Should the government be paying investment fees on $3 trillion of tax-deferred retirement assets?*

Mattia Landoni† and Stephen P. Zeldes‡

First version: March 26, 2016
This version: September 7, 2019

Internet Appendix available at https://ssrn.com/abstract=3138303

Keywords: asset management, fees, taxes, retirement savings
JEL: D14, G11, G23, G28, H21, J26, J32

*The authors would like to acknowledge the helpful comments of Dan Bergstresser, Patrick Bolton, Emiliano Catonini, Kent Daniel, Philippe d’Astous, Michael Halling, Glenn Harrison, Daniel Hemel, Charles Jones, David Laibson, Pierre-Carl Michaud, Brett Myers, Emi Nakamura, Thomas Philippon, Jim Poterba, Jon Steinsson, Simon Straumann; seminar participants at the Columbia Business School Finance Free Lunch, the Columbia Macro Lunch, CEIBS, Baruch College, Texas A&M Mays Business School, the Federal Reserve Bank of Chicago, EIEF, Yonsei University, and the Federal Reserve Bank of Boston; and conference participants at the Red Rock Finance Conference 2016, ESSF M Gerzensee 2017, the World Finance Conference 2017, DCIIA 2017, the EUROFIDAI Paris December 2017 meeting, and the European Finance Association 2018 meeting. We thank Abdullah Al-Sabah and Matt Hochhauser for research assistance. Stephen Zeldes is an external advisor at FeeX.com, and is grateful to Eyal Halahmi, Yoav Zurel, and the rest of the team there for their help understanding and measuring investment management fees.

†Assistant Professor of Finance, Edwin L. Cox School of Business, Southern Methodist University
‡Benjamin M. Rosen Professor of Finance and Economics, Columbia Business School, Columbia University; and NBER
Should the government be paying investment fees on $3 trillion of tax-deferred retirement assets?

Abstract (100 words) – In a standard benchmark, both individuals and the government are indifferent between traditional tax-deferred retirement accounts and “front-loaded” (Roth) accounts. We add investment fees to this benchmark and show that individuals are still indifferent but the government is not. By deferring tax revenue through traditional accounts, the government increases demand for asset management services by $3 trillion and pays $19.5 billion in corresponding annual fees. We construct a general equilibrium model with imperfect competition in the asset management industry and examine the incidence and welfare implications of the added demand. In our model, tax deferral produces a larger asset management industry, lower tax revenue, and lower social welfare.

Abstract – Governments incentivize retirement saving by allowing individuals to contribute to tax-advantaged accounts in which the returns to financial assets receive special tax treatment. In accounts with “back-loaded” taxation, the individual contributes pretax money and pays taxes when the money is withdrawn. In accounts with “front-loaded” taxation, the individual contributes aftertax money and pays no future taxes. Under some simplifying assumptions, a standard benchmark result is that each account type results in the same cash flows for the individual, and the same present value of tax revenue for the government, even though back-loaded taxation results in more assets under management (AUM). We add investment management fees to the benchmark model and show that the equivalence result breaks down. Assuming fees are fixed as a percent of AUM, we show that individuals are still indifferent to the timing of taxation but the government is not. Under back-loaded taxation, the government implicitly owns a share of all retirement accounts and is effectively paying investment fees on this share, something it avoids under front-loaded taxation. By deferring tax revenue through traditional accounts, the government increases the demand for asset management services by $3 trillion and pays $19.5 billion in corresponding annual fees. We construct a general equilibrium model with imperfect competition in the asset management industry and examine the incidence and welfare implications of the added demand. The answer depends both on the nature of the cost function for asset management services, and on the nature of market competition. In our model, tax deferral produces a larger asset management industry, lower tax revenue, and lower social welfare.
1 Introduction

Retirement savings systems around the world incorporate tax incentives designed to increase saving and enhance retirement security. The traditional and most common incentive system is tax deferral: the U.S. alone has $23.5 trillion of tax-deferred retirement assets in both employer-based accounts (including defined-benefit plans and defined-contribution plans, or “401(k)s”) and individual accounts (“IRAs”). A tax deferral system works by back-loading taxation, i.e., exempting contributions to retirement accounts from current income taxation and then taxing the principal and returns upon withdrawal.

Although the vast majority of retirement assets is held in traditional accounts, an alternative system in which taxes are front-loaded is becoming increasingly widespread. Under this type of plan, referred to as “Roth”, contributions are made with after-tax income, but then neither the principal nor returns are taxed at any point in the future. During the formulation of the 2017 U.S. tax reform, Congress considered including provisions for “Rothification”, i.e., a shift away from traditional accounts in favor of front-loaded taxation (see, e.g., Tergesen and Rubin, 2017). Although these provisions were not included in the final tax reform law (Public Law 115-97 of 12/22/2017), the debate raised public awareness about the alternative ways of structuring retirement tax incentives.

Much of the U.S. debate focused on the political economy aspects of the choice, i.e., whether front-loading tax revenue with a Roth system would encourage irresponsible fiscal policy. We abstract from this debate. Instead, our contribution is to highlight another important channel through which the timing of taxation affects welfare outcomes: investment funds, recordkeepers, and financial advisors (“asset managers” for brevity) charge fees. These fees are typically proportional to assets under management (AUM). By deferring tax revenue with a traditional system, the govern-

---

1In the U.S., Roth accounts, named after the U.S. senator who originally proposed them, were made available as an additional option via Roth IRAs in 1997 and Roth 401(k)s in 2001. Roth IRA contributions have been higher than traditional ones since 1999, and Roth assets have climbed to about 9% of total IRA assets. Similar trends are occurring in Canada and the U.K., who started with tax-deferred accounts and later introduced front-loaded ones as an additional option. In 2015, the U.K. Treasury launched a formal consultation on, among other things, whether one or the other system is preferable (Osborne, 2015; Buttonwood, 2015).
ment generates an additional $3 trillion in assets under management corresponding to the amount in retirement accounts that will be used to pay future taxes. We estimate that these asset management services cost the government $19.5 billion annually in fees. We argue that the services create little value for the government and therefore we refer to the corresponding fees as an “implicit subsidy” to the financial industry. A traditionally-defined subsidy drives a wedge between the prices received by producers and those paid by consumers, whereas the “subsidy” here takes the form of artificial demand that creates a shift in the demand curve and helps producers cover their fixed costs. Further, we solve a general equilibrium model in which asset management services are represented as differentiated products and show that tax deferral produces a larger asset management industry, lower tax revenue, and lower social welfare.

Our findings have important implications for the recent literature on the growth and the optimal size of finance. Recent evidence shows an upward secular trend in the size of the asset management industry as a fraction of gross domestic product (Philippon and Reshef, 2012; Greenwood and Scharfstein, 2013), largely driven by increased asset valuations. However, the increased scale did not result in lower percentage fees (Malkiel, 2013; Philippon, 2015). Existing explanations for these facts rely on time variation in regulation, moral hazard, or information frictions (Philippon and Reshef, 2012; Bolton et al., 2016). We propose a simple alternative mechanism: because of monopolistic competition, enhanced scale could result in firm and product proliferation instead of lower prices. Further, we show one additional mechanism—tax deferral—that further contributes to increase the scale of the finance industry and in our model has clear welfare consequences: it makes the industry too large. The implications of this mechanism go beyond retirement accounts, because the tax code grants tax deferral treatment to other forms of income, such as capital gains and deferred compensation. However, we focus on retirement accounts because they are the largest and cleanest instance of tax deferral. The growth of tax-deferred retirement accounts, and of the resulting implicit subsidy, alone may be responsible for as much as one-seventh of the relative growth in traditional asset management since 1980.

Section 2 begins with a standard benchmark equivalence result. Under a few simplifying assumptions, including the constancy of the tax rate across working and retirement years, and in
Figure 1: **Benchmark equivalence result.** Traditional is equivalent to a Roth account plus an implicit government account equal to the balance of deferred taxes. Without asset management fees, the two accounts yield the same cash flows for individuals, and the same future value of cash flows for the government. A 150% return is approximately equal to the total return on a 30-year Treasury bond (3.10% for 30 years).

The first two panels of Fig. 1, labeled Roth and Traditional, provide a simple numerical example of this equivalence. The third panel, labeled “Traditional (As If),” provides intuition for the equivalence result by decomposing a Traditional account into a Roth-like individual account plus an implicit government account equal to the balance of deferred taxes. An individual with $100 in a tax-deferred retirement account who faces a 30% tax rate in retirement could be seen as owning $70 in a Roth-like account, with the government owning the remaining $30.

We add one additional realistic element to this benchmark model: fees paid on retirement accounts. We show that individuals remain indifferent between the two types of account, but the fees paid on retirement accounts can affect the equivalence.
Figure 2: **Fee nonneutrality.** Traditional is equivalent to a Roth account plus an implicit government account equal to the balance of deferred taxes. With asset management fees, the two accounts yield the same cash flows for individuals, but the Roth account yields a superior future value for the government thanks to lower total fees. A 150% return is approximately equal to the total return on a 30-year Treasury bond (3.10% for 30 years); 20% fees are the future-value equivalent of 0.80%/year (the value we calibrate in this paper) for 30 years.

The U.S. government’s implicit account is large. In Section 4, we estimate its size as the total amount of tax-deferred assets in defined-contribution (DC) plans and IRAs ($15.4 trillion) times
20%, our rough estimate of the average tax rate on retirement account withdrawals, leading to our
title figure of $3 trillion of retirement assets.\(^3\) We conservatively estimate overall asset-weighted
investment fees (including asset-level fees, account-level fees, individual-level fees and implicit
trading costs) to be about 80 basis points (bps). We assume that 21% of fees paid by the gov-
ernment are recovered via corporate taxation of the asset managers. Multiplying $3 trillion by
\[0.80\times(1-0.21)\], we reach our estimate of $19.5 billion per year—a cost for the government, and
an annual subsidy to the asset management industry in the form of extra demand for asset manage-
ment services.

We produce our estimate of investment fees by adjusting and combining a disparate range of in-
dustry and academic estimates. To our knowledge, it is the first estimate that is both comprehensive
and asset-weighted. In order to estimate the total investment fees incurred by the average dollar
invested in DC plans and IRAs, we construct our own original estimate of explicit advisory fees
on IRAs and convert existing estimates of account-level fees, asset-level fees, and trading costs to
asset-weighted. Most published estimates of mutual fund fees and overall net-of-fees performance
are equal-weighted at the fund level in order to measure average fund quality or average manager
skill. Similarly, most published estimates of trading costs are volume-weighted in order to measure
trader execution skill or overall market liquidity. However, weighting by assets is necessary for
our purposes because lower-cost funds tend to have more AUM (Hubbard et al., 2010).

Our discussion so far has abstracted from features of Traditional accounts that could potentially
be beneficial to the government. If benefits exist, they should be subtracted from the cost. We de-
scribe the potential direct benefits to the government of the asset management services provided
on the implicit government account, but we are skeptical about their importance. For instance, it is
possible that some of the additional fees paid are compensation for services provided by asset man-
gers, such as customized asset allocations (e.g., target-date funds or industry funds) and possibly
better performance. While these services create value for individuals, we argue that many of the po-

\(^3\)Our estimate of assets excludes the $0.5 trillion in the federal government’s Thrift Savings Program (TSP), whose
fees are negligible. It also excludes defined-benefit (DB) plans, although a parallel argument applies to these plans
as well. Including corporate and state and local government DB plans would add $7.5 trillion of tax-deferred money,
increasing our estimate of the implicit government account by 50%.
potential benefits cancel out in the aggregate for the government, which implicitly holds a fraction of all retirement portfolios. Traditional accounts could also potentially bring indirect benefits. First, roughly two-thirds, or $2 trillion, of the U.S. government’s implicit portfolio under Traditional is invested in equities, thus giving the government substantial exposure to the higher risk and higher expected return of the stock market. We argue, similar to the literature on Social Security Trust Fund equity investing, that the added government exposure under Traditional would be beneficial only if the government desires it and cannot obtain it in a less expensive way. Second, active asset managers may create a positive externality if they make prices more efficient and do not fully capture through trading the value they generate (Grossman and Stiglitz, 1980; Greenwald and Stiglitz, 1986). Since a substantial fraction of tax-deferred assets are managed passively, Traditional accounts may only work as a blunt tool to subsidize price discovery because the extra assets benefit active and passive managers indiscriminately. Fourth, issuing government debt and transferring the proceeds to private agents who would otherwise borrow (or want to borrow) through private markets has been shown in some contexts to lower the costs of intermediation and be efficient (Barro, 1974; Woodford, 1990; Aiyagari and McGrattan, 1998; Heathcote, 2005). In our context, however, the transfer is instead saved in retail retirement accounts generating substantial investment costs, so debt issuance adds to intermediation costs and is inefficient.

Our partial equilibrium results from Section 2 show that a hypothetical switch from Roth to Traditional results in increased demand for asset management services, and our results from Section 4 quantify the cost paid by the government for these services. These results, however, leave many questions open: what is the right amount of asset management services? Does the cost to the government translate into cheaper services for individuals saving for retirement or higher profits for asset managers? To answer questions about equilibrium fees, profits, employment in the asset management industry, and welfare, we turn to a general equilibrium analysis.

4Moreover, it is unclear how effective retail investment funds are at price discovery. Gârleanu and Pedersen (2018) argue that in a noisy market for asset management services small investors should index and active management should be left to large investors who have the resources to screen managers.

