-tax, real return on the riskless asset is therefore given by

\[
\frac{r^b}{1 + \tau_d} = 1 + \frac{[(r^b + 1)(1 + \pi) - 1](1 - \tau_d)}{(1 + \pi)} - 1. 
\]

The real return on the risky asset is given by

\[
r_t^r = 1 + g_t + d - 1 + \pi
\]

where \( \mu^s \) is the average, real, before-tax equity premium, and \( \varepsilon_t^s \) follows an i.i.d. \( N(0, \sigma^2_s) \), potentially correlated with the labor income shocks \( (\ln N_t, \ln U_t) \). The random real gross stock return \( (r_t^r) \) is comprised of a constant nominal dividend yield \( (d) \) and a stochastic nominal capital gain \( (\varepsilon_t^s) \), deflated by the inflation rate:

\[
r_t^r = 1 + g_t + d - 1 + \pi
\]

These two components are taxable at different rates. More specifically, nominal capital gains are taxed at the rate \( \tau_g \), whereas nominal dividends are taxed at the rate \( \tau_d \). For simplicity, we assume that all income is taxed at source. The after-tax real return on the risky asset is given by

\[
\frac{r_t^r}{1 + \tau_d} = 1 + g_t + d(1 - \tau_d) - 1 + \pi
\]

3.4. Budget constraint and wealth dynamics

Securities can be kept in two accounts: a tax-deferred retirement account (TDA) and a regular taxable account (TA). In the TA, there is no deferral of dividend and interest income taxes and all taxes are assumed to be paid at source. In the retirement account no taxes are withheld and the investor is free to rebalance her portfolio without creating a tax liability. Throughout working life the investor contributes to the TDA a fraction \( \alpha' \) of before-tax earnings. We set the maximum contribution rate equal to 20% and we do not allow for early withdrawals from TDAs prior to retirement.9 After the age of 70, the investor faces a minimum withdrawal rate equal to the inverse of her life expectancy, according to the current IRS regulation. She pays taxes on the withdrawals at the income tax rate \( (\tau_0) \), and bequests are assumed to be tax-exempt.11

3.4.1. Direct stockholders

We start by considering the budget constraint for households that have already paid the fixed cost, and therefore can invest in equities in their taxable account. Let us first consider the working life period \( (t < K) \). Let \( \alpha^t_t \) denote the share of wealth invested in stocks in the retirement and liquid taxable accounts, respectively. Also, let \( h_t \) be the fraction of income paid in housing expenses, which is taken exogenously from the data.12

The wealth dynamics equations for the liquid taxable account \( (W_t^L) \) and for the retirement or tax-deferred account \( (W_t^R) \) are given by13:

---

7 The motivation for this is that the tax-deferred account bypasses the search costs, inertia and set-up costs associated with direct investment in equities.

8 Capital gains in the U.S. are only taxed upon realization. Modeling this feature requires two additional state variables and substantially complicates the model.

9 Technically, it is possible to borrow against TDA account or withdraw money before retirement. However in practice there are restrictions and non-trivial costs of withdrawal and borrowing. First, withdrawals to finance consumption are not allowed unless justified by “hardship”, e.g. loss of employment, court judgments, illness. Second, withdrawing is usually only possible after a long vesting period (up to 5 years) and the amount withdrawn is subject to 10% penalty and income tax. Third, some TDAs do not allow borrowing, and those that allow it may require the transfer of the collateral balances into safe annuities. Finally, loans against TDA usually carry interest rates comparable to those on 10–15 year mortgages.

10 Some calibrations were re-computed allowing for early withdrawals \( (k_t < 0) \) subject to a 10% penalty and we found that the main quantitative implications remained unchanged. In the simulations households virtually never take advantage of early withdrawal possibility, consistent with findings of Dammon et al. (2004).

11 Under the current U.S. tax code estates are subject to a generous exemption \((\$1M)\), implying that most bequests are effectively tax exempt. Therefore, a zero tax rate is more representative of the average household (both in the SCF and in the model) than the corresponding statutory rate.

12 Details are given in the calibration section. In the model we match financial wealth accumulation and do not consider housing assets explicitly for computational reasons. Specifically, endogenizing housing equity is a formidable computational problem in itself, even without considering TDA-related decisions. See for example recent papers by Cococo (2005) and Yao and Zhang (2005). However, we indirectly account for housing equity accumulation by calibrating an exogenously-specified housing expenses that are paid out of labor income.

13 If the household is paying the fixed cost for exactly this period then we need to subtract \( F \times P_t \) from \( W_t^R - C_t - k_t Y_t (1 - \tau_0) \) in the first budget constraint.
\[
W_{t+1}^L = \left[ \alpha_L^t \left( 1 + r_{t+1}^L \right) + (1 - \alpha_L^t) (1 + \rho^L) \right] (W_t^L - C_t - k_t Y_t (1 - \tau_d)) + (1 - \tau_d) (1 - h_{t+1}) Y_{t+1}, \tag{8}
\]

\[
W_{t+1} = \left[ \alpha_L^t \left( 1 + r_{t+1}^L \right) + (1 - \alpha_L^t) (1 + \rho^L) \right] (W_t^L + k_t^* Y_t) \tag{9}
\]

where \( k_t^* \) includes an employer matching contribution, (one-to-one) up to a maximum \( (k_e) \), and is thus given by

\[
k_t^* = \min(2k_t, k_t + k_e). \tag{10}
\]

After retirement \( (t \geq K) \) these equations change. Let \( Q_t \) denote the withdrawal from the retirement account and let \( A_t \) denote the investor's life expectancy. The equations become:

\[
W_{t+1}^L = \left[ \alpha_L^t \left( 1 + r_{t+1}^L \right) + (1 - \alpha_L^t) (1 + \rho^L) \right] (W_t^L - C_t + (1 - \tau_d) Q_t) + (1 - \tau_d) (1 - h_{t+1}) Y_{t+1}, \tag{11}
\]

\[
W_{t+1} = \left[ \alpha_L^t \left( 1 + r_{t+1}^L \right) + (1 - \alpha_L^t) (1 + \rho^L) \right] (W_t^L - Q_t), \tag{12}
\]

subject to the constraint

\[
Q_t \geq \frac{1}{A_t} W_t^L, \quad t \geq 70, \tag{13}
\]

which imposes a minimum withdrawal rate equal to the inverse of life expectancy for households 70 years of age or older.