5Another source of the government’s comparative advantage in borrowing is private agents’ demand for safe and liquid assets (Holmström and Tirole, 1998; Krishnamurthy and Vissing-Jorgensen, 2012). However, tax deferral creates both supply (new government borrowing) and demand (new private wealth).
In Section 5 we present a two-period, general equilibrium model with fixed costs of asset management that generate increasing returns to scale. Absent other frictions, economies of scale would lead to a monopoly, which is inconsistent with empirical evidence on the large number of firms. We therefore model competition among asset management firms (“funds”) as monopolistic competition with differentiated products and free entry (Salop, 1979). In this model, a switch from Roth to Traditional continues to increase assets under management. This increase in assets does not result in lower fees, even in the extreme case in which each fund only needs a fixed amount of labor to operate, and therefore could costlessly expand to manage the additional assets. Instead we find that, with logarithmic utility, total equilibrium fees are a constant fraction of assets under management and there is an increase in (i) the aggregate dollar fees collected, (ii) the equilibrium number of funds, and (iii) employment in the asset management industry.

Finally, we examine how the larger asset management industry that arises under Traditional accounts affects social welfare in our model, defined as the aggregate utility of all individuals, by comparing the quantities that arise in the market equilibrium to those chosen by a planner. We find that the Roth equilibrium in the model has too many funds (i.e., too many resources devoted to asset management) due to the “business-stealing” effect (Mankiw and Whinston, 1986), and that a shift from Roth to Traditional further increases the number of funds and thus lowers social welfare in our model.

Our results have implications for public policy questions related to retirement saving. The primary question is whether it is appropriate for the government to mandate, subsidize, or otherwise encourage a shift towards Roth accounts. Although we point out one important difference between Roth and Traditional retirement accounts, there are potentially other important factors not captured by our model that could affect the relative desirability of the two types of account, such as progressive taxation, behavioral biases, and political economy considerations. Potential alternative policies include attempting to reduce the overall level of fees, either directly, by leveraging the government’s position as a stakeholder, or indirectly, via stricter fiduciary standards for retirement savings accounts. Section 6 addresses these public policy issues and concludes.
2 The benchmark model and the impact of fees

In this section we begin by describing the standard equivalence result (e.g. Brady, 2012) that, assuming proportional (flat) taxation and no time variation in the tax rate, optimizing individuals under Roth can and will choose the same consumption allocation (both during work life and during retirement) as they would under Traditional. In addition, the present value of government revenue is identical under Roth and Traditional.

Next, we add fees to the model, assuming that asset managers provide their services in exchange for a percentage fee \( f \), charged as a fixed proportion of account size. Under this assumption, we show that the equivalence result remains for individuals but breaks down for the government due to its extra spending on asset management services under Traditional. Our fixed-fee assumption can be interpreted in several ways. For instance, our result could hold exactly if the total cost of producing asset management services were proportional to aggregate assets and firms set prices proportional to marginal costs. Another possibility is that the economy is a small open economy, and any changes in the assets are immaterial for the scale at which the asset managers operate.\(^6\) In Section 5, we relax this assumption and examine a general equilibrium model.

2.1 Base assumptions and notation

Money earned and saved for retirement can be taxed at three points: when earned (labor income), as returns are generated (investment income), and when paid out of the account in retirement (account withdrawals). We assume three proportional tax rates, one for labor income \( (\tau_L) \), one for intermediate investment returns \( (\tau_I) \), and one for retirement income \( (\tau_R) \). Tax rates do not vary, either across time or with the level of income.\(^7\)

\(^6\)For instance, suppose that the same asset managers operate in countries A and B, but A’s retirement assets are $99 and B’s retirement assets are $1, and total fees are $1 or 1%. Next, suppose that B’s assets double to $2, but total dollar fees remain constant at $1. In this scenario, percent fees only drop to 0.99%, and almost all the economies of scale accrue to the investors of country A.

\(^7\)In practice, the tax system is instead progressive (i.e., marginal tax rates increase with income), so that even if the tax rate schedule is constant over time, a lower level of income in retirement would imply \( \tau_L > \tau_R \). This is our main motivation for allowing distinct tax rates. Progressivity also introduces additional complications: when coupled with uncertain labor income or asset returns, marginal tax rates become stochastic. Moreover, individuals could then prefer
Table 1: Different tax treatment of retirement savings.

<table>
<thead>
<tr>
<th>Account type</th>
<th>Label</th>
<th>Type of taxation</th>
<th>Tax on initial contribution</th>
<th>Tax rate on investment returns</th>
<th>Tax on retirement payouts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taxable</td>
<td>TTE</td>
<td>Front-loaded</td>
<td>$\tau_L$</td>
<td>$\tau_I$</td>
<td>0</td>
</tr>
<tr>
<td>Traditional</td>
<td>EET</td>
<td>Back-loaded</td>
<td>0</td>
<td>0</td>
<td>$\tau_R$</td>
</tr>
<tr>
<td>Roth</td>
<td>TEE</td>
<td>Front-loaded</td>
<td>$\tau_L$</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 1 compares three possible account types: Taxable, back-loaded taxation (Traditional) and front-loaded taxation (Roth). Based on the timing of taxation, these accounts are also conventionally referred to as TTE, EET, and TEE, respectively. For instance, a common taxable account is TTE because earned income is taxable, investment returns are taxable, but account distributions in retirement are exempt.

2.2 Benchmark: equivalence between front-loaded and back-loaded taxation

We assume that an individual’s savings are held in an account and invested in the one asset in positive supply, government bonds, paying a known return of $r$. We hold government spending fixed, assume that the individual has the same pretax labor earnings under either account type, and abstract for now from details such as contribution limits, withdrawal penalties, and required having both Traditional and Roth accounts, because their simultaneous existence allows people to take advantage of temporarily low marginal tax rates by timing withdrawals, converting to Roth, and/or (until recently) recharacterizing a Roth conversion. We briefly discuss progressive taxes in the conclusion.

The EET/TEE notation is standard in publications by the World Bank (Holzmann and Hinz, 2005; Whitehouse, 2007) and the OECD (Antolín et al., 2004; OECD, 2015a). We follow the World Bank (Holzmann and Hinz, 2005) and Beshears et al. (2017) in using the terms “front-loaded” and “back-loaded” to refer to the timing of the taxation. A source of potential confusion is the earlier use of the terms “front-loaded” and “back-loaded” to refer to the timing of the tax break. Since the tax break for Roth accounts does not occur upfront, some of those involved in the discussion of the 1997 law that introduced Roth accounts referred to them as “back-loaded IRAs” (Committee on Finance of the U.S. Senate, 1997). Several authors including Thaler (1994) and Burman et al. (2001) follow this latter convention instead.
minimum distributions.\footnote{In practice, Roth accounts have less restrictive effective contribution limits (Burban et al., 2001) and fewer restrictions on withdrawals. We also ignore for now any behavioral factors that could cause individuals to choose a consumption plan that differs from the optimum computed here under one or both accounts. We discuss some of these factors in the conclusion.} We also assume that the tax rates are the same under Traditional and Roth (i.e., $\tau_{L}^{Roth} = \tau_{L}^{Trad} = \tau_{L}$), an assumption we relax later in our general equilibrium analysis. We normalize the initial account contribution under Traditional to be $1$ which, because it is tax-deductible, can be funded with $1$ of pre-tax income.

Table 2 shows the initial and future cash flows for both the individual and the government. With a Traditional account, the government has no revenue upfront, and the individual’s account balance is $1$. At time $T$, when the individual retires and the account is liquidated, the balance, $(1 + r)^T$, is paid out and taxed. The government receives a fraction $\tau_{R}$ of the balance, and the individual receives the remaining fraction $1 - \tau_{R}$.

For the Roth account, assume for the moment that the individual contributes $1 - \tau_{L}$, the amount that remains after paying taxes on $1$ of pre-tax income (i.e. the contribution amount such that initial consumption would be the same as under Traditional.) No subsequent taxation happens, and therefore at time $T$ the individual keeps the entire balance, $(1 - \tau_{L})(1 + r)^T$. Consider first the case in which $\tau_{R} = \tau_{L}$. In this case, the individual’s ending wealth is the same under both account types. This means that for \textit{any} initial consumption, the corresponding ending wealth (and thus retirement consumption) will be the same under Roth as it would be under Traditional. Since the set of all feasible consumption paths is the same under Roth and Traditional, it must be the case that the individual’s \textit{optimal} consumption plan (i.e. both initial and retirement consumption) will be the same under the two systems.\footnote{Define $Y$ as initial income, $C_0$ as initial consumption, and $C_1$ as final consumption (equal to terminal wealth). Then for any $C_0$, $C_1^{Roth} = [Y(1 - \tau_{L}) - C_0](1 + r)^T$, and it is easy to show that $C_1^{Trad} = [Y(1 - \tau_{L}) - C_0](1 - \tau_{R})/ (1 - \tau_{L})(1 + r)^T$. Define $S$ as initial saving (also account size). Note that for any $C_0$, $S^{Roth} = Y(1 - \tau_{L}) - C_0$, and $S^{Trad} = [Y(1 - \tau_{L}) - C_0]/(1 - \tau_{L})$, so that $(S^{Trad} - S^{Roth}) / S^{Trad} = \tau_{L}$. If $\tau_{R} = \tau_{L}$, then for any choice of $C_0$, $C_1^{Trad} = C_1^{Roth}$.}

The government’s cash flow differs across plans—with Roth accounts, revenue $\tau_{L}$ is received up front, whereas with Traditional accounts, the revenue is deferred with interest until the future $((1 + r)^T \tau_{L})$. However, assuming that the government discount rate is equal to the interest rate on
If $\tau_R = \tau_L$

<table>
<thead>
<tr>
<th>Account</th>
<th>Initial balance</th>
<th>Future balance</th>
<th>Final payout</th>
<th>Initial revenue</th>
<th>Future revenue</th>
<th>PV @ r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional</td>
<td>1</td>
<td>$(1 + r)^T$</td>
<td>$(1 + r)^T \cdot \frac{1 - \tau_L}{1 - \tau_R}$</td>
<td>0</td>
<td>$(1 + r)^T \cdot \tau_R$</td>
<td>$\tau_R$</td>
</tr>
<tr>
<td>Roth</td>
<td>$1 - \tau_L$</td>
<td>$(1 - \tau_L) \cdot \frac{1}{(1 + r)^T}$</td>
<td>$(1 - \tau_L) \cdot \frac{1}{(1 + r)^T}$</td>
<td>$\tau_L$</td>
<td>0</td>
<td>$\tau_L$</td>
</tr>
<tr>
<td>Traditional - Roth</td>
<td>$\tau_L$</td>
<td>$\tau_L \cdot \frac{1}{(1 + r)^T}$</td>
<td>$(1 + r)^T \cdot \frac{1}{(\tau_L - \tau_R)}$</td>
<td>$-\tau_L$</td>
<td>$(1 + r)^T \cdot \tau_R - \tau_L$</td>
<td>$\tau_R$</td>
</tr>
</tbody>
</table>

Table 2: **Benchmark cash flows under Traditional and Roth.** With flat taxes, and assuming that the tax rate on retirement income ($\tau_R$) is the same as the tax rate on labor income ($\tau_L$), the individual has the same retirement wealth both with a Traditional and a Roth account. Government revenue is also constant in present value, assuming that the government’s discount rate is the same as the return on government debt ($r$). Government bonds, the time-0 present value of revenue under Traditional is $\tau_L$, i.e., the same as the immediate revenue under Roth. The government will therefore be indifferent in a present value sense between the accounts. We refer to the two systems as equivalent, since individual consumption patterns will be the same under each, and the government will be indifferent in a present value sense (issuing more government bonds under Traditional at time 0, but collecting all the revenue necessary to cover interest and principal at time $T$).

Although we wait until Section 5 to discuss general equilibrium results, we note here that the equivalence result above would continue to hold in general equilibrium. In other words, a shift from Roth to Traditional would leave the equilibrium interest rate and all consumption allocations unchanged.\footnote{For every dollar in a Traditional account, the government faces a revenue shortfall relative to Roth of $\tau_L$, and so must issue an extra amount $\tau_L$ of bonds, adding to the existing supply. At the same time, the account balance is $\tau_L$ larger relative to Roth. Since the account is invested in government bonds, this creates additional demand for government bonds equal to $\tau_L$. Thus, the higher desired private saving is exactly offset by the lower government saving, leaving desired national savings and the equilibrium interest rate the same under Traditional and Roth.}

Next, consider the case in which $\tau_R \neq \tau_L$. We can now decompose a Traditional balance into
three virtual accounts (rather than the two described in the introduction) as follows:

\[ V_{t}^{Trad} = (1 + r)^t \left[ \left( 1 - \tau_L \right) + \left( \tau_L - \tau_R \right) + \tau_R \right] \]

As before, the first term is a “Roth-like” account of size \(1 - \tau_L\), as if the individual had contributed to a Roth account, and the third term is the “implicit government account” of size \(\tau_R\) that represents the claim the government has on the account due to future taxes when the investor takes distributions from the account. The middle term is new, and represents a “government matching account” of size \(\tau_L - \tau_R\). This account is akin to an employer matching program—the government matches every dollar the individual saves under Traditional with a grant \(\tau_L - \tau_R\), the difference in tax rates between labor income and retirement income. If \(\tau_R < \tau_L\), the match is positive and therefore individuals prefer Traditional to Roth and also face a stronger incentive to postpone consumption. The government, on the other hand, must pay for the match and therefore has a lower PV of revenue under Traditional. If \(\tau_R > \tau_L\), there is a negative implicit match and the results are reversed. Thus, when \(\tau_R \neq \tau_L\), the equivalence result described above no longer holds.\(^{12}\)

### 2.3 Fee nonneutrality in the benchmark model

We now examine the effects of introducing asset management fees. We start with the benchmark model, and simply add the assumption that at time 0 an asset management firm charges fees equal to a fixed proportion \(f\) of account size. We assume initially that the government does not tax the corporate income of asset managers (\(\tau_C = 0\)).