We also impose the following borrowing and short sales constraints

\[
\alpha_L^t \in [0, 1], \quad W_{t+1}^L \geq 0, \tag{14}
\]

\[
\alpha_L^t \in [0, 1], \quad W_{t+1}^L \geq 0, \tag{15}
\]

\[
k_t^* \in [0, 0.2], \tag{16}
\]

At the time of death, all funds are withdrawn from the TDA untaxed and are paid to the beneficiary together with the remaining cash on hand balance in the TA.\(^{14}\)

3.4.2. Indirect stockholders

Households that have not yet paid the fixed entry cost are subject to the same constraints except that they can only invest in stocks in their tax-deferred retirement accounts. Therefore, when solving their dynamic programming problem, we only have to replace Eqs. (8) and (11) with

\[
W_{t+1}^L = \left( 1 + \rho^L \right) (W_t^L - C_t - k_t Y_t (1 - \tau_d)) + (1 - \tau_d) (1 - h_{t+1}) Y_{t+1}, \tag{17}
\]

and

\[
W_{t+1} = \left( 1 + \rho^L \right) (W_t^L - C_t + (1 - \tau_d) Q_t) + (1 - \tau_d) (1 - h_{t+1}) Y_{t+1}, \tag{18}
\]

respectively, thus eliminating \( \alpha_L^t \) as a choice variable.

3.5. Estimation

3.5.1. Preference parameters

We use preference heterogeneity and Epstein–Zin (1989) and Weil (1990) preferences to generate predictions that are consistent with observed direct stock market non-participation and substantial wealth accumulation for households that choose to participate directly in the stock market. We consider two groups of agents which will allow us to match the participation decisions, and the wealth accumulation conditional on participation. Both groups have the same preference curvature parameter \( \rho = 4.0 \) in the baseline calibration, but they have different EIS and different discount factors.\(^{15}\) We use the mortality tables of the National Center for Health Statistics to parameterize the conditional survival probabilities.

3.5.2. Labor income process and housing expenditures

It is common practice to estimate different labor income profiles for different education groups (college graduates, high-school graduates, and households without a high-school degree). The deterministic labor income profile reflects the hump shape of earnings over the life-cycle of college-educated households that dominate the class of stockholders in the U.S. economy according to the SCF data. The corresponding parameter values for the hump-shape are taken from Cocco et al. (2005), while the standard deviations of the idiosyncratic shocks \((\sigma_L = 10\% \text{ and } \sigma_n = 10\%) \) are taken from Carroll et al. (1997). The replacement ratio \((\lambda) \) is set at 60% of net earnings which is slightly lower than the 66% estimated by Cocco et al. (2005), since households with TDAs have lower average state pension replacement ratios. Finally, to account for housing expenses

\(^{14}\) We omit estate taxation issues for simplicity.

\(^{15}\) We could have allowed for heterogeneity in risk aversion \((\rho) \), but we have decided to keep the heterogeneity in the model to a minimum. As shown in Gomes and Michaelides (2005) and Alan (2006), households with lower risk aversion will accumulate much less wealth and thus will be much less likely to pay the participation cost and invest in stocks.
we subtract exogenous housing payments from the household’s labor income. These exogenous housing payments are taken from Gomes and Michaelides (2005), who compute using PSID data the ratio of annual, housing-related expenditures to annual labor income as a function of age.

3.5.3. Asset returns, taxes, participation cost and contributions

The real bond return $r^b$ is set at 2%. For the stock return process we consider a mean equity premium $(\mu^d)$ equal to 4% and a standard deviation $(\sigma^d)$ of 20%. We consider 4% as opposed to the historical 6% to take into account transaction costs (e.g. mutual fund and brokerage fees). The nominal dividend yield $d$ is set at 3.2%.\(^{16}\) The proportional tax on dividends and labor income $(\tau_d)$ is 25% and the tax on capital gains, $\tau_c$, is 20%. We consider a labor income tax of 25% to reflect the average income tax of the typical household in our sample. Inflation rate is set at 2.8% corresponding to the average from CRSP data from 1989 to 2004, the period covered by the SCF sample.

The fixed cost of stock market entry is set at 5% of current permanent income $(P_t)$.\(^{17}\) As previously discussed, we view the participation cost as having two important components: the actual monetary transaction cost associated with acquiring information, and the opportunity cost corresponding to the time required for collecting and processing this information. If we assume that the cost is 100% monetary then, taking an average U.S. wage of $38000/year, this number implies a one-time fee of $1900. Pailla (2007) and Vissing-Jørgensen (2001) have estimated the implied participation costs from household-level consumption Euler equations. They obtain per-period costs in the $75 to $200 range. In present-value terms our calibration is quite reasonable when compared with those number: for the vast majority of households the cost in our calibration is much lower than the present value of these alternative estimates. If instead we assume that the cost is 100% opportunity cost of time then, for an average of 252 working days per year, this number implies a one-time cost of 12.6 days.\(^{18}\) More generally, any convex combination of these two costs is valid, depending on one’s priors. For example, $950 and 6.3 days seems reasonable to us, but naturally this is a subjective choice.

We set the cap on employer matching $(k_e)$ equal to 3%, based on two sources of information. First, numbers from the SCF indicate that approximately only 50% of TDAs have employer-matching. Second, Mitchell and Dykes (2003) report that, for those plans with some degree of matching, the mode of employer contributions is 6%. The correlation between stock returns $(\varepsilon_t)$ and permanent labor income shocks $(\ln N_t)$ is set equal to 0.15, based on the results from Campbell et al. (2001), while the correlation with transitory labor income shocks $(\ln U_t)$ is set equal to 0 (the magnitude of this correlation does not affect saving or portfolio choice decisions).

3.5.4. Distribution of years since first access to TDA

According to the data, even among stockholders, a significant fraction of households does not have a tax-deferred account. Moreover, even within those that do have a TDA, most of them have only had one for a small number of years. Therefore, to replicate the average wealth accumulation of households aged 50 to 65, for example, we cannot assume that all of them have had access to a TDA since age 20.

From the 2004 SCF we have information on the number of years each household has contributed to an employer-provided retirement plan, including the years of contributions in any previous plans provided that these have been rolled over to start the current one. Thus, this variable provides us with a lower bound on the number of years that the household has had access to a TDA $(\theta)$.\(^{19}\) We can now use this variable to compute a distribution of $\theta$ and use this as an exogenous input for the model.

For simplicity, we consider groups with three different values of $\theta$. Table 2 reports the different groups and their corresponding population weights, both in the SCF and in the model. The values of $\theta$ used in the model correspond roughly to the mean of each interval in the data, and produce an average $\theta$ of 8.2, which is very close to the number in the data (8.4). Finally, we set the percentage of households without a TDA to 21% to approximate the corresponding figure from the SCF for direct stockholders, since nonstockholders without a TDA are excluded from our analysis.

4. Estimation and model results

Given the complexity in numerically solving a life-cycle model with two continuous state variables (wealth in TA and TDA), four continuous policy functions (consumption/saving, the portfolio allocations in the two separate accounts and the contribution rate in the TDA) and a binary participation decision, we undertake a two-state estimation procedure that is common in this literature (see, for example Cagetti, 2003; De Nardi et al., 2006). We first pick the calibrated parameters in accordance with the empirical evidence (as described in the previous subsections) and then given these chosen parameters, we subtract exogenous housing payments from the household’s labor income. These exogenous housing payments are taken from Gomes and Michaelides (2005), who compute using PSID data the ratio of annual, housing-related expenditures to annual labor income as a function of age.

\(^{16}\) We have scaled down the historical dividend yield by the same factor as the equity return.

\(^{17}\) Alan (2006) estimates a (one-time) stock market participation cost between 3%–5% ($F$ in our model) of the permanent component of labor income $(P_t)$, in the context of a (life-cycle) structural model with the same labor income process.

\(^{18}\) Naturally this refers to the daily working time only, not the full 12.6 days since, in the model, the cost only leads to a reduction in income. Otherwise we would have also considered a decrease in marginal utility, due to the decrease in leisure, which would then allow us to calibrate a lower value of the fixed cost.

\(^{19}\) This constitutes a lower bound both because some households might not have rolled over previous accounts or they had started non-employer provided accounts, such as IRA’s, earlier.
we pick the structural parameters to minimize the distance between the median wealth to earnings ratio over working life between the model and the data. We estimate these structural parameters \( \{ \beta, \psi \} \) separately for the DS and IS. For both groups we match six moments: for both the TA and the TDA we match the median wealth to earnings ratio for the three different age groups: between ages 22 and 35, between ages 36 and 50 and between ages 51 and 65. It is important to note that matching medians versus means is an important choice given the skewness of the wealth distribution. The results in Cagetti (2003) illustrate how, in a CRRA model, attempting to match mean wealth rather than median wealth results in preference parameters that generate more wealth accumulation (higher risk aversion for instance). Given the skewness of the wealth distribution in the data, we therefore base our estimation on matching the median wealth to earnings ratio over the working life cycle. The numerical method for solving for the policy functions given the structural parameters is described in Appendix B, while details of the estimation procedure are discussed in Appendix C.

Table 3 reports the results for the IS and DS respectively. The IS are more impatient than the DS (the discount factor is 0.88 for the IS and 0.9875 for the DS). Moreover, the DS have a greater EIS than the IS; the latter have an EIS \( \psi \) equal to 0.25 and the latter an EIS equal to \( \psi = 0.6 \). Our results are intuitive in the sense that more impatient households accumulate less wealth in the TA. The heterogeneity in EIS is also consistent with empirical evidence based on an Euler equation approach. Specifically, Vissing-Jørgensen (2002) and Malloy et al. (2005) obtain higher estimates of the EIS for richer stockholders, as the DS turn out to be in the data.

In our model households are heterogeneous with respect to their labor income histories and this will impact their participation decisions and wealth accumulation. However, such a mechanism alone is not enough to generate significant wealth heterogeneity, and it would not suffice to endogenize the participation decision. Without preferences heterogeneity the model fails to generate a balanced mix of households who invest in stocks in TAs from young age and households who remain indirect stockholders throughout working life.\(^{20}\)

4.1. Wealth accumulation and contribution rates

Table 4 (Panel A (B) for DS (IS)) shows the median wealth to labor income ratios in the model and the data for different age groups. The results from the model are obtained by simulating 150,000 individual life histories and taking the medians across households and across age groups. We report medians instead of means since the wealth distribution exhibits significant skewness.

Young households face a high expected future labor income, against which they cannot borrow. Therefore, they prefer to consume most of their income, and a very modest saving is done for precautionary reasons. Saving for retirement starts later on in the life-cycle, and is mostly driven by the higher EIS and the need for long-run, life-cycle consumption smoothing. Therefore, the higher EIS households save more and therefore have a stronger incentive to participate in the stock market through the TA account (these are the DS households). The accumulation in the TDA for DS matches well with its empirical counterparts. Nevertheless, the model does predict a larger wealth accumulation through the TA than there is in the data for the second and third part of the lifecycle.\(^{21}\) One potential reason for this discrepancy with the data could be due to real

---

\(^{20}\) The same result is shown by Gomes and Michaelides (2005) in a different context.

\(^{21}\) While the lower discount factor may potentially help to match the TA accumulation better there is a tradeoff with matching TDA at the same time. Discount factor affects TA and TDA simultaneously and lowering it would lead to lower balances in both accounts. The EIS also has impact on both accounts with relatively more importance for the TDA. For lower discount factor, EIS would have to be higher to better match TDA. The estimation procedure optimally chooses the combination of the two considering weighted errors.
Table 4
Wealth to earnings ratios (medians) from simulations and the 2004 SCF data. Only working age medians are used in the estimation. Estimated preference parameters for DS are $\beta = 0.9875$ and $\psi = 0.6$ and for IS they are $\beta = 0.88$ and $\psi = 0.25$. The risk aversion is set to $\rho = 4$ for both groups. The rest of the parameters are described in the calibration Section 3.5.

<table>
<thead>
<tr>
<th>Age</th>
<th>Model</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Taxable</td>
<td>Tax-Def.</td>
</tr>
<tr>
<td>&lt;36</td>
<td>0.12</td>
<td>0.14</td>
</tr>
<tr>
<td>36–50</td>
<td>1.34</td>
<td>0.41</td>
</tr>
<tr>
<td>51–65</td>
<td>3.19</td>
<td>1.32</td>
</tr>
</tbody>
</table>

Table 5
Wealth to earnings ratios percentiles (25th/75th) from simulations and the 2004 SCF data. Only working age medians are used in the estimation. Estimated preference parameters for DS are $\beta = 0.9875$ and $\psi = 0.6$ and for IS they are $\beta = 0.88$ and $\psi = 0.25$. The risk aversion is set to $\rho = 4$ for both groups. The rest of the parameters are described in the calibration Section 3.5.