Table 3 shows our results. Both accounts are funded with $1 of pre-tax income and grow at a net-of-fee rate of \((1 + r) (1 - f) - 1\). The left panel of Table 3 calculates the final payouts for the

\(^{12}\)If \(\tau_R \neq \tau_L\), then for any given \(C_0\), \(C_1^{Trad} - C_1^{Roth} = \left[ Y(1 - \tau_L) - C_0 \right] \left( \tau_L - \tau_R \right) / \left( 1 - \tau_L \right) (1 + r)^T\) and so the equivalence will not hold.
individual. Under Traditional, the initial balance, equals 1, and the final aftertax distribution from the account is \((1 + r)^T (1 - f)^T (1 - \tau_R)\). Under Roth, the initial balance is \(1 - \tau_L\), and the final distribution from the account is \((1 - \tau_L) (1 + r)^T (1 - f)^T\).

Consider first the case \(\tau_R = \tau_L = \tau\), a necessary condition for the benchmark equivalence result. For any given initial consumption, while the individual’s final retirement wealth is lower with fees than it was without, it is still the same across Traditional and Roth. Therefore, as before, the feasible sets of consumption plans are identical under Roth and Traditional, and so individuals will choose the same optimal consumption (initial and retirement) under Roth and Traditional and will be indifferent between the two systems.13

Next consider the government. The right panel of Table 3 calculates the present value of tax revenue. Even with asset management fees on retirement accounts, the government discounts cash flows at \(r\), the interest rate it pays on outstanding debt, because this is its marginal opportunity cost of funds. As in the benchmark case, the timing of tax revenue cash flows differs between Traditional and Roth. Unlike in the benchmark case, however, the present value of these cash flows is also different. Under our assumption that \(\tau_L^{Roth} = \tau_L^{Trad} = \tau_L = \tau_R = \tau\), the government has unambiguously lower present value of tax revenue under Traditional:

\[
P\left(\text{Tax Revenue}^{Trad}\right) - P\left(\text{Tax Revenue}^{Roth}\right) = -\tau [1 - (1 - f)^T] < 0.
\]

This formula has an intuitive interpretation: \(\tau\) is the initial size of the government’s implicit account under Traditional, and \(1 - (1 - f)^T\) is the fraction of the account that gets eroded by fees. Note that if the government’s present value budget constraint holds under Roth, it will no longer hold under Traditional, due to the drop in revenue that arises from paying fees. Therefore, unlike in the no-fee case, there will not be an equivalence in general equilibrium—some tax rates will have to be higher under Traditional, which will imply different equilibrium paths of consumption and/or interest rates. We discuss this further in Section 5.

Up to this point we have not considered corporate taxation. We now show that taxing the
Table 3: Present value of tax revenue under Traditional and Roth with fees and no corporate taxes. An asset manager charges proportional fees $f$ on the account. Assuming that the tax rate on retirement income ($\tau_R$) is the same as the tax rate on labor income ($\tau_L$), the individual has the same retirement wealth both with a Traditional and a Roth account. However, government revenue is lower with Traditional, assuming that the government’s discount rate is the same as the return on government debt ($r$).
income of asset managers at a rate $\tau_C$ shrinks, but does not eliminate the difference in present value of government revenue between Traditional and Roth accounts. We focus on the simple but conservative case in which the additional assets under Traditional result in additional fee revenues but no additional costs. Therefore every extra dollar of asset manager revenue turns into an extra dollar of pretax profits and an extra $\tau_C$ dollars of tax revenue.\footnote{If the additional assets result in additional costs, only a fraction of the revenue turns into profits for the asset manager. However, most asset manager costs would equal income for employees (e.g. additional hours worked) or other entities connected with the asset manager (e.g. additional bid-ask spreads and fees paid to a broker-dealer). This income could then be taxed at some rate, possibly higher than $\tau_C$, or escape taxation altogether. We abstract from these complexities.} Then, the government receives a stream of corporate tax revenues growing at a rate $(1 + r) (1 - f) - 1$ (the same rate as the account balance), effectively recapturing a fraction $\tau_C$ of the fees paid on its implicit account:\footnote{For a Traditional account, the present value of corporate tax revenues is equal to $\tau_C [1 - (1 - f)^T]$, i.e., the asset manager’s one-period revenue flow $f (1 + r)$, times the corporate tax rate $\tau_C$ (to obtain the government’s one-period corporate tax revenue flow), times a growing annuity term $1/[f (1 + r)] \cdot [1 - (1 - f)^T]$. Similarly, the present value of corporate tax revenues for Roth is $(1 - \tau_L) \tau_C [1 - (1 - f)^T]$, and therefore a Traditional yields an additional $\tau_L \cdot \tau_C [1 - (1 - f)^T]$ in corporate tax revenues. Adding this term to (2), we obtain (3).}

\[
\text{PV (Tax Revenue}^{\text{Trad}}\text{)} - \text{PV (Tax Revenue}^{\text{Roth}}\text{)} = \nonumber
\]
\[
= -\tau [1 - (1 - f)^T] (1 - \tau_C) < 0. \tag{3}
\]

Summarizing, when $\tau_R = \tau_L$, the investor is still indifferent between Traditional and Roth because both accounts are eroded by fees in equal proportions. However, the government is not indifferent, because fees drive a wedge between its discount rate and the expected return on its implicit account. If the government were to shift from Traditional to Roth, it would receive the revenue upfront and take on lower debt. By instead leaving the money in its implicit account, the government keeps paying an interest rate $r$ on the outstanding debt, but receives a net-of-fees return $(1 + r) (1 - f) - 1 (< r)$.

The government’s loss of tax revenue is not simply a consequence of the fact that in a Roth account fees are paid with after-tax money, whereas in a Traditional account fees are paid with pre-tax money, i.e., “deductible.” To examine this possibility, in Section 2 of the Internet Appendix we consider two alternative hypothetical accounts: a “fee-deductible Roth”, in which individuals
receive a deduction for fees, and a “fee-nondeductible Traditional”, in which the government taxes
the individual based on the gross-of-fees balance. By comparing outcomes among these four
account types we decompose the difference between Traditional and Roth into two components:
fee deductibility, accounting for a fraction $1 - \tau_R$ of the revenue loss, and the sheer existence of
additional assets, accounting for the remainder.

Next we consider the case $\tau_R \neq \tau_L$ (for simplicity returning to $\tau_C = 0$). As in the model
without fees, the individual is not indifferent between Traditional and Roth (preferring Traditional
if $\tau_R < \tau_L$) and the Traditional account can be still decomposed into three virtual accounts: a Roth-
like account, a government matching account, and the implicit government account. The difference
between the PV of government revenue under Traditional and Roth in Equation (3) can be rewritten
as

$$PV \left( \text{Tax Revenue}^{Trad} \right) - PV \left( \text{Tax Revenue}^{Roth} \right) = -\tau_R \left[ 1 - (1 - f)^T \right] - (\tau_L - \tau_R).$$

The first term represents the extra fees that the government pays under Traditional on its implicit
account. The second term is the cost to the government of the incremental match under Traditional
(which is greater in absolute value than the PV of the match the individual receives, due to the extra
fees). The difference between the PV of revenue received by asset managers under Traditional and
Roth equals $\tau_L \left[ 1 - (1 - f)^T \right]$. Asset managers receive additional fee revenue under Traditional
because at the same initial consumption level, the size of the account under Traditional is $\tau_L$ greater
than what it would be under Roth, and they receive fees on this extra amount, regardless of whether
the extra assets are owned by the individual (due to the government match) or the government.

In the presence of both fees and corporate taxes, the difference in the PV of government revenue
is

\[\text{(4)}\]

16 We thank Mariacristina DeNardi for suggesting we examine the possibility of a “fee-nondeductible Traditional”. A situation similar to a “fee-deductible Roth” was possible under the U.S. Internal Revenue Code until 2017 as, subject to limitations, Roth account advisory fees could have been paid using outside money and claimed as a deduction for investment expenses. Such possibility was eliminated by the 2017 tax reform.

17 The value of the match to individuals under Traditional (equal to the present value of the increment to the individual’s retirement wealth under Traditional) is equal to $(1 - f)^T (\tau_L - \tau_R)$. 

16
\[
\text{PV (Tax Revenue}^{\text{Trad}}\text{)} - \text{PV (Tax Revenue}^{\text{Roth}}\text{)} = \\
= -\tau_R \left[1 - (1 - f)^T \right] \left(1 - \tau_C \right) - (\tau_L - \tau_R) \left\{ 1 - \tau_C \left[1 - (1 - f)^T \right] \right\}.
\]

Expression (5) is the same as (4), except for the effect of corporate taxation. The first term is multiplied by a factor \(1 - \tau_C\) because the government recovers a fraction \(\tau_C\) of the fees paid on its implicit account. The second term is the cost of the match multiplied by \(1 - \tau_C \left[1 - (1 - f)^T \right]\), because the government levies corporate taxes on the fees paid by the individual on the match.

### 2.4 Adding a risky asset

So far we have assumed that there is only one asset in the economy: government bonds. In reality, however, roughly two-thirds of the government’s $3 trillion implicit account is invested in stocks, mostly through actively and passively managed stock funds. To analyze this, consider the base case above with \(\tau_R = \tau_L = \tau\), and suppose now that there are two assets: the government bond yielding \(r\) and a risky asset (stocks) that has a stochastic return \(\tilde{r}_s\) and expected return \(r_s > r\). Individuals choose the share \(a\) of the portfolio that is held in stocks.

Holding tax and interest rates constant, and for the moment ignoring fees, it is straightforward to show that the individual would choose the same initial consumption and portfolio share \(a\), and thus identical (risky) future consumption, under Roth and Traditional and would therefore be indifferent between the two. As before, the timing of the government’s cash flows differ between Roth and Traditional, but now the incremental delayed cash flows under Traditional are will be positively correlated with realized stock returns, and thus have both higher expected value \textit{and higher risk} than in the benchmark case.\(^{18}\) A shift from Roth to Traditional is equivalent to an increase in the supply of government bonds, an equal increase in the private demand for government bonds, \textit{and} a portfolio swap in the government’s implicit account that raises the demand for stocks and lowers the demand for bonds.

\(^{18}\)Under Roth, the tax revenue is \(\tau\) at time 0. Under Traditional, the tax revenue at time \(T\) is a random variable equal to \([a(1 + \tilde{r}_s)^T + (1 - a) (1 + r)^T]\). The expected tax revenue is equal to \([a(1 + r_s)^T + (1 - a) (1 + r)^T]\), which is increasing in \(a\).
If the government faces no binding constraints on its holdings of stocks, then the appropriate rate for discounting the expected cash flows from the government’s implicit stockholdings will be $r_s$, and so the present value of tax revenues will continue to be the same under the two systems and the benchmark equivalence result will continue to hold.

The government could be constrained from holding more stocks under Roth, for example due to political constraints as democracies are usually averse to direct government holdings of productive assets (see, e.g., Che and Qian, 1998).\footnote{These constraints are not universal, however, as some governments do hold trillions of dollars of professionally managed risky investments via sovereign wealth funds. An August 2018 estimate by the Sovereign Wealth Fund Institute (https://www.swfinstitute.org/sovereign-wealth-fund-rankings/) places total sovereign wealth fund assets at $8 trillion. Holdings of risky assets vary by fund, but as of Q1 2019 Norway’s Government Pension Fund Global, the world’s largest sovereign wealth fund, held just over two-thirds of its assets in equities (https://www.nbim.no/en/the-fund/).} If the government faced such a constraint, and it were binding, then the value of an extra dollar of stock due to a shift from Roth to Traditional would be based on a discount rate $r'_s$, lower than the equilibrium discount rate $r_s$, for a marginal benefit equal to approximately $r_s - r'_s$.\footnote{If the government is constrained, then its intertemporal budget constraint will no longer be satisfied after a switch from Traditional to Roth or vice versa. If future taxpayers see fully through the government veil, it is possible that a “super” Ricardian equivalence result could hold in general equilibrium. This would require future taxpayers under Traditional to realize that the government will have to change taxes in the future based on realized stock returns up to then. This tax policy would amount to an implicit stock position for individuals, who would reduce their own stock holdings today. In this scenario, demand and supply of stocks and bonds could remain in balance with no change in interest rates, stock expected returns, or household consumption.} However, it is not clear that such a constraint, even if present, would necessarily be binding. The government already has substantial implicit exposure to equities through the tax system due to the positive correlation between tax revenue cash flows and stock market performance, and it may not be advantageous for it to increase this exposure further (Auerbach, 2004). In this scenario, the government could face a binding constraint under Traditional that precluded it from reducing its equity exposure, which would imply $r_s - r'_s < 0$. A full analysis would require a general equilibrium model and a complete specification of the nature of the constraints. The issues that arise here are similar to those in the literatures on whether the Social Security trust fund should invest in equities.\footnote{See, for example, Bohn (1990), Geanakoplos et al. (1999), Abel (2001), Diamond and Geanakoplos (2003), and Lucas and Zeldes (2009). See also the related discussion in Romaniuk (2013).}

How are the above results influenced by the presence of investment fees? In the absence of
constraints, the government could likely replicate the implicit government stock holdings under
Traditional with a much less expensive combination of a Roth system and a sovereign wealth fund.
An investment management firm could manage a largely passive portfolio that could mimic the
holdings of aggregate retirement accounts (or the aggregate stock market) at the likely cost of only
a few basis points. If, however, the government faced political or other constraints on doing this,
then one would need to compare the benefits from the added stock holdings under a Traditional
system (relative to Roth) to the loss from the higher fees paid on the extra assets held in retirement
accounts.