<table>
<thead>
<tr>
<th>Age</th>
<th>Model</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Taxable</td>
<td>Tax-Def.</td>
</tr>
<tr>
<td>&lt;36</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>36–50</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>51–65</td>
<td>0.25</td>
<td>0.67</td>
</tr>
</tbody>
</table>

Table 4 compares the range of possible outcomes predicted by the model relative to their empirical counterparts using the 25th and 75th percentiles. With some exceptions, the model generally has difficulty matching the dispersion of wealth observed within age cohorts. Comparing Panel A and Panel B we can see that the ranges for the IS in both accounts are more consistent with the data. Overall, most ranges for the data are wider than the ones predicted by the model illustrating that the model does not generate as much inequality (conditional on participation status) as there is in the data. Note also that the 25th percentiles in TA wealth for DS are considerably higher than the medians for IS from Table 4. This underscores why it would not be possible to use a single set of preference parameters to generate wealth heterogeneity consistent with stock market participation decisions across the two groups.

It is worth exploring how wealth accumulation profiles are affected by the participation cost. While we do not attempt a full sensitivity analysis with re-estimation, we investigate the economic intuition behind the impact of the fixed cost. We use a lower participation cost of $F = 2.5\text{%}$ to recompute and simulate the model for DS and IS parameters. The main effect of the lower cost is to trigger stock market participation earlier in the life cycle and for lower wealth levels. The results for DS are virtually unaffected by this change. The IS have a weaker savings motive and they accumulate a low level of wealth in the taxable account and therefore optimally choose not to invest in equities. Comparing these results with those obtained for DS (Panel A) we find something more consistent with the data. Comparing Panel A with Panel B of Table 4 we find that, as we would expect, the differences in wealth accumulation are stronger in the TA because of the participation constraint.

The IS have a weaker savings motive and they accumulate a low level of wealth in the taxable account and therefore optimally choose not to invest in equities. Comparing these results with those obtained for DS (Panel A) we find something more consistent with the data. Comparing Panel A with Panel B of Table 4 we find that, as we would expect, the differences in wealth accumulation are stronger in the TA because of the participation constraint.

The model therefore may overstate the desired level of financial assets.

The IS have a weaker savings motive and they accumulate a low level of wealth in the taxable account and therefore optimally choose not to invest in equities. Comparing these results with those obtained for DS (Panel A) we find something more consistent with the data. In the model this is matched by the fact that IS have a lower discount factor than the DS. Consequently, they care less about retirement savings, which is consistent with the low wealth accumulation throughout the lifecycle observed in the data. Comparing Panel A with Panel B of Table 4 we find that, as we would expect, the differences in wealth accumulation are stronger in the TA because of the participation constraint.

Table 5 compares the range of possible outcomes predicted by the model relative to their empirical counterparts using the 25th and 75th percentiles. With some exceptions, the model generally has difficulty matching the dispersion of wealth observed within age cohorts. Comparing Panel A and Panel B we can see that the ranges for the IS in both accounts are more consistent with the data. Overall, most ranges for the data are wider than the ones predicted by the model illustrating that the model does not generate as much inequality (conditional on participation status) as there is in the data. Note also that the 25th percentiles in TA wealth for DS are considerably higher than the medians for IS from Table 4. This underscores why it would not be possible to use a single set of preference parameters to generate wealth heterogeneity consistent with stock market participation decisions across the two groups.

It is worth exploring how wealth accumulation profiles are affected by the participation cost. While we do not attempt a full sensitivity analysis with re-estimation, we investigate the economic intuition behind the impact of the fixed cost. We use a lower participation cost of $F = 2.5%$ to recompute and simulate the model for DS and IS parameters. The main effect of the lower cost is to trigger stock market participation earlier in the life cycle and for lower wealth levels. The results for DS are virtually unaffected by this change. This is because the households with DS parameters already accumulate...
human capital is still quite high (see, for example, Jagannathan and Kocherlakota, 1996). As retirement approaches, the young households invest a larger component of their financial wealth in stocks: the implicit riskless asset in the form of stocks (Cocco, 2005) and also reduce attractiveness of the stock market and lower the cost required to deter participation from household portfolios (e.g. real estate, private businesses). Introducing such assets may reduce portfolio allocation to closer to the model. One reason that the model may overstate portfolio allocation is because we omit other sources of risk wave of the SCF portfolio allocations in stocks for DS are higher by as much as 20–30 percentage points and are much lower because now the IS groups consist of less wealthy households. The medians of TDA wealth for the three groups of IS are only [0.0, 0.0, 0.29] under the lower cost assumption which is underestimating the data by a larger amount than the original calibration. We conclude from this exercise that our results for IS are sensitive to the value of the fixed cost and that for lower cost it would be harder to match retirement savings.

4.2. Portfolio allocations

Compared to wealth to earnings ratios, we have found that the portfolio allocations between stocks and bonds are more sensitive to the time period for the SCF survey. Specifically, the 2004 SCF produces more conservative portfolios than the 2001 one in both accounts and this is probably a response to the stock market crash in 2000. Unlike the wealth to earnings profiles we therefore view the portfolio shares as indicative rather than strict life-cycle profiles.

We first consider the portfolio decisions of the IS. They hold very little financial wealth in their TAs because they have a lower discount factor and their retirement saving is done only through their TDAs. Therefore, a small cost of stock market participation is sufficient to prevent these households from investing in equities. As a result their only relevant portfolio decision occurs in the TDA account. The age profiles of the mean portfolio allocation to stocks in the TDA for the IS are shown in Table 6. The profile in the data varies in the range of 50–68% over the life cycle. With low financial wealth and given the equity premium, the model predicts a full specialization in stocks in the TDA, a more aggressive stance compared to the data. Interestingly, the 2001 SCF numbers were closer to the ones the model predicts, the medians for IS in the TDAs were in the range of 90–100%.

The asset allocation decisions of the DS are shown in Table 7. Our generated portfolios are more aggressive (higher proportion invested in stocks) in both accounts than their empirical counterparts. Similar to the case of IS the previous wave of the SCF portfolio allocations in stocks for DS are higher by as much as 20–30 percentage points and are much closer to the model. One reason that the model may overstate portfolio allocation is because we omit other sources of risk from household portfolios (e.g. real estate, private businesses). Introducing such assets may reduce portfolio allocation to stocks (Cocco, 2005) and also reduce attractiveness of the stock market and lower the cost required to deter participation for a given wealth level.