3 Investment fees ($f$)

In this section we construct a comprehensive, asset-weighted estimate of explicit fees (both asset
level and account level) and trading costs for DC plans and IRAs. We rely on existing industry
estimates for explicit fees on DC plans and for explicit asset-level fees on IRAs, construct our own
original estimate of explicit advisory fees on IRAs, and combine academic studies to construct an
estimate of trading costs for DC plans and IRAs. Our overall estimate of $f$ is 80 basis points (bps),
obtained as the asset-weighted average of 67 bps for DC plans and 92 for IRAs. Our results are
summarized in Table 4 and described in more detail below. Additional detail on our estimates is
provided in Section 3 of the Internet Appendix.

3.1 Explicit fees

An individual saving for retirement faces at least two types of explicit fees: asset-level fees and
account-level fees. Asset-level fees are charged based on what financial products the account

---

22 As an example of this, BlackRock currently manages the assets of the Thrift Savings Plan at very low cost.
23 Note that Traditional accounts hold bonds that also pay fees. The entire balance in the government account pays
fees at a rate $f$, but only a fraction $a$ is invested in stocks, for a cost $f/a$. Thus, $r_s - r'_s > f/a$ is required for the
first dollar of additional equity exposure to be beneficial. For instance, with fees equal to 0.80% (our asset-weighted
estimate), $a = 2/3$ (a value close to the actual asset allocation of U.S. retirement accounts) and an expected return
on stocks of 8%, the government’s discount rate on the marginal dollar of stocks would need to be at least 1.2%
($= 0.80 / (2/3)$) lower than the expected return on stocks in order for the benefits of a marginal shift to Traditional
accounts to outweigh the costs.
Panel A. Overall Fee Breakdown (bps).

<table>
<thead>
<tr>
<th>Tax-Deferred Assets</th>
<th>1. Explicit Fees</th>
<th>2. Trading Costs</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC Plans ($7.8 trillion)</td>
<td>All-in fees</td>
<td>50*</td>
<td>17†</td>
</tr>
<tr>
<td>IRA ($9.3 trillion)</td>
<td>Asset-level fees</td>
<td>35*</td>
<td>16†</td>
</tr>
<tr>
<td>Advisory fees</td>
<td>41‡</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Weighted Average (A, B) | 64 | 16 | 80 |

(*) based on asset-weighted industry estimates; (†) estimate derived from volume-weighted academic estimates of trading costs using industry estimates of fund turnover; (‡) original estimate based on SEC 10K filings.

Panel B. Asset-Level Fees and Trading Costs Breakdown (bps).

<table>
<thead>
<tr>
<th>Asset Class</th>
<th>DC Plans</th>
<th>IRA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equity</td>
<td>59%◊</td>
<td>26.4%*</td>
</tr>
<tr>
<td>Bonds</td>
<td>10%◊</td>
<td>7.7%*</td>
</tr>
<tr>
<td>Balanced**</td>
<td>28%◊</td>
<td>10.1%*</td>
</tr>
<tr>
<td>Cash</td>
<td>3%◊</td>
<td>2.5%*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ind.Sec</th>
<th>DC Plans</th>
<th>IRA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equity</td>
<td>17^</td>
<td>29.6%*</td>
</tr>
<tr>
<td>Bonds</td>
<td>27^</td>
<td>7.6%*</td>
</tr>
<tr>
<td>Cash</td>
<td>0^</td>
<td>8%*</td>
</tr>
</tbody>
</table>

| Other | 0^ | 8.1%* | 65† |

Average | 17 | 35 | 16 |

(*) Our estimation based on ICI and EBRI data (Copeland, 2016; ICI, 2018); (†) Assumed same as DC, reported in (BrightScope and ICI, 2018); (^) Assumed negligible; (§) Based on Deloitte and Brightscope/ICI studies (Rosshirt et al., 2014; BrightScope and ICI, 2018); (◊) Based on academic estimates of volume-weighted costs (Anand et al., 2012 and Busse et al., 2018 for equity; Bessembinder et al., 2018, Choi and Huh, 2017, and Goldstein and Hotchkiss, 2018 for bonds) and industry turnover estimates (BrightScope and ICI, 2018 for equity; Rowley and Dickson, 2012 and Novick et al., 2016 for bonds); (◊) Based on ICI research (Collins et al., 2016; ICI, 2018); (¶) Original estimate based on SEC 10K filings information of top 21 IRA providers; (□) assumed same as funds; (**) assumed to be 60% equity and 40% bonds.

Table 4: Overview of our estimates of investment fees and costs.
money is invested in, and include both ongoing fees (mutual fund expense ratios) and one-time fees (front or back-end loads). Some of these fees are paid to asset managers (e.g., mutual fund sponsors, insurance companies, and issuers of structured notes) and some (typically revenue from sales loads and/or 12b-1 fees that are included in the expense ratio) are paid to distribution channels (e.g., mutual fund brokers, securities brokers, and advisors to 401(k) plans).

Account-level fees include account maintenance fees in IRAs and DC plans charged by the recordkeeper or account provider, typically as a fixed dollar amount per account. They also include advisory fees charged by financial advisers for providing expertise to individuals in asset allocation, estate and tax planning, and other services covering one or multiple accounts belonging to the same individual or household. These fees, sometimes also referred to as “wrap” fees, are generally charged as a percent of the total value of advised assets.

Because we care about the total fees paid by the government, it is important that we use asset-weighted rather than equal-weighted estimates. For DC plans (row 1, column 1), we rely on two recent asset-weighted estimates made by industry participants in partnership with the industry trade association, the Investment Company Institute. Both estimates focus on 401(k) accounts, and thus exclude 403(b) and other plan types, and they also focus on the largest and most efficient plans. Deloitte (Rosshirt et al., 2014) estimates the “all-in fee” at 58 bps, while BrightScope (2018) estimates “total plan costs” at 37 bps. Because the BrightScope study excludes about $1 trillion or 27% of total assets held in the smallest, and likely most expensive, plans, we put more what more weight on the Deloitte estimates and use 50 bps as our estimate of total account and asset-management fees.

Next we consider explicit fees on IRAs (row 2, column 1). IRA asset-level fees are estimated as an asset-weighted average by type of product (mutual funds, individual securities, and other) and asset class (stocks, bonds, money markets, and other) based on industry estimates. As shown in Panel B of the table, about 45% of IRA assets is invested in individual securities or cash without.

---

24The BrightScope estimate is based on filings by audited plans, which generally means plans with 100 or more participants. The Deloitte estimate is survey-based excluding plans with less than $1 million in assets and oversampling large plans, yielding representation of roughly 97% of the assets (36% of the plans) within the universe of plans filing Form 5500 with the Department of Labor.
explicit fees other than trade commissions, about 47% is held in mutual funds with asset-weighted average fees of 61 bps, and the remaining 8% is invested in other products like real estate funds, commodity funds, etc. whose fees are estimated at 65 bps. This yields an asset-weighted estimate of asset-level IRA fees of 35 bps.

IRA account-level fees, particularly advisory fees, are challenging to estimate. There is no systematic source of information on what fraction of investors participate in programs of paid advisory services or on the size of the fees paid by those who participate, and to the best of our knowledge no comprehensive estimate of these fees is available. We construct original “supply-side” estimates of account-level AUM-based advisory fees by relying on SEC filings and other publicly available information. Most of the largest IRA providers are exchange-listed discount brokers, mutual fund families, and the retail arms of large banks and insurance companies, who disclose the relevant information in their annual reports (SEC 10K forms). Using industry sources, we identify the top 20 IRA account providers by assets under administration. For 18 of the top 20 (corresponding to 74.9% of total IRA assets and 97.9% of top-20 assets), we are able to estimate percent advisory fees as total revenue from AUM-based advisory programs divided by total client assets, regardless of assets actually enrolled in fee-based advisory programs. We estimate average IRA advisory

---

25 Using ICI data on IRA holdings of mutual funds and EBRI data on overall asset allocation in IRAs (Collins et al., 2016; Copeland, 2016; ICI, 2018), we estimate that 58.5% of IRA assets are held in mutual funds (28.8% equity, 8.4% bonds, 10.7% balanced, and 10.5% money market), 33.4% in individual securities and 8.1% in other investments. For mutual funds held in IRAs, the ICI reports asset-weighted average fees of 66, 49, and 65 bps for equity, bonds and balanced funds respectively. Having no data specifically on money-market funds and “other” investments held in IRAs, we use the corresponding DC plan figures (25 bps and 65 bps respectively) reported by BrightScope (BrightScope and ICI, 2018). The resulting average fees (61 bps) are substantially higher than those of mutual funds held in DC plans (47 bps, estimated with an analogous calculation based on ICI data), reflecting higher distribution fees due to the dispersed nature of individual accounts. The assumption of no fees for individual securities is conservative, because some of these securities may be structured notes whose payoff bears a complex relation to the performance of the underlying assets. Although more complex and more difficult-to-understand financial products need not necessarily have higher prices, a growing literature on shrouded prices (Gabaix and Laibson, 2006; Carlin, 2009; Henderson and Pearson, 2011) suggests that this is typically the case.

26 For instance, if half of the clients pay 100 bps in advisory fees and the other half does not use a fee-based advisor, our methodology produces an estimate of 50 bps. Note that this methodology excludes distribution-related fees (a source of revenue for IRA account providers), which are already counted as part of asset-based fees. Two providers are large, privately held discount mutual fund families. For one, we are able to obtain the relevant information from press reports, and for the other we simply use the average of discount mutual fund families and brokers. Finally, two nonprofit providers are excluded because we are not able to obtain the relevant information.
<table>
<thead>
<tr>
<th>IRA provider type</th>
<th>Market share</th>
<th>Advisory Fee</th>
<th>No.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overall</td>
<td>Within top 21</td>
<td>Revenue (bps)</td>
</tr>
<tr>
<td>Full-service broker or MF family</td>
<td>33.9%</td>
<td>44.2%</td>
<td>58</td>
</tr>
<tr>
<td>Discount broker or MF family</td>
<td>35.6%</td>
<td>46.4%</td>
<td>30</td>
</tr>
<tr>
<td>Insurance company</td>
<td>2.4%</td>
<td>3.1%</td>
<td>141</td>
</tr>
<tr>
<td>Total</td>
<td>71.9%</td>
<td>93.7%</td>
<td>47</td>
</tr>
</tbody>
</table>

Table 5: Estimates of advisory fees in IRAs

Full-service brokers are large banks with a wealth management arm. Full-service mutual fund (MF) families are firms that hold a substantial fraction of assets under management in actively managed mutual funds. Discount brokers are companies that offer a brokerage account and a simple set of tools to manage one’s portfolio, and discount mutual fund families are firms that offer predominantly low-cost index funds.

fees to be 41 bps (see Table 5 for this estimate and more information on its derivation).

3.2 Trading costs

Both mutual funds and individual investors who directly trade individual securities incur trading commissions and also trading costs in the form of bid-ask spreads, defined as the difference between the buy price and the sell price. Trading commissions are not included in explicit fees such as the expense ratio of mutual funds (Livingston and Zhou, 2015). Mutual funds also incur costs due to market impact, defined as adverse price moves caused by one’s trades. Because of their size, unique disclosure requirements, and liquidity needs, mutual funds trade more predictably than other investors; as a result, they can be front-run and face adverse price pressure (Ben-Rephael et al., 2011; Shive and Yun, 2013). Predictability is especially a problem for index funds who trade mechanically to rebalance and incorporate changes in the index (Pedersen, 2018). Market impact and bid-ask spreads are not straightforward to assess even for the fund itself, and therefore they are rarely or never disclosed, but they are reflected in returns.\textsuperscript{27}

We derive our own estimate of trading costs because we are not aware of any asset-weighted

\textsuperscript{27}To further complicate the picture, some broker-dealers offer “soft-dollar” arrangements under which they provide clients (i.e., funds) with services such as research reports in exchange for their business (Conrad et al., 2001; Livingston and Zhou, 2015). These arrangements reduce the funds’ explicit expenses to the detriment of execution quality, and thus muddy the distinction between explicit fees and trading costs.
estimates. We first estimate trading costs at the asset class level, and then we construct separate estimates for IRAs and DC plans using their respective asset allocations. Our estimate is based on the following approximation:

Annual trading costs = Trading costs per unit of volume × 2 × Annual turnover.

Trading costs per unit of volume are measured as explicit commissions plus execution shortfall as a percentage of trade size. Execution shortfall is a standard measure of execution quality, defined as the difference between the actual execution price and a reference price observed at the time the order is placed. Turnover is defined as the lesser of a fund’s gross purchases and sales of securities divided by the fund’s average net assets, so that 2 × Annual turnover is a lower bound to total volume of trading as a fraction of total fund assets.

For stock funds, recent estimates place trading cost at roughly 26 basis points per unit of volume (Anand et al., 2012; Busse et al., 2018). The average turnover of US equity funds (active and passive combined) is 44%, but only 32% for mutual funds held in 401(k) accounts (BrightScope and ICI, 2018) because of a greater prevalence of low-turnover index funds. Assuming the latter figure applies to both DC plans and IRAs, we obtain annual trading costs for equity mutual funds of $26 × 2 × 32% = 17$ bps, lower than typical pronouncements by industry insiders.\(^{28}\)

For bond funds, we conservatively use cost estimates for the largest corporate bond trades. Unlike in the case of stocks, large bond transactions have a lower cost per unit of volume, suggesting that execution shortfall is driven less by market impact and more by search frictions. A recent, comprehensive estimate (Bessembinder et al., 2018, Table III) places transaction costs on corporate bond trades of $5 million and up in the 2012–2014 period at roughly 17 bps of trade size for a round-trip, or 8.5 bps per unit of volume, consistent with other recent works (Choi and Huh, 2017; Goldstein and Hotchkiss, 2018).