In terms of life cycle behavior, the predictions of the model are similar to previous results in the literature. Specifically, young households invest a larger component of their financial wealth in stocks: the implicit riskless asset in the form of human capital is still quite high (see, for example, Jagannathan and Kocherlakota, 1996). As retirement approaches, the

---

Table 6

<table>
<thead>
<tr>
<th>Age</th>
<th>Model</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;36</td>
<td>1.00</td>
<td>0.60</td>
</tr>
<tr>
<td>36–50</td>
<td>1.00</td>
<td>0.68</td>
</tr>
<tr>
<td>51–65</td>
<td>1.00</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Table 7

<table>
<thead>
<tr>
<th>Age</th>
<th>Model Taxable</th>
<th>Data Taxable</th>
<th>Model Tax-Def.</th>
<th>Data Tax-Def.</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;36</td>
<td>0.50</td>
<td>0.63</td>
<td>0.79</td>
<td>0.75</td>
</tr>
<tr>
<td>36–50</td>
<td>1.00</td>
<td>0.65</td>
<td>0.78</td>
<td>0.64</td>
</tr>
<tr>
<td>51–65</td>
<td>0.99</td>
<td>0.65</td>
<td>0.66</td>
<td>0.56</td>
</tr>
</tbody>
</table>

---

23 A small fraction of about one percent remains nonstockholders until retirement.

24 It is important to emphasize that these results do not control for time or cohort effects. We are just reporting cross-sectional results without taking a view on the relative importance of time or cohort effects. Ameriks and Zeldes (2001) obtain very different portfolio profiles when controlling for time effects versus controlling for cohort effects. Without any prior view on their relative importance it is impossible to determine the correct profile.

25 The exception is the young group which includes households who did not yet pay the participation cost. However, all of them pay the cost by the age of 35 and conditional on participation hold portfolios 100% invested in stocks.
present value of future labor income decreases. Households respond by reducing their exposure to the stock market. This pattern is visible in overall portfolio allocation (not reported) and in each account separately both in the model and in the data.26

Across the two accounts we can observe that TA is more heavily invested in stocks compared to the TDA. This is driven by the tax efficiency motive since interest on bonds is taxed at a higher rate compared to stocks’ capital gains.27 In the data, the average portfolio allocation includes both bonds and stocks in both accounts (Ameriks and Zeldes, 2001 and Bergstresser and Poterba, 2004). Our model partially addresses this issue and is able to generate some bonds in the TA due to the presence of uninsurable labor income risk, the illiquidity of TDA and the stock market participation cost.28

4.3. Contribution rates

Table 8 (Panel A (B) for DS (IS)) reports the average contribution rate in the TDA. The contribution rates increase over the life cycle for both DS and IS and also increase if the number of years of access to the TDA are reduced (controlled by variable $\theta$ in the table). These contributions are especially high for the DS (Panel A) who care more about retirement savings due to their higher discount factor and EIS.29 Nevertheless, despite their very high contribution rates from age 35 onwards, Table 4 shows that most wealth is still accumulated in the TA. The data also shows this seemingly counterintuitive pattern. This occurs because within each age group, there are households that have only had access to the TDA for a small number of years, and therefore they do not have much wealth in that account.30 This shows the importance of accurately calibrating the distribution of $\theta$ if we want to match the data. Overall, our results are broadly consistent with their empirical counterparts (columns 3–4 in Table 4).

The IS (Panel B) qualitatively behave in a similar way as the DS (contribution increasing over the life cycle and decreasing in $\theta$, assuming the zero constraint is not binding). Nevertheless, a key difference involves the magnitude of the contribution with the IS contributing substantially lower amounts than the DS. Thus, even though the IS have a stronger incentive to use their TDAs because they do not directly invest in equities, they still have lower contribution rates than the DS.

5. The impact of tax-deferred accounts on consumption and savings

We can now use our estimated model to study the household-level effects of introducing tax-deferred accounts. Adding a TDA in the model increases the household’s life-time resources through the deferral of taxation and the employer contribution matching. Ignoring the contribution, to keep the intuition simpler, TDAs will generate an income and substitution effect, and household preferences will determine which one will dominate. Moreover, the presence of undiversifiable idiosyncratic income risk creates a precautionary savings demand that induces savings in the liquid TA.

---

26 Once the participation cost is paid, in the model the portfolio allocation in the TA is essentially 100% in equities until approximately age 65 so the declining pattern is only evident after that.

27 See for example, Dammon et al. (2004) who show that when the capital gains tax is lower than the interest income tax (as is the case in the U.S., and in our model), it is tax-efficient to receive capital gains income in the taxable account and interest income in the tax-deferred account.

28 The cross-sectional average also appears to contradict tax-efficiency because some households without TDAs hold balanced portfolios in the TA, as they only get access to the TDA later in the lifecycle. Therefore, in a cross section the average portfolio allocation in the TA and in the TDA would be mixed, even if individual asset location were tax-efficient. Consistent with this, Bergstresser and Poterba (2004) find from the SCF that at the individual level the asset location efficiency is much better than it appears from the cross-sectional averages. They report that relatively small fraction of households have tax-efficient asset location and the dollar size of average inefficiency is small.

29 We do not compare the contribution rates implied by the model with the data since there is no comprehensive survey data on household contribution rates as a function of age.

30 For example, consider a 55-year old household that has only had access to a TDA at age 50. Naturally, in the last 5 years, she will have contributed significantly to this account. However, the vast majority of her (retirement) savings will still be in her taxable account where she had been saving for about 30 years already.
In Appendix D we construct a simple 3-period model to illustrate this intuition and show how consumption in early and later periods will be affected with the introduction of a TDA. With CRRA preferences, our results in Table 9 (calibration details can be found in the appendix) show that for EIS greater than 0.5, there are additional net savings in the early periods of the life cycle (lower consumption) and higher consumption in retirement. On the other hand, the income effect starts to dominate for EIS less than 0.5.\(^\text{31}\) The simpler model suggests that TDA can potentially encourage additional savings and by doing so stimulate the transfer of resources from working life to retirement which is the primary goal of the tax subsidy embedded in such accounts. This conclusion depends on the parameterization of the model and this is why it is important to consider a fully specified model with plausible parameters which generate realistic life-cycle savings and consumption profiles as discussed in the previous sections.

We now turn to the full model to evaluate the impact of introducing TDAs on household-level decisions. We do this by comparing a household that has access to the tax-deferred account since age 20, with another one that never has access to this account. For the purpose of this analysis we exclude the employer matching feature, to focus exclusively on the properties of the TDA, namely the trade-off between tax benefits and illiquidity. Probably the most important policy question regarding TDAs is whether these accounts only crowd-out taxable savings, or whether the preferential tax treatment effectively induces extra savings for retirement, and if so, by how much? In our model, in the absence of tax benefits, the optimal investment in the TDA would be zero, since this account would be dominated by the TA which offers liquidity before retirement. Therefore, by comparing the two models, we are measuring the net impact of this tax incentive on household-level savings. We will perform this experiment for the two sets of parameter values that we have previously estimated as being representative for the population of DS and IS.

First we will consider the DS parameter values. Fig. 1, upper panel, shows the ratio between the consumption for a household with a TDA and the consumption of the same household without a TDA. The lower panel reports similar comparisons but for wealth accumulation. Here we plot both the ratio between total wealth with and without TDA, and between total liquid wealth with and without the TDA. Naturally, for the household without access to a TDA, liquid wealth and total wealth are the same, so the denominator is the same in both cases. We can see, from Fig. 1, that early in life the household saves slightly more than before, but that this effect is small: less than 2\% of consumption in additional savings is made. The lower panel of Fig. 1 shows that these additional savings are being invested in the TDA. In fact, liquid wealth is actually less than in the absence of the retirement account (by 15\% to 20\%). Therefore, early in life, there is a non-trivial crowding out-effect from introducing the TDA, and there are also some small additional net savings. At middle age, the household with a TDA starts to enjoy the benefits of higher wealth accumulation (due to the tax deferral benefits), and as such consumption starts to increase fairly rapidly. As we can see, the presence of a TDA has allowed the household to significantly increase consumption during retirement and, to a smaller extent, just before retirement. From this perspective these accounts are fairly successful: they promote a transfer of resources to the later part of the life-cycle, which is what they are designed to achieve.

Fig. 2 plots the results of the same comparison, but now using the preference parameters estimated for the IS. In this case we observe that consumption early in the life-cycle is virtually the same with or without the TDA. Consumption begins to rise during working life from ages 40–49 and the peak of the rise occurs in retirement when household gains access to the TDA. Comparing with the previous case we find that, the increase in consumption from age 60 onwards is much higher for DS due to their additional savings in the TDA. As we can see from the lower panel of Fig. 2, consistent with consumption behavior, total wealth does not increase for the younger consumers. As the retirement motive begins to dominate savings decisions, the IS households substantially reduce balances in the TA and begin to save in the TDA. Due to faster wealth accumulation in the TDA from the tax deferral, total wealth rises relative to the case without a TDA. Thus, a higher consumption level can be afforded at retirement but it occurs entirely at the expense of crowding out savings in the TA.

\(^{31}\) The discount factor plays some role in determining this threshold. Lower discount factor results in lower threshold, but large changes are required for quantitative effect to be significant. We compared baseline from Appendix D with \(\beta = 0.71(=0.9875)^{20}\) to a similar 3-period model with substantially lower value \(\beta = 0.54(=0.9720)^{20}\) and found that the corresponding EIS threshold where income effect dominates shifted to around 0.45.
Our conclusion from these experiments is two-fold. First, the TDAs are effective in increasing consumption during the retirement period. In part this occurs due to the tax subsidy and in part due to increased savings. However, the second effect is very small. Second, the savings increase is generated only for the households with a sufficiently high savings incentive (high EIS parameter), as is the case of the DS. When the EIS is low (as estimated for the IS), consumption remain unchanged. While consumption is still higher in retirement than without a TDA, this occurs due to deferral of taxes and not due to an increase in net savings. This result presents a challenge as it shows that TDAs may fail to induce additional savings for households who do not save much in their absence.

6. Concluding remarks

We have analyzed a model of life-cycle consumption and portfolio allocation with a taxable and a tax-deferred account (TA and TDA) in the presence of uninsurable labor income risk and borrowing constraints. The model generates plausible wealth accumulation and portfolio choice for direct and indirect stockholders conditional on age and account tax status and generates an endogenous separation between these two groups. Our findings support the hypothesis that TDAs increase wealth accumulation. However, these accounts only induce marginal increases in net savings, and mostly crowd-out wealth accumulation in the taxable account. The tax benefits of the TDA generate additional wealth accumulation, for the same savings rate, and are used to generate higher consumption during retirement, consistent with the objective of these accounts. However, households with the lowest savings incentives are exactly the ones that will respond less to these accounts. In particular, those households that were not accumulating significant retirement wealth will continue to do so.\footnote{This effect is likely to hold in broader population which includes non-stockholders without TDA. They save even less than IS and therefore the impact of TDAs on their savings is potentially negative.} An implication of this result is that policies which have a goal of increasing participation in defined contribution pensions have to target lower wealth households to be efficient. One potentially effective way of achieving this result is to make automatic enrollment with the ability to opt out (see for example discussion in Benratzi and Thaler, 2007).

The results in the paper could naturally change if we were to consider alternative extreme preference parameter values. However, those values would not generate realistic life-cycle savings and consumption profiles as shown in the paper.
Fig. 2. Relative change in consumption and wealth from introducing TDA for indirect stockholders.

highlights the importance of having a well-calibrated/estimated model for these quantitative evaluations. Our model has several natural extensions, which can be implemented in future research. First, we use a flat tax rate while in reality the tax code is progressive. Modeling this feature would likely increase the benefits from the TDAs as retirees tend to face lower tax rates than workers. Second, capital gains realizations in the TA could be explicitly considered, as in Dammon et al. (2004). This would incorporate the value of the tax-timing option and its interaction with labor income risk. Third, bequests have been identified as important for matching the skewness of the wealth distribution but we feel a more detailed analysis of the retirement decision is needed to introduce this motive. Finally, the general equilibrium implications of introducing TDAs in a heterogeneous agents model with aggregate uncertainty remain to be explored.

Acknowledgments

We would like to thank Joao Cocco, Franck de Jong, Otto van Hemert, Deborah Lucas, Randi Rosenblatt, Chester Spatt, Harold Zhang, Dirk Krueger (the editor), two anonymous referees and seminar participants at the Federal Reserve Board of Governors, Econometric Society meetings in Boston, U of Cincinnati, U of Houston, LSE, UT Dallas, Toulouse, and the NBER Summer Institute Capital Markets and the Economy workshop for helpful comments and discussions. Polkovnichenko acknowledges support from a McKnight Business and Economics research grant at the University of Minnesota and hospitality of the Financial Markets Group at LSE. All errors and omissions are sole responsibility of the authors.

Appendix A. The survey of consumer finances data

The SCF is probably the most comprehensive source of data on U.S. household assets. The SCF uses a two-part sampling strategy to obtain a sufficiently large and unbiased sample of wealthier households (the rich sample is chosen randomly using tax reports). To enhance the reliability of the data, the SCF makes weighting adjustments for survey non-respondents; we used these weights to compute the values reported in the tables. The specific names in the codebook and the acronyms used by the net worth program supplied by the SCF are used below.