\(^{28}\)A managing director for Morningstar (Phillips, 2013) states that in the five years prior to March 31, 2013 “the average U.S. large-cap equity fund, on an asset-weighted basis, trails the market index by its expense ratio plus ... 25 basis points.” Bogle (2014) guesstimates trading costs of 50 bps for active equity funds, and negligible for passive equity funds. Assuming that active funds’ market share is equal to their overall U.S. market share (roughly 50%), Bogle’s figure implies asset-weighted trading costs of roughly 25 bps.
Assessing turnover for bond funds is less straightforward than for stock funds. Vanguard (Rowley and Dickson, 2012) estimates that the asset-weighted average portfolio turnover of open-end bond funds ranges from 90% for index funds to 193% for active funds. Using AUM figures from Blackrock (Novick et al., 2016) we calculate asset-weighted average turnover of 178%, reflecting the predominance of active funds. We multiply this turnover by 1.8, rather than by 2, to reflect the fact that some reported turnover may be driven by reinvestment of coupon and principal payments, rather than sales and purchases. Thus, our estimate of trading costs for bond funds is $8.5 \times 1.8 \times 178\% = 27$ bps.

Our final, asset-weighted estimate of trading costs assumes 17 bps for equity funds, 27 bps for bond funds, and 0 bps for money market funds, own-company stock and other investments. For individual securities, in the absence of data, we assume the same trading costs in each of the three corresponding asset classes. Based on the overall asset allocation in DC plans and IRAs, we estimate total asset-weighted implicit trading costs of 17 bps and 16 bps respectively.

4 Calibration: the subsidy to asset managers under Traditional

In a Traditional system, the government owns an implicit account of size $S \cdot \tau_R$, where $S$ is the aggregate amount of tax-deferred retirement savings, and $\tau_R$ is the effective tax rate on retirement payouts. This implicit account incurs investment fees (explicit fees plus trading costs) at an annual average percentage rate $f$. We assume that the government recovers a fraction $\tau_C$ of these fees.

---

29 When selling a security and purchasing another, a fund trades twice; when reinvesting a coupon or principal payment, a fund trades only once. If reported turnover is defined as the lesser of purchases on one hand, and sales plus issuer payments on the other, the “2” coefficient would overstate trading costs. In practice, however, many bond funds and indices have rules that cause them to sell bonds before maturity. For instance, “most flagship Bloomberg Barclays Aggregate, High Yield, Inflation-Linked and Emerging Markets Indices have a minimum time to maturity” (Barclays, 2017). In particular, the most widely followed bond index (the Barclays Capital US Aggregate Index) has annual turnover is 42.0%, of which only about one-fifth is due to coupon and principal paydowns (Tucker, 2011). Accordingly, our adjusted coefficient counts one-fifth of the volume once instead of twice ($2 \times \frac{4}{5} + 1 \times \frac{1}{5} = 1.8$). To the extent that any funds already report turnover net of issuer payments, our adjustment is conservative. In principle, a similar adjustment should apply to stocks as well, but it would be minimal because of the lack of principal paydowns and an average dividend yield of less than 2%.

30 In DC plans, 69% of assets is allocated to equity funds, 20% to bond funds, and 11% to other investments (Collins et al., 2016). In IRAs, 62% is allocated to equity, 19% to bonds, and 19% to other investments (Copeland, 2016).
via corporate taxation of the profits of asset managers and other intermediaries. The unrecovered fraction amounts to an implicit subsidy, i.e., a transfer from the government to asset managers that would not exist under Roth.

We estimate the total size of the subsidy, based on Equation (5) from the previous section, with several simplifications. First, we only focus on the first term, the fees paid on the implicit government account.\footnote{The second term of Equation (5) (the “match”) exists regardless of fees if and only if \( \tau_R \neq \tau_L \), and therefore an estimate should only include the difference between the cost of the match with fees and the cost of the match without fees. In practice, this difference is small. Moreover, in principle with Roth accounts the government could simply provide an explicit match. In this case, the revenue impact of the match would be the same as under Traditional, eliminating this term from the difference.} Second, we focus on the annual flow of revenue being lost to fees rather than the present value of all foregone revenue, because we want to measure the subsidy and its impact in a way that can be compared to the current government budget. Thus, a simple estimate of the annual subsidy can be calculated as:

\[
\text{Annual subsidy} = S \cdot \tau_R \cdot f \cdot (1 - \tau_C).
\]  

(6)

We next calibrate Equation (6), first for the U.S. and then for the six additional countries with the largest Traditional retirement assets.

4.1 Tax-deferred retirement assets in the U.S. \((S)\)

Table 6 summarizes the composition of tax-advantaged retirement assets in the U.S. Total retirement assets amount to $26.6 trillion. We estimate \(S\) as the total amount of tax-deferred assets in IRAs and DC plans. To be conservative, we exclude DB plans ($7.5 trillion), which are also tax-deferred and to which a similar argument applies. We also exclude annuities ($2.1 trillion) because their special tax treatment entails only a small amount of tax deferral. The remainder of retirement assets ($17.0 trillion) includes two main components: employer-sponsored defined contribution retirement accounts such as 401(k) and 403(b) plans (DC plans), and individual retirement accounts (IRAs), with roughly $8 and $9 trillion of assets respectively. A large majority of IRA assets were initially accumulated in DC plans: according to the ICI Retirement Market Statistics (Tables 9 and 31...
<table>
<thead>
<tr>
<th>$ billion</th>
<th>Total</th>
<th>Roth (TEE)</th>
<th>Traditional (EET)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total retirement assets</td>
<td>26,623</td>
<td>1,053</td>
<td>23,454</td>
</tr>
<tr>
<td>Individual retirement arrangements (IRAs)</td>
<td>9,263</td>
<td>834</td>
<td>8,429</td>
</tr>
<tr>
<td>Defined contribution (DC) plans</td>
<td>7,780</td>
<td>219</td>
<td>7,561</td>
</tr>
<tr>
<td>401(k) and 403(b)</td>
<td>6,338</td>
<td>199</td>
<td>6,139</td>
</tr>
<tr>
<td>Other private-sector DC</td>
<td>540</td>
<td>0</td>
<td>540</td>
</tr>
<tr>
<td>Thrift Savings Plan (TSP)</td>
<td>580</td>
<td>9</td>
<td>571</td>
</tr>
<tr>
<td>457</td>
<td>323</td>
<td>10</td>
<td>313</td>
</tr>
<tr>
<td>Annuities</td>
<td>2,116</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Defined benefit (DB) plans</td>
<td>7,464</td>
<td>0</td>
<td>7,464</td>
</tr>
</tbody>
</table>

Table 6: U.S. Retirement assets by type of account. Source: ICI Retirement Market Statistics 2018q2 (totals and Roth IRAs) and own estimates (Roth DC plans). 401(k), 403(b), and TSP are standard DC plans sponsored respectively by private-sector employers, non-profit employers, and the federal government. Other private-sector DC plans include Keogh, profit-sharing, thrift-savings, stock bonus, and money purchase plans. 457 plans are tax-advantaged deferred compensation arrangements available for certain employers in the United States. DB Plans exclude $1,680 billion of U.S. government employee DB plans which are required by law to be invested in U.S. government obligations.
10), 90% of IRA inflows from 1995 to 2015 were rollovers, primarily from DC plans. From these assets we further remove $0.6 trillion of assets in the federal government’s Thrift Savings Plan (TSP), whose fees are negligible, and $1.0 trillion of Roth assets. This results in an estimated amount $S = $15.4 trillion.

4.2 Tax rates ($\tau_R$ and $\tau_C$)

Because of progressive taxation, the effective rate effective tax rate on retirement payouts, $\tau_R$, is challenging to estimate - it is neither the marginal tax rate on retirement income, nor the average, because individuals may have retirement income from sources other than tax-deferred accounts. Rather, the appropriate definition of $\tau_R$ is the present value of future taxes that will be paid on all Traditional balances accumulated as of today ($S$), minus the present value of future taxes that would be paid if Traditional distributions were not taxable, as a fraction of the present value of all taxable distributions.

We estimate $\tau_R$ using data on retirement wealth reported in the 2013 Survey of Consumer Finances (SCF) and the U.S. marginal tax schedule in place at the time of the survey, assuming they remain constant moving forward from 2013. Our resulting estimate of the effective tax rate on retirement wealth is 25.8%. This estimate could be overstated because it is obtained assuming individuals do not adjust the timing of their withdrawals to take advantage of fluctuating marginal tax rates, but it could also be understated because it does not take into account the taxable part of Social Security benefits and any labor income. As a check, we also reverse-engineer present-value

---

32 The ICI Retirement Market Statistics report that $834 billion of IRA assets is in Roth IRAs. For DC plans, we roughly estimate total Roth assets to be $219 billion based on information on Roth adoption rates in T. Rowe Price’s 2017 Reference Point and Vanguard’s How America Saves 2017 reports, together with ICI information on contribution flows, and the fact that Roth options were introduced in 2001 for 401(k) and 403(b) plans. We are not aware of data on Roth options for other private-sector DC plans.

33 For each SCF observation in which the head of household is aged 65 to 74 we calculate $W^T$, taxable wealth, and $W^{Trad}$, tax-deferred retirement wealth. We assume baseline retirement taxable income to be equal to taxable wealth times a constant nominal rate of return $r = 3\%$ ($Y_{Baseline} = W^T r$). In addition to this baseline income, we assume that the individual uses $W^{Trad}$ to withdraw an equal nominal amount each year for 20 years (computed as a fixed annuity stream using the same rate $r = 3\%$. This implies that $Y_{Combined} = W^T r + W^{Trad} / \left( \{1 - (1 + r)^{-T}\}/r \right)$. We then calculate total dollar tax on the baseline income ($T_{Baseline}$) and on the combined total income ($T_{Combined}$), and calculate $\tau_R$ as $\frac{\sum_i(T_{Combined,i} - T_{Baseline,i})w_i}{\sum_i(Y_{Combined,i} - Y_{Baseline,i})w_i}$, where $w_i$ are the SCF sampling weights for household $i$.
tax expenditure estimates published by the federal government (Office of Management and Budget, 2014) or its staff (Lurie and Ramnath, 2011) and obtain a range for the effective marginal tax rate (based on different studies) of 20%–30%. As a conservative estimate we use 20%, the lowest of all our estimates. Note that this estimate also excludes any state income taxes on Traditional retirement payouts. The size of the implicit government account is therefore $S \cdot \tau_R = 3$ trillion.

For $\tau_C$, the corporate tax rate, we simply use the top statutory corporate tax rate of 21%.

### 4.3 Calibration

Using the calibrated inputs in Eq. (6), we obtain our estimate:

\[
\text{Annual subsidy} = S \cdot \tau_R \cdot f \cdot (1-\tau_C) = \\
= $15.4 \times 20\% \times 0.80\% \times (1-21\%) = $19.5 \text{ billion.}
\]

This amount is equal to 4.4% of the size of the total interest expense ($363.4 billion) or 1.6% of the federal budget deficit ($984.4 billion), and is about the size of the 2018 budget of NASA ($20.7 billion).\(^{34}\)

As discussed above in Section 4.1, our estimate of assets under management excludes $7.5 trillion of tax-deferred assets in state and local government and corporate defined-benefit pension plans. Although these assets do not belong to any individual in particular, they are subject to the exact same tax deferral benefit: the contribution is made with pretax money, and benefits are taxed only when paid out. Therefore, defined-benefit plan assets can also be decomposed into an employees’ account and a government account earmarked to pay future taxes. While defined-benefit plans are likely to be more efficient investment vehicles than defined contribution plans or IRAs, they still incur positive investment fees. Accounting for the government’s implicit share of DB assets would increase our estimate of the government’s account by another $7.5 \times 20\% = $1.5$ trillion. Assuming lower fees for DB plans (47 bps instead of 80 bps),\(^{35}\) the estimated subsidy rises

---


\(^{35}\)This rough estimate is implied by Table II of Dyck and Pomorski (2011). We define total plan fees as “Overall
Table 7: Estimated subsidy to the asset management industry in seven countries with the largest Traditional retirement assets. Fees are the asset-weighted average of money market, equity and fixed-income mutual fund fees based on overall (not retirement-only) asset allocation in that country. For each country, $\tau_R$ (the tax rate on retirement income, and therefore the fraction of Traditional assets that implicitly belong to the government) is calculated as the average tax rate faced by a person earning the average retirement income with no other income. $\tau_C$, the corporate tax rate, is simply the top statutory tax rate. Sources: see text.