We construct a measure of non-financial income to match the process for $Y_{it}$ (earnings) in the model. Non-financial income is defined as the sum of wages and salaries (X5702), business/practice/farm income (X5704), rent and royalties (X5714), unemployment or worker’s compensation (X5716), child support and alimony (X5718), food stamps and welfare
income (X5720), Social Security or other pensions, annuities, or other disability or retirement programs (X5722) and other income (X5724).

We next construct measures of bonds and stocks held in the taxable and tax-deferred accounts. Bonds in the TA are made up of SAVING and MMA (savings and money market accounts), CDs (certificates of deposit), GBMUTF (government bond mutual funds), OBMUTF (other bond mutual funds), BOND (corporate bonds), SAVBND (saving bonds), TFBMUTF (tax free bond mutual funds) and COMUTF (combination mutual funds), for which we assume that half is allocated to bonds. We also include annuities (ANNUIT) and trusts (TRUSTS) that are allocated to bonds. We subtract non-credit card and non-residential real estate debt (variables ODEBT and OTHLOC which include unsecured loans and loans secured by pensions).

Stocks in the taxable account consist of STOCKS (directly held stocks), STMUTF (stock mutual funds), half of COMUTF (combination mutual funds), annuities (ANNUIT) and trusts (TRUSTS) that are allocated to stocks. Bonds (stocks) in the TDA include bonds (stocks) from IRA/KEOGH plans, bonds (stocks) in other future pensions (FUTPen) and bonds (stocks) in account-type retirement plans for which we have information on asset allocation. For example, for the first pension plan of the respondent we require that variable X4216 is less than or equal to 18 (various account-type plans) and X4324 equals 1, 2, 3 or −7 (known asset allocation), and for all the other pension plans of the respondent and spouse the requirement is the same. The shares of wealth invested in stocks in each account are constructed by averaging across the shares of individual households. To construct total taxable wealth, we add bonds and stocks in the taxable accounts, checking and call accounts (CHECKING plus CALL) and subtract revolving credit card debt (CCBAL). To construct total TDA wealth we add bonds and stocks in the TDAs.

Appendix B. Numerical solution

We exploit the scale-independence of the maximization problem and rewrite all variables as ratios to the permanent component of labor income \((P_t)\). The laws of motion and the value function can then be rewritten in terms of these normalized variables, and we use lower case letters to denote them (for instance, \(w_t = \frac{w_t}{P_t}\)). This normalization allows us to reduce the number of state variables to five: liquid wealth in the taxable account, accumulated wealth in the tax-deferred account, current transitory income shock, participation status, and age. The problem is solved (lower case variables normalized by \(P_t\)) as follows. The household needs to decide whether to incur the fixed cost at time (age) \(t\) or not and therefore compares the two value functions associated with direct stock market participation or not:

\[
v_t(w_t^L, w_t^F, u_t, i_t) = \text{MAX}_{0,1}\{v_t(w_t^L, w_t^F, u_t, i_t = 0), v_t(w_t^L, w_t^F, u_t, i_t = 1)\}
\]

where \(i_t = 1\) and \(i_t = 0\) denote direct and indirect participation, respectively. In turn,

\[
v_t(w_t^L, w_t^F, u_t, i_t = 1) = \text{MAX}_{c_t, k_t, \alpha_t^L, \alpha_t^F}\{(1 - \beta)c_t^{1-1/\psi} + \beta \left(\frac{P_{t+1}}{P_t}\right)^{1-\rho} \left(p_t\left[v_{t+1}(w_t^L, w_t^F, u_t, i_t = 1)\right]\right)^{1-\rho} + (1 - p_t)\beta(w_t^{L+1} - 1/\psi)\}
\]

and for indirect stock market participation

\[
v_t(w_t^L, w_t^F, u_t, i_t = 0) = \text{MAX}_{c_t, k_t, \alpha_t^L, \alpha_t^F}\{(1 - \beta)c_t^{1-1/\psi} + \beta \left(\frac{P_{t+1}}{P_t}\right)^{1-\rho} \left(p_t\left[v_{t+1}(w_t^L, w_t^F, u_t, i_t = 0)\right]\right)^{1-\rho} + (1 - p_t)\beta(w_t^{L+1} - 1/\psi)\}
\]

We solve the model recursively backwards starting from the last period. In the last period \((t = T)\) the policy functions are trivial and the value function corresponds to the bequest function.\(^\text{34}\) We need to solve for four control variables in every year: the fraction of taxable wealth being saved, or equivalently, current consumption \((c_t)\), the fraction of the taxable portfolio allocated to stocks \((\alpha_t^L)\), the fraction of retirement wealth allocated to stocks \((\alpha_t^F)\) and the contribution rate \((k_t)\). For every age \(t\) prior to \(T\), and for each point in the state space, we optimize using grid search. From the Bellman equation the optimal decisions are given as current utility plus the discounted expected continuation value \((E_t v_{t+1}(\cdot))\), which we can compute since we have just obtained \(v_{t+1}\). We perform all numerical integrations using Gaussian quadrature to approximate the distributions of the innovations to the labor income process and the risky asset returns. We discretize the state-space along the two continuous state variables and use tensor product splines to perform the interpolation of the value function for points which do not lie on the state space grid, with more points used at lower levels of wealth where the value

\(^{33}\) For annuities, trusts, mutual funds, pension funds and IRA and KEOGH plans in which the allocation between bonds and stocks is mixed, we assume an equal allocation between bonds and stocks.

\(^{34}\) Note that in \(T - 1\) the life expectancy is exactly one year and the agent is required to withdraw all retirement funds and invest them in the taxable account so that \(w_t^F = 0\).
function has high curvature. Equivalently, we use a more dense set of grid points for low values of wealth for the two accounts because the consumption function exhibits a kink at the points where liquidity constraints are no longer binding. Once we have computed the value of each alternative we pick the maximum, thus obtaining the policy rules for the current period. Substituting these decision rules in the Bellman equation, we obtain this period’s value function \( v_t(\cdot) \), which is then used to solve the previous period’s maximization problem. This process is iterated until \( t = 1 \).

### Appendix C. Estimation details

To compute the standard errors of the reported medians in Table 3 we employ a bootstrap procedure based on the 2004 SCF data set, treating all the replicated data from SCF as actual observations. Specifically, for 1000 replications we generate pseudo-datasets by randomly drawing from the specified population (DS or IS for example) with replacement. To draw a particular observation we first generate a uniform random number between zero and one and then pick the observation in the data set that is closest to this draw based on the cumulative distribution of the data that is constructed based on the empirical weights provided by the survey. We repeat this procedure up to 0.8 \( \times N \) where \( N \) is the original sample size (given the sample size our results eventually do not depend on this choice). Based on this sample we compute the median and store the results. Having repeated this procedure for 1000 times, we compute the standard errors of the medians reported in Table 3.

We provide estimates of the structural parameters using Method of Simulated Moments Estimator (MSM) of Duffie and Singleton (1993). The structural parameters \( \hat{\theta} \) are determined as:

\[
\hat{\theta} = \text{Argmin}_{\theta} D' S^{-1} D.
\]

Let \( Y_t \) and \( \tilde{Y}_t \) denote the observations at time \( t \) of the actual and simulated endogenous variables, respectively. Let \( T \) be the sample size of the observed series whereas \( T \cdot H \) data points are simulated to compute moments from the structural model. For the latter, let \( Y_{[T]} \) and \( \tilde{Y}_{[TH]} \) denote the vectors of actual and simulated endogenous variables of length \( T \) and \( TH \), respectively. We have:

\[
D = \left( \frac{1}{T} \sum_{i=1}^{T} \text{median}(Y_t) - \frac{1}{TH} \sum_{i=1}^{TH} \text{median}(\tilde{Y}_t) \right).
\]

When simulating from the structural model we simulate around twenty times the actual data set.

The asymptotically efficient optimal weighting matrix \( S^{-1} \) equals the inverse of the variance-covariance matrix of the data. Following Appendix B in De Nardi et al. (2006), we use a diagonal weighting matrix for \( S^{-1} \) with the elements along the diagonals being the variance of each median, which have been computed using the previously described bootstrap procedure. To compute the standard errors we need to be able to compute the derivative of \( D \). Usually this is achieved using a numerical derivatives method but unfortunately in this instance the derivative cannot be computed using this procedure because of the presence of non-differentiabilities in the moment conditions. De Nardi et al. (2006) show how this difficulty can be sidestepped and the derivative can be written as a product of the derivative of the median and the density function of wealth evaluated at the median. We use a non-parametric density estimator to evaluate the conditional (on age) probability density function of assets at the median.

Given the significant computational complexity of the model we optimize using grid search. It is also important, when conducting the estimation for the DS, to exclude any households happen not to pay the fixed cost and thus do not invest in stocks in their TA for the calculation of the medians. In other words, we might have some non-stockholders or IS, which have the same preference parameters as the DS, but naturally we do not want to use their wealth profiles to compute the moments implied by model, as otherwise those would not be comparable with the moments in the data (where we only consider DS).\(^{35}\)

### Appendix D. Three-period model

To understand the intuition for the impact of TDA accounts we first build a simple 3-period model which can be solved quickly in a spreadsheet. The model includes key elements from our structural model and is used to investigate consumption and savings changes from introducing TDA account. There are 3 time periods \( t = 0, 1, 2 \): young, middle-aged and retirement. Households have CRRA utility \( U(c) = \frac{c^{1-\psi}}{1-\psi} \) and maximize

\[
U = u(c_0) + E_0 \left\{ \beta u(c_1) + \beta^2 u(c_2) \right\}.
\]

\(^{35}\) This also raises the important issue of what percentage of non-stockholders and IS should we allow for a given preference specification that can be consider as admissible in the estimation. We decide to consider a 10% cutoff, implying that any combination of preference parameters for which 90% or more of the households in the model are indeed DS is accepted, while all others are discarded.
Labor income and social security payments are as follows:

\[ y_0 = 1, \quad y_1 = (1 + g)\tilde{\eta}, \quad y_2 = \lambda y_1 \]

where \( g \) is the growth rate of income from young to middle-aged period, \( \tilde{\eta} \) is a random income shock and \( \lambda \) is the income replacement ratio at retirement. Households can save in one asset with deterministic return \( r \) and may do so in the regular taxable account and in a tax-deferred account. Contributions in the TDA are not taxable but the entire account balance must be withdrawn at time \( t = 2 \) and is subject to labor income tax. Tax rate on labor income is \( \tau \) and tax on interest income is \( \tau_d \). Interest income tax on earning in taxable account is paid each period. Denote as \( s_t^T \) and \( s_t^R \) the amount deposited at time \( t \) in taxable and retirement accounts respectively. The budget constraints are as follows:

\[
\begin{align*}
    c_0 + s_0^T &= (y_0 - s_0^R)(1 - \tau), \\
    c_1 + s_1^T &= (y_1 - s_1^R)(1 - \tau) + s_0^T (1 + r(1 - \tau_d)), \\
    c_2 &= y_2(1 - \tau) + s_1^T (1 + r(1 - \tau_d)) + (1 - \tau)(s_1^T (1 + r) + s_0^R (1 + r)^2).
\end{align*}
\]

For consistency with the full model we do not allow withdrawals from, and borrowing against, the TDA and also preclude borrowing in taxable account:

\[
s_t^T \geq 0, \quad s_t^R \geq 0.
\]

The solution of the model can be written as a system of non-linear equations and inequalities, depending on whether borrowing and withdrawal constraints are binding.\(^{36}\) Because the inequalities are binding in some states (e.g. in the low income state) for the relevant parameterizations of the model, we solve the model in a spreadsheet using solver tool and use it to analyze comparative statics across the cases with and without access to the tax-deferred account.

We parameterize the model consistent with our calibrations and estimation results from previous sections. We adjust all parameters appropriately for the 3-period structure by assuming 20 year periods. We set \( \beta = 0.987520, \quad r = (1 + 0.02)^20 - 1, \quad g = (1 + 0.2)^20 \) and the volatility of income shock \( \sigma_y = 0.15\sqrt{20} \). We model income shock as a two-state random variable \( \tilde{\eta} = 1 \pm \sigma_y \). The tax rates on labor and interest income are set at 25%. We vary the EIS parameter from \( \psi = 1 \), the case of log utility, to \( \psi = 0.2 \) to span the estimates we obtained in the paper.

The results are presented in Table 9. In the case of log utility, consumption is lower by 2% and 1.4% in the first and second periods respectively and higher by 9.8% in retirement, when the agent has access to the TDA. Since income is the same across the two models, it follows that savings are higher in this case. This also shows in wealth accumulated at period \( t = 1 \) which is 21.5% higher when TDA is available. For the value of EIS \( \psi = 0.6 \) which we estimated for our group of direct stockholders, the effect is still there although smaller. As we decrease the EIS, the effect of TDA on savings during working life is reversed. For values of the EIS of \( \psi = 1/2 \) and lower, the consumption during working life is higher when the TDA is available.

References


\(^{36}\) In the unconstrained case there are four first order conditions determining savings in taxable and tax deferred accounts in periods \( t = 0, 1 \). However when some constraints are binding, the respective first order condition becomes an inequality instead.