<table>
<thead>
<tr>
<th>Country</th>
<th>Data Year</th>
<th>Retirement Assets $b</th>
<th>% Deferred</th>
<th>Govt. Acct. Size $b</th>
<th>Fees $b</th>
<th>$\tau_R$</th>
<th>$\tau_C$</th>
<th>Subsidy $b</th>
<th>% GDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td>2018</td>
<td>16,464</td>
<td>94%</td>
<td>20%</td>
<td>3,084</td>
<td>0.80%</td>
<td>21%</td>
<td>19.5</td>
<td>0.10%</td>
</tr>
<tr>
<td>Canada</td>
<td>2015</td>
<td>1,003</td>
<td>86%</td>
<td>15%</td>
<td>129</td>
<td>2.06%</td>
<td>15%</td>
<td>2.3</td>
<td>0.15%</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>2015</td>
<td>950</td>
<td>86%</td>
<td>20%</td>
<td>41</td>
<td>1.45%</td>
<td>20%</td>
<td>0.7</td>
<td>0.02%</td>
</tr>
<tr>
<td>Netherlands</td>
<td>2015</td>
<td>108</td>
<td>100%</td>
<td>39%</td>
<td>41</td>
<td>1.41%</td>
<td>25%</td>
<td>0.4</td>
<td>0.06%</td>
</tr>
<tr>
<td>Switzerland</td>
<td>2015</td>
<td>945</td>
<td>100%</td>
<td>40%</td>
<td>38</td>
<td>1.29%</td>
<td>18%</td>
<td>0.4</td>
<td>0.06%</td>
</tr>
<tr>
<td>Australia</td>
<td>2015</td>
<td>1,797</td>
<td>55%</td>
<td>3.4%</td>
<td>34</td>
<td>1.10%</td>
<td>30%</td>
<td>0.3</td>
<td>0.02%</td>
</tr>
<tr>
<td>Japan</td>
<td>2015</td>
<td>112</td>
<td>100%</td>
<td>2.6%</td>
<td>3</td>
<td>1.47%</td>
<td>30%</td>
<td>0.0</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

In Table 7, we carry out the same back-of-the envelope calculation for the seven countries with the largest dollar amounts of tax-deferred assets (the rows are ordered by the size in dollars of the implicit government account). We use data on all existing types of tax-advantaged retirement plans and their tax treatment from the Organization for Economic Cooperation and Development (OECD, 2015a,b), estimates of their magnitudes from various sources (2015-2018), average retirement income from each country’s statistical office, information on basic deductions, personal tax brackets, and the corporate tax rate from each country’s tax authority, and fee estimates from Morningstar (Alpert et al., 2013) and other sources.\textsuperscript{36} For consistency with our U.S. estimates, we exclude defined benefit (DB) pension plans from the calculation. With or without DB plans, the asset-class-level costs” plus “Plan-level administrative costs” and calculate an approximate asset-weighted average of 31 bps based on the average plan size in each of the size quintiles. To this figure we add our estimate of trading costs of 16 bps.\textsuperscript{36}Our estimate of $\tau_R$ is a rough lower bound, equal to the average tax rate faced by a person earning the average retirement income with no other income.

---

\textsuperscript{36}
U.S. has the world’s largest retirement assets, and therefore leads the list. However, other countries have substantial amounts of DB retirement assets (United Kingdom, Netherlands and Japan), and omitting DB leads to an important underestimate of the size of the implicit government account. In the case of United Kingdom and Netherlands, this underestimate meaningfully affects the estimated subsidy.

Each of the components of the subsidy has substantial variation across countries. For instance, although Switzerland, Australia and Japan have significant tax-deferred assets, the estimated subsidy is small simply because under current tax law retirement payouts are lightly taxed. Canada has the second-largest subsidy in dollar terms ($2.3 billion) and the largest as a fraction of GDP (0.15%), driven by the surprisingly large fees charged by Canadian funds (2.06%).

4.4 Does the government receive value for the fees it pays?

In the calibration above, we have ignored any potential benefits the government obtains in exchange for the fees paid on its virtual account, including direct portfolio services, alpha, and economy-wide benefits. We argue that the first two are likely to be small and leave a discussion of the third primarily to our general equilibrium model in 5.

4.4.1 Portfolio services

One important service that individual investors receive is basic portfolio management, i.e., asset managers build and maintain large portfolios of securities. However, basic portfolio management services are an inexpensive commodity. The U.S. federal government’s Thrift Savings Plan (TSP) pays no explicit asset management fees for its stock funds, all of which are index funds, instead only sharing the revenue from securities lending.37 Many major private-sector mutual funds families offer passive index funds charging expense ratios of 4 bps or less.

---

37 According to plan documents (Thrift Savings Plan, 2018), all TSP funds had administrative expenses of 4.6 bps in 2017, but these expenses were incurred for recordkeeping and similar services at the plan level, and not for portfolio management at the fund level. The portfolio manager for all TSP funds received as only compensation a fraction of securities lending revenues amounting to roughly 1.0 bps of AUM for those funds that actually hold securities.
Second, a substantial fraction of the cost of investing in mutual funds and other structured products consists of distribution fees. The government, however, does not benefit from distribution services.

Third, individuals may pay fees to advisers who help them identify a suitable asset allocation and to asset managers who create products that help implement such an allocation. Examples of such products include funds focusing on style (conservative/aggressive, depending on the investor’s risk tolerance), or personal situations (target-date funds for individuals retiring in 2030). Although individual investors may experience real benefits from a more tailored asset allocation, these benefits largely cancel out in aggregate for the government, which holds a fraction of all individual accounts.

Finally, part of asset-level distribution fees and individual-level advisory fees may cover financial advice that helps individuals find lower-fee funds. This advice does create value for the government which incurs lower fees on its implicit account. However, the value of this advice is already reflected in the observed level of fees.

In sum, if the government wanted to obtain exposure to stocks or any other securities, it could invest directly in a passive vehicle at very low cost. Such a vehicle would not provide those services that the government does not need, but is currently paying for, in Traditional accounts.

4.4.2 Alpha

Actively managed funds have significantly higher fees and trading costs than passive index funds. However, it is possible that actively managed funds also earn higher expected returns. For instance, in Berk and Green’s (2004) model, active money management pays for itself and investors are

---

38For instance, according to its 2017 annual report, Invesco (one of the largest listed U.S. investment management companies) had AUM of $875 billion with operating revenues of $5 billion (57 bps). This amount, which does not include implicit trading costs, is not far from our explicit fee estimate. Invesco’s major operating costs were personnel ($1.5 billion or 18 bps), third-party distribution, service and advisory expenses, including passing through of 12b-1 and similar fees ($1.5 billion or 17 bps), and marketing ($123 million or 1.4 bps). Marketing and distribution costs are close to one-third of revenue, similar to the estimate by Roussanov et al. (2018).

39To the extent that individuals’ allocations do not cancel out perfectly, the average asset allocation in tax-deferred accounts may differ slightly from the market portfolio. For instance, Boguth and Simutin (2018) show that the average active fund has a beta of is 1.08, i.e., a slight tilt towards riskier stocks. If this departure from the market portfolio is optimal for the government, even this modest benefit could be replicated at a much lower expense.

32
indifferent between active and passive funds in equilibrium. On its face, this would imply that the overall net of alpha cost of active management should not be measured but rather assumed to be equal to the cost of passive management.\footnote{In a recent working paper, Roussanov et al. (2018) show that adding marketing and search costs to a model like Berk and Green (2004) has the potential to explain the distribution of fund size. In their model, active management does not pay for itself.}

For our purposes, however, alpha must represent real value that funds in aggregate create and capture by trading. Active funds held in tax-deferred accounts must either be systematically winning a zero-sum game against other market participants, or they must be making prices more efficient \textit{and} capturing the resulting value.

In practice, the existing empirical evidence suggests that, in aggregate, mutual funds lose the zero sum game against other market participants (Shive and Yun, 2013; Ince et al., 2018, etc.). For active funds in particular, all published studies we are aware of point to a negative aggregate alpha. In the words of Fama and French (2010), “The aggregate portfolio of actively managed U.S. equity mutual funds is close to the market portfolio, but the high costs of active management show up intact as lower returns to investors.” However, Fama and French and many other studies are based on fund-level factor regressions—typically because their goal is to capture manager skill—and thus essentially equal-weighted. Figures comparable to our asset-weighted estimate are provided by two studies. French (2008) constructs a measure of the aggregate cost of equity investing and concludes that passive investing has a cost of roughly 0.10% and active investing an additional 0.67% of the total capitalization of the U.S. stock market. Berk and van Binsbergen (2015) compare active funds against “investable” benchmarks (i.e., retail shares of Vanguard funds), estimating a value-weighted net alpha of -12 bps. This cost of active management is in addition to the cost of investing in the specific passive benchmark (18 bps plus any implicit trading costs).\footnote{According to Petajisto (2009), the S&P 500 index itself has implicit trading costs of 20-28 bps.} While the two studies disagree in their conclusions, neither study implies that individuals should be indifferent between active and passive funds. Moreover, neither study includes account- or plan-level fees and individual-level fees. Adding these fees to either estimate would imply total investment fees of the same order of magnitude as ours, though somewhat lower.
Finally, if aggregate alpha is positive, the value created by the additional funds in the government’s implicit account is likely to be small. Recent theoretical and empirical work on industry-level diminishing returns to scale in asset management (Pástor and Stambaugh, 2012; Pástor et al., 2015) suggests that every additional dollar of assets in the industry (and every additional fund that becomes viable because of it) creates less and less value. Upon a switch from Roth to Traditional, the additional government assets would create little additional value.

### 4.4.3 Economy-wide benefits

The government’s implicit ownership of assets as a result of Traditional accounts entails a few possible economy-wide benefits. First, in the presence of otherwise binding constraints, a potential benefit arises from the increased government ownership of stocks. We have discussed the value of government ownership of stock in Section 2.4, and argued that it could be either positive or negative, depending on the nature of the constraints.

Second, a potential positive externality arises from the added demand for the services of active managers, who could contribute to price discovery and cannot appropriate the entire value from making capital market prices more efficient. It remains an open question, however, whether the proportional scaling up in demand for active management from a hypothetical switch from Roth to Traditional would actually aid price discovery and improve the efficiency of capital markets, as this may depend on whether firms simply scale up each trade or put more resources into actively choosing trades.

Finally, consumers may receive benefits from more tailored asset services that could arise as a result of the expansion of the number of asset management firms. We address this possibility in the general equilibrium model in the next section, where we explicitly incorporate the benefits of additional asset management firms.
5 A general equilibrium model of retirement savings and asset management

In this section, we build and solve a simple, two-sector general equilibrium model to address four questions. First, are the aggregate resources devoted to asset management higher under Traditional than under Roth? Second, who bears the incidence of these higher costs? Third, how do fees charged vary across systems? And fourth, which system (Roth or Traditional) results in higher social welfare? Our model includes an asset management industry with fixed and variable costs, allows for firm entry and exit, and is compatible with the basic empirical evidence about cost structure, market structure, and competition in the asset management industry. If all asset management costs were variable and proportional to assets under management, our partial equilibrium resultson the extra aggregate resources devoted to asset management under Traditional would carry through to general equilibrium. To stress-test this result, and to be consistent with evidence suggesting economies to scale as firms expand AUM, we allow for the possibilities that some or all asset management costs at the firm level are fixed and that new asset management firms generate some value to consumers. We find that Traditional nevertheless involves higher resources devoted to asset management and lower social welfare than Roth accounts.

5.1 Modeling choices

We begin by describing the nature of the choices we make in constructing the general equilibrium model, including cost structure, market structure and competition, and firm entry.

5.1.1 Cost structure

A key modeling choice relates to the existence and magnitude of economies of scale in the asset management industry. The academic literature has considered two related but distinct sources of economies of scale in asset management: costs and performance.

On the cost side, the existence of economies of scale with assets under management seems
uncontroversial: most empirical studies agree that larger funds and larger plans have lower administrative costs per dollar of AUM (Baumol et al., 1990; Latzko, 1999; Coates and Hubbard, 2007; Gao and Livingston, 2008; Hubbard et al., 2010; Dyck and Pomorski, 2011). Statements by industry insiders also support this idea: Kahn (2002) quotes the director for portfolio review at a major fund family as saying that the “marginal cost of managing increasing dollars is minimal.”

On the performance side, the picture is more nuanced. On one hand, Dyck and Pomorski (2011) show that larger DB plans outperform smaller ones at least in part because of access to a better investment opportunity set. On the other hand, for active mutual funds, the existence of diseconomies of scale has been theorized both at the fund level (Perold and Salomon, 1991; Berk and Green, 2004; Gabaix et al., 2006) and the industry level (Pástor and Stambaugh, 2012). In practice, the available evidence shows that size is negatively associated with performance both at the fund level (Chen et al., 2004; Yan, 2008; Berk and van Binsbergen, 2015) and at the industry level (Pástor et al., 2015). Since we do not directly model investment performance, we must represent any economies of scale as occurring on the cost side. We do so by using a simple cost function equal to the sum of a fixed component and a variable component that is proportional to assets under management.

5.1.2 Market structure and competition

There is abundant evidence that investors are not very sensitive to the price of asset management services. Hortacsu and Syverson (2004) note that in 2000 there existed 82 S&P 500 index fund share classes (50 distinct funds) with large dispersion in fees (an interquartile range of 98 bps). We update their analysis and find that price dispersion persists: in 2016 there existed 79 share classes (46 distinct funds) with a fee interquartile range of 102 bps.\footnote{Christoffersen and Musto (2002) show evidence of fee dispersion for money market mutual funds. Cooper et al. (2018) show that fee dispersion has increased over the last 20 years in a comprehensive study of equity mutual funds with homogeneous holdings.} Although funds with lower fees tend to have higher market shares (Hubbard et al., 2010), multiple studies point to the continued existence of dominated funds, such as funds that are more costly and underperform (Bergstresser et al., 2009; Gil-Bazo and Ruiz-Verdú, 2009), structured equity products with negative expected returns (Henderson and Pearson, 2011), and “closet indexers” that charge fees as if they are active
All proposed explanations for price insensitivity point to information frictions or outright inertia: marketing (Roussanov et al., 2018), search costs (Hortacsu and Syverson, 2004), captive DC plan participants (Ayres and Curtis, 2015; Pool et al., 2016), shrouded prices of complex financial products (Gabaix and Laibson, 2006; Carlin, 2009), investors’ inability to precisely observe the quality of fund management (Gil-Bazo and Ruiz-Verdú, 2008; Gärleanu and Pedersen, 2018), investors’ unwillingness to sever relationships with brokers or trusted advisors (Bergstresser et al., 2009; Hubbard et al., 2010; Gennaioli et al., 2015), or even irrationality (Elton et al., 2004).

To capture limited price sensitivity, we build a model of monopolistic competition with differentiated products based on Salop (1979). In this model, funds face a downward-sloping demand function, i.e. if they raise their fees, their demand falls, but it does not fall to zero. Our approach is conservative in that every additional fund entering the market creates value for some investors, i.e. there are no dominated funds. Alternative choices such as informational or behavioral frictions would have likely generated stronger welfare losses under Traditional relative to Roth.\(^{43}\)

Another important feature of asset management markets is the existence of tiered pricing (decreasing marginal fees) based on account size, which results in decreasing average fees as a function of account size. To approximately capture this feature, we assume funds set a two-part fee structure with a fixed component and a variable component proportional to assets.

5.1.3 Entry

Empirical evidence and casual observation suggest low barriers to entry in the mutual fund industry (Hubbard et al., 2010; Baumol et al., 1990). In 2018 alone, 345 new mutual funds, 237 new exchange-traded funds, and 40 new fund sponsors entered the industry (ICI, 2019).\(^{44}\) To reflect

\(^{43}\)For instance, we could have based our analysis on a search-based model like Anderson and Renault (1999) or Roussanov et al. (2018). These models, like almost every model of firm competition based on search frictions, results in high prices and either excessive entry or excessive expenditure on marketing. Other alternatives include a model with captive demand and shrouded fees or one based on trust such as Gennaioli et al. (2015).

\(^{44}\)A similar situation is reflected in the non-mutual fund segments of the asset management industry. For instance, the majority of the leading third-party retirement plan administrators were established in the past 25 years (see Plansponsor’s 2016 TPA survey of 1,070 administrators available at http://www.plansponsor.com/2016-Third-Party-Administrator-Survey/).
Figure 3: **Spatial competition.** $N$ funds are uniformly spaced along a circle of circumference 1, as in Salop (1979). Individuals are distributed uniformly along the circle, and prefer funds located at a closer distance.

This evidence, we assume that there is no entry cost (other than the fixed operating cost), and funds enter the market until profits are zero. Because our model allows for product differentiation, the assumption of a dominant fixed cost component does not result in a monopoly.

### 5.2 Demand: individuals

Our model is based on Salop’s (1979) circular city, which we embed in a two-period economy. There is a unit continuum of individuals $i \in [0, 1)$ uniformly distributed over a circle of circumference 1 (Figure 3). When individuals are young ($t = 0$), they work, consume, and save for retirement. When old ($t = 1$), individuals retire and consume all of their accumulated savings (net of any taxes), leaving no bequest.

Young individuals are endowed with one unit of labor, which they supply inelastically. Individuals can work either in the consumption goods sector or in the asset management sector, and are
indifferent (at the same wage) between working in the two sectors. The production technology is linear, i.e., the marginal (and average) product of labor is $\omega$. Since labor markets are frictionless, $\omega$ also equals the wage of each worker.

Individual savings, $S_i$, must be allocated to one of $N$ firms producing asset management services (“funds”), indexed by $j \in \{1, 2, \ldots, N\}$. Funds are situated on the same circle at locations $t_1$, $t_2$, $\ldots$, $t_N$. Each individual’s utility depends negatively on the “distance” from their chosen fund. A low distance can be thought of as literally low physical distance from the nearest branch, but also ease of finding, availability (e.g., the fund is part of a small menu of preselected funds, as is the case for many retirement plans), trust, a preference for portfolio characteristics of funds, or even non-portfolio characteristics such as the level of customer service.

We assume that the utility of individual $i$ is

$$u_i (C_{0,i}, C_{1,i}, d_{i,j}) = \ln C_{0,i} + \delta \ln C_{1,i} - \gamma \ln d_{i,j}. \tag{8}$$

Individuals derive utility from current consumption ($C_{0,i}$) and future consumption ($C_{1,i}$) discounted by a factor $\delta$, and derive disutility from the distance between their own location $i$ and the location $t_j$ of their chosen fund $j$, $d_{i,j} \equiv |i - t_j|$. On the circle, the distance between two points is defined as the shortest possible distance. We choose this logarithmic utility specification because it is both economically sensible and tractable, yielding easy-to-interpret expressions for the quantities of interest.$^{45}$

$^{45}$We are not aware of any works using the circular city model that feature either an intertemporal choice problem or logarithmic distance disutility. Salop (1979) and many subsequent works feature linear consumption utility and linear or quadratic distance disutility (see Gong et al., 2016 for a review). We choose logarithmic consumption utility because linear consumption utility does not permit us to study the savings decision. We choose logarithmic distance disutility for tractability, and because it allows for an interpretation of “distance” as something whose importance increases with the individual’s wealth.

One potential problem with logarithmic distance disutility is that individuals living at exactly zero distance from their fund have infinite utility. This is not technically problematic. The planner’s problem is well-defined because the integral of $u (\cdot)$ over the $[0, 1)$ circle is finite, and the market equilibrium is unaffected by these individuals because their utility can still be maximized with respect to $C_0$ and $C_1$, and firms do not find it profitable to charge infinite (or 100%) fees. However, if one were to find this setup philosophically problematic, the same results could be obtained by assuming that no individuals live in an $\varepsilon$-neighborhood centered around each fund and examining the limit as $\varepsilon \to 0$. 

39
5.3 Supply: Funds

The unit of production of asset management services is the “fund”, a generalist firm that produces asset management services. We ignore the existence of multiple layers of financial intermediation (e.g., recordkeepers, asset managers, fund families, subadvisors, securities brokers, etc.). We assume a linear storage technology with exogenous non-stochastic return \( r \), and allow storage to be positive or negative. Each fund invests its assets in this storage technology and/or in riskless government bonds (which by arbitrage pay return \( r \) as well) and incurs costs described below. Following Salop (1979), we assume the \( N \) funds are evenly distributed around the circle, as shown in Figure 3, and that \( N \) can be noninteger.\(^{46}\)

We assume that all costs are incurred at time 0, and equal

\[
Costs_j = \phi + c \cdot AUM_j,
\]

where \( Costs_j \) equals total costs (measured in goods) incurred by firm \( j \), \( \phi \) is the fixed component of costs, and \( c \cdot AUM_j \) is the variable component of costs, equal to a fraction \( c \) of assets under management of firm \( j \). Since the marginal product of labor is \( \omega \), the amount of labor going to the asset management sector is

\[
L \equiv \frac{N \cdot \phi + c \cdot AUM}{\omega},
\]

where \( AUM \equiv \sum_{j=0}^{N} AUM_j \), and the amount of labor going to the consumption goods sector is \( 1 - L \).

\(^{46}\)These assumptions lighten the exposition considerably and are consistent with the existing literature. It is easy to show that equidistant funds maximize social welfare and a planner would choose this location pattern. Economides (1989) derives location endogenously in a three-stage game, albeit at the cost of restrictions on the utility function and greater complexity. Requiring \( N \) to be integer can lead to a situation in which \( N \) funds make positive profits, and \( N + 1 \) funds make negative profits. This possibility, which could have important consequences if \( N \) were small, is analyzed by Mankiw and Whinston (1986).
5.4 The planner’s problem

We characterize the social optimum in the model by assuming the planner directly chooses individual consumption \( \{C_{0,i}\} \) and \( \{C_{1,i}\} \) and the number of funds \( N \) to maximize the sum of individual utilities, so that

\[
U = \max_{\{C_{0,i}\},\{C_{1,i}\},N} \int_0^1 (\ln C_{0,i} + \delta \ln C_{1,i} - \gamma \ln d_{i,j}) \, di,
\]

subject to the resource constraint

\[
C_1 = [(\omega - C_0 - G) (1 - c) - \phi N] (1 + r),
\]

where \( C_t \equiv \int_0^1 C_{t,i} \, di \) is aggregate consumption at time \( t \in \{0, 1\} \), \( G \) is an exogenously given amount of public expenditure, and \( r \) is the exogenously given return on the storage technology. Since we assume that no production occurs in period 1, the only way the planner can allocate consumption in period 1 is via the asset management industry. The planner incurs variable asset management costs at rate \( c \) on \( \omega - C_0 - G \).\(^{47}\)

It is optimal for the planner to give equal consumption to all individuals (\( C_{0,i} = C_0 \) and \( C_{1,i} = C_1 \)) because consumers have identical preferences and productivity, and because the planner’s objective function is concave, separable in its arguments (consumption and distance), and unaffected by inequality in utility. Since all funds are identical except for location, it is also optimal for the planner to allocate each consumer to the nearest fund. Then, using the assumption that the \( N \) funds are located equidistantly along the circle, the planner’s objective function simplifies (see the Internet Appendix for a derivation) to

\[
U = \max_{C_0,N} \ln C_0 + \delta \ln C_1 + \gamma \ln N - \gamma (1 + \ln 2).
\]

\(^{47}\)For simplicity, we assume that the planner incurs variable costs on all of \( \omega - C_0 - G \), including the resources devoted to fixed costs of asset management.
There are two first-order conditions. The first is

\[ u'(C_0^*) = \delta (1 + r) (1 - c) u'(C_1^*), \tag{14} \]

where \( u'(C_t^*) \equiv 1/C_t^* \). This equation says that the marginal cost in terms of lost utility of giving the consumer one less dollar of period-0 consumption must equal the marginal benefit of allocating that dollar plus interest and net of the marginal asset management costs to period 1 consumption.

The second first-order condition, related to the choice of \( N^* \), is

\[ \phi \cdot \delta (1 + r) u'(C_1^*) = \frac{\gamma}{N^*}, \tag{15} \]

and says that at the optimum the cost of adding another fund, \( \phi \), expressed in units of utility, must equal \( \gamma/N \), the utility benefit from adding that fund that arises due to the drop in the average distance.

Combining and simplifying the first-order conditions yields the following optimal quantities:

\[ C_0^* = \frac{1}{1 + \delta + \gamma} (\omega - G), \tag{16} \]

\[ C_1^* = \frac{\delta}{1 + \delta + \gamma} (\omega - G) (1 - c) (1 + r), \tag{17} \]

\[ N^* = \frac{\gamma}{1 + \delta + \gamma} (\omega - G) (1 - c) \frac{1}{\phi}. \tag{18} \]

Note that the distaste for distance enters into the formulae for optimal consumption in each period, because it affects the optimal number of firms, and thus the aggregate resources available for consumption. Note also that \( N^* \phi \), the total allocation to the fixed costs of asset management, does not depend on the size of the fixed cost \( \phi \).
5.5 The market economy

In the market equilibrium, individuals pay taxes, choose consumption \((C_{0,i} \text{ and } C_{1,i})\) and private saving \((S_i)\), and allocate their savings to a fund \((j)\) to maximize utility. Note that since this is a two-period non-overlapping generations model, the flow of private saving equals the stock of accumulated assets and we use the same notation for each. The \(N\) funds set their fees to maximize profits taking competitors’ choices as given. We assume that funds cannot engage in price differentiation based on distance, and that each fund \(j\) charges each of its customers a two-part fee structure: a fixed component \(F_j\) and a variable component \(f^v_j\) per dollar of saving.\(^{48}\)

5.5.1 Government spending and taxes

At time 0, the government spends an exogenously given amount \(G\), taxes individuals’ labor income at a rate \(\tau_L\), and grants individual \(i\) an income tax deduction of \(\tau_S\) for each dollar contributed to retirement saving \((S_i)\). At time 1, the government taxes retirement income at a rate \(\tau_R\). The values of \(\tau_S\) and \(\tau_R\) depend on the specific incentive system (Roth or Traditional). Under Roth, the government grants no deduction \((\tau_{S}^{\text{Roth}} = 0)\) and does not tax retirement income \((\tau_{R}^{\text{Roth}} = 0)\). Under Traditional, the government grants a deduction for savings at the same rate as labor income \((\tau_{S}^{\text{Trad}} = \tau_L)\).\(^{49}\) For simplicity, we assume that aggregate firm profits \(\Pi = \sum_{j=1}^{N} \pi_j\) are rebated at time 0 to consumers (an equal amount per consumer) and are taxed at the personal level only, at the same rate as labor income.

5.5.2 The individual’s problem

At time 0, individuals earn labor income \(\omega\) and receive profits \(\Pi\). Each individual pays taxes \(T_{0,i}\), consumes \(C_{0,i}\), and saves and invests the remainder \(S_i = \omega + \Pi - T_{0,i} - C_{0,i}\) to finance retirement.

\(^{48}\) Within time 0, a number of events happen. Individuals receive their income (wage \(\omega\) and fund profits \(\Pi\)) on which they pay taxes and choose their consumption \(C_0\) and saving \(S\). Individuals hand over their savings to funds, which incur fixed costs \(\phi\) and variable costs \(cS\), and charge fees \(F + f^vS\).

\(^{49}\) Note that in our setup with fixed labor supply labor income taxes induce no distortions, whereas capital income taxes introduce a wedge in the Euler equation and discourage saving. A Roth retirement account system eliminates the wedge and achieves the same Euler equation as in the planner’s problem.
consumption \( C_{1,i} \).\( T_{0,i} \) can be written as \( \tau_L (\omega + \Pi) - \tau_S S_i \), which implies the following budget constraints

\[
C_{0,i} + S_i = (\omega + \Pi) (1 - \tau_L) + \tau_S S_i,
\]

\[
C_{1,i} = [S_i (1 - f_j^u) - F_j] (1 + r) (1 - \tau_R).
\]

Note that (19) can be rewritten as\(^{50}\)

\[
S_i = \frac{(\omega + \Pi) (1 - \tau_L) - C_{0,i}}{1 - \tau_S}.
\]

We assume for simplicity that any legal limits on the amount of saving that individuals can put into either type of retirement account, if present, are large enough that they are never binding. We also assume that the individual must allocate all savings to one fund, \( j \). Moreover, to rule out what Salop (1979) calls a “supercompetitive” equilibrium (the situation in which a fund charges sufficiently low fees to be competitive with funds other than its immediate neighbors), we assume that the individual only considers the two nearest funds. Thus, the individual chooses \( C_{0,i} \) and \( j \) to maximize utility:

\[
U_i = \max_{C_{0,i,j}} \ln C_{0,i} + \delta \ln C_{1,i} - \gamma \ln d_{i,j},
\]

subject to the budget constraints (19) and (20). The individual’s first-order condition with respect to consumption is

\[
u'_i(C_{0,i}) = \delta (1 + r) \left(1 - f_j^u\right) \left(1 + \tau_M\right) u'_i(C_{1,i}),
\]

where \( u'_i(C_{t,i}) \equiv 1/C_{t,i} \), and \( \tau_M \) is the government “match” rate, equal to \((\tau_L - \tau_R)/(1 - \tau_L)\) under Traditional and 0 under Roth, that arises due to the difference in working and retirement tax rates.\(^{51}\) This first-order condition differs from that of the planner in two ways. First, a non-zero match changes the tilt in the consumption path \((C_{1,i}/C_{0,i})\), with a positive match increasing and a

\(^{50}\)Equation (21) implies \( S_i^{Roth} = (\omega + \Pi) (1 - \tau_L^{Roth}) - C_{0,i} \), and \( S_i^{Trad} = \omega + \Pi - C_{0,i}/(1 - \tau_L^{Trad}) \).

\(^{51}\)Under Traditional, each dollar of initial-period consumption that is given up generates an extra \( 1/(1 - \tau_L) \) dollars of saving. As described in Section 2.2 above, each dollar of saving is effectively matched by a government contribution equal to \( \tau_L - \tau_R \).

44
negative match decreasing the tilt. Second, for the planner the cost of future consumption relative to today’s consumption depends only on the return on assets net of variable costs \([(1 + r)(1 – c)]\), whereas for the individual it depends on the return on assets net of variable fees \([(1 + r)(1 – f^v)]\). If variable fees include a markup over variable costs \((f^v > c)\), this reduces the tilt in the optimal consumption path of the market economy (relative to the planner), counteracting the effect of the saving subsidy itself.

For the choice of a fund there is no first-order condition. Individuals simply pick the one out of \(N\) funds that gives them the highest utility:

\[
j^* = \arg \max_{j \in \{1, 2, \ldots, N\}} \ln C_{0,i} + \delta \ln C_{1,i} - \gamma \ln d_{i,j},
\]

where both \(C_{0,i}\) and \(C_{1,i}\) are functions of \(j\). In choosing a fund, individuals are willing to pay higher fees (and thus through the budget constraint have lower consumption) for the convenience of lower distance \(d_{i,j}\). Solving for optimal \(S_i\) yields:

\[
S_i = \frac{\delta}{1 + \delta} (\omega + \Pi) \left( \frac{1 - \tau_L}{1 - \tau_S} \right) + \frac{1}{1 + \delta} \left( \frac{F_j}{1 - f_j} \right).
\]

\[\tag{25}\]

### 5.5.3 Competition between funds and the fund’s problem

Each of the \(N\) funds chooses \(f^v_j\), its variable fee, and \(F_j\), its fixed fee, to maximize profits \(\pi_j\). The fund understands that its fee choices affect both its market share, \(q_j\), and the saving choices of each of its investors, which we denote as \(S_j\),\(^{52}\) but it takes aggregate profits and competitors’ choices as given. Profit \(\pi_j\) is calculated as fee revenue, \((F_j + f^v_j S_j)q_j\), minus variable costs, \(cS_jq_j\), minus the fixed cost \(\phi\). Thus, the fund’s problem is:

\[
\max_{f^v_j, F_j} \left[ F_j + (f^v_j - c) S_j \right] q_j - \phi,
\]

\[\tag{26}\]

\(^{52}\)Conditional on having chosen fund \(j\), all of its investors (which are identical in all respects except for their position on the circle) will make the same consumption and saving choices.
where $f_j^v - c$ is the markup of variable fees over marginal cost. The fund’s first-order conditions are

\[
\frac{\partial \pi_j}{\partial f_j^v} = S_j \cdot q_j + \left[ F_j + (f_j^v - c) S_j \right] \cdot \frac{\partial q_j}{\partial f_j^v} + (f_j^v - c) \cdot \frac{\partial S_j}{\partial f_j^v} \cdot q_j = 0, \tag{27}
\]

\[
\frac{\partial \pi_j}{\partial F_j} = q_j + \left[ F_j + (f_j^v - c) S_j \right] \cdot \frac{\partial q_j}{\partial F_j} + (f_j^v - c) \cdot \frac{\partial S_j}{\partial F_j} \cdot q_j = 0. \tag{28}
\]

Raising variable fees $f_j^v$ by 1 basis point has three effects. First, the fund gains some revenue on its existing market share ($S_j \cdot q_j$); second, the fund loses some market share ($\partial q_j / \partial f_j^v < 0$) and the associated margin over variable costs ($F_j + (f_j^v - c) S_j$); third, the investors that choose to remain with fund $j$ may change saving ($\partial S_j / \partial f_j^v > 0$), so that the fund earns the markup over marginal cost on the additional assets.\footnote{This last term is a second-order effect due to the existence of fixed fees. If $F_j = 0$, we recover the familiar result under logarithmic utility that upon a change in variable fees $f_j^v$, income and substitution effects cancel out and saving is unaffected (see equation (25)).} Raising fixed fees $F_j$ by 1 cent has the same three effects. Equations (27) and (28) say that the fund’s optimal fee level is such that the three effects balance out.

### 5.6 Market equilibrium

Individuals choose the amount of optimal saving, and a fund to manage it, given the fee structures. Firms set their fee structure taking into account the functions determining consumer saving and firm choice. The market equilibrium described below is a set of values ($N, f^v, F, S_{0,i}, j$) such that funds are maximizing their profits, individuals are maximizing their utility, the individual and government budget constraints are satisfied, and markets clear. We start by solving for the equilibrium with $N$ fixed, and then later allow $N$ to be determined by endogenous entry.

#### 5.6.1 Equilibrium fees, consumption, and saving

We assume a symmetric equilibrium in which each fund has market share $q_j = 1/N$ and therefore total assets under management of $S/N$. We assume here that firms can choose both $f_j^v$ and $F_j$; later in the paper we consider what happens if firms are restricted to set one of these fees to zero.
In the Internet Appendix, we show that equilibrium fees are

\[ f_j^v = f^v = c, \]  

(29)

\[ F_j = F = \frac{2\gamma}{1 + 2\gamma + \delta} (1 - c) (\omega + \Pi) \frac{1 - \tau_L}{1 - \tau_S}. \]  

(30)

In equilibrium, the variable component of fees is equal to variable costs, i.e. there is no markup over variable costs; funds cover all fixed costs and extract all surplus via the fixed component of fees. This is a well-known result for two-part pricing schedules (see, e.g., Oi, 1971; Schmalensee, 1981; Tirole, 1988, Ch. 3). When variable fees equal variable costs, investors’ demand for asset management services is not distorted and funds have relatively more surplus available to extract. Firms charge fixed fees based on customers’ willingness to pay, which depends on lifetime resources and the extent to which fees are effectively deductible from taxes. Because of logarithmic distance disutility, fees depend on the number of firms only via \( \Pi \) (directly, and via the tax rates).

Since all individuals have the same values of \( C_{0,i}, C_{1,i}, \) and \( S_i \), aggregate values (\( C_0 \equiv \int_0^1 C_{0,i}di, \) \( C_1 \equiv \int_0^1 C_{1,i}di, \) \( S \equiv \int_0^1 S_idi \)) are equal to the individual values. Solving yields

\[ C_0 = \frac{1}{1 + 2\gamma + \delta} (\omega + \Pi) (1 - \tau_L), \]  

(31)

\[ C_1 = \frac{\delta}{1 + 2\gamma + \delta} (\omega + \Pi) (1 - \tau_L) (1 - c) (1 + r) (1 + \tau_M), \]  

(32)

\[ S = \frac{2\gamma + \delta}{1 + 2\gamma + \delta} (\omega + \Pi) \frac{1 - \tau_L}{1 - \tau_S}. \]  

(33)

where \( C_0, C_1, S, \) and each of the tax rates depend both on the regime (Roth or Traditional) and on the number of firms \( N \).

Combining and solving yields equations for aggregate fees, profits, and saving:

\[ F_j = F = \frac{\omega - \phi N}{1 + 2\gamma + \delta} \cdot \frac{1}{1 - c} \cdot \frac{1 - \tau_L}{1 - \tau_S} \cdot 1. \]  

(34)
\[ \Pi = \sum_j \pi_j = F - \phi N = \frac{\omega - \phi N}{1 + \frac{2\gamma}{1+2\gamma+\delta}} (1 - c) \frac{1}{1 - \tau_L} - \omega, \]  
(35)

\[ S = \frac{(2\gamma + \delta) (\omega - \phi N)}{(1 + 2\gamma + \delta) \frac{1}{1 - \tau_L} - 2\gamma (1 - c)}. \]  
(36)

Combining the above equations yields the result that

\[ F_j = F = \frac{2\gamma (1 - c)}{2\gamma + \delta} S \]  
(37)

i.e. fixed fees are proportional to saving. This implies that total fee revenue per unit of saving, \( f = f^v + F/S \), equals \((2\gamma + c\delta) / (2\gamma + \delta)\), a constant that is independent of both \( N \) and of whether we are under a Roth or Traditional system. This demonstrates that the assumption underpinning the results in Section 2 (that total fee revenue is proportional to saving) can emerge endogenously in a general equilibrium setting, even in the presence of fixed costs.

5.6.2 Tax policy

In order to satisfy its time-0 budget constraint, the government borrows an amount \( B \) at the market interest rate \( r \) to cover its deficit, so that

\[ B = G + S \tau_S - (\omega + \Pi) \tau_L. \]  
(38)

At time 1, the government taxes retirement income at a rate \( \tau_R \) to satisfy the time-1 budget constraint:

\[ B (1 + r) = [S(1 - f^v) - F] (1 + r) \tau_R, \]  
(39)

where \( f^v \) and \( F \) are the equilibrium values. Putting these together yields the government’s single (intertemporal) budget constraint:

\[ G = (\omega + \Pi) \tau_L - (F/S + f^v) \cdot (\tau_R \cdot S) - S(\tau_S - \tau_R), \]  
(40)
where the left side is government expenditure and the right side is the present value of government revenue.

Under Traditional, \( \tau_S = \tau_L \), so government revenue has three terms: the tax on labor income and profits, minus the fixed and variable fees it pays on its implicit account, minus the implicit government match on saving. The government has two policy variables (the labor income tax rate \( \tau_L^{Trad} \) and the retirement income tax \( \tau_R^{Trad} \)), and one intertemporal budget constraint, leaving it with one degree of freedom. We discuss what the government does with this degree of freedom below in subsection 5.6.5.

Under Roth, both \( \tau_S \) and \( \tau_R \) equal zero. Because the government receives no revenue at time 1, it cannot borrow \( (B^{Roth} = 0) \). Its only option is to balance the budget at time 0 by setting the tax rate on labor income equal to the ratio of government expenditure to the sum of wage income and profits:

\[
\tau_L^{Roth} = \frac{G}{\omega + \Pi^{Roth}}.
\]

Using the equations for optimal fees (29) and (37), optimal saving (33), and the expression for aggregate profits (35), we obtain:

\[
\tau_L^{Roth} = \frac{(G/\omega) (1 + \delta + 2c\gamma)}{(1 + \delta + 2\gamma)(1 - N\phi/\omega) - 2(1 - c)\gamma G/\omega} \tag{41}
\]

\[
\tau_L^{Trad} = \frac{(G/\omega) (1 + \delta + 2c\gamma)}{[1 + \delta (1 - c)] (1 - N\phi/\omega)} \tag{42}
\]

It is easy to show with some algebra that \( \tau_L^{Trad} > \tau_L^{Roth} \) for every \( N \) if \( c = 0 \), and for every \( N \) such that profits are positive if \( c > 0 \).

5.6.3 Comparison of equilibrium quantities at fixed \( N \)

At any given number of funds \( N \), one can see from (34) - (36) that fixed fees, profits, and assets under management are higher under Traditional than under Roth. The intuition is that at any given fee level the individual is less sensitive to fixed fees under Traditional than under Roth, because

\[54\] If profits are negative, it is possible for a lower tax rate to result in higher tax revenue.
under Traditional the government is effectively paying a fraction $\tau_R$ of the fixed fee (this occurs through the second term of (28). The higher fixed fees under Traditional (again holding $N$ constant), generates higher firm profits.

5.6.4 Entry

Next, we examine the entry decision. For any given $N$, the marginal fund exits if its profits are negative and enters if profits are positive conditional on entry. For example, if we start at positive profits, entry will cause $N$ to increase, and the resulting added competition and higher aggregate fixed costs together cause equilibrium fixed fees ($F$), saving ($S$), and aggregate profits ($\Pi$) to fall.

Since we assume that $N$ can be noninteger, funds enter the market until equilibrium profits $\pi_j$ equal zero, $j = 1, 2, \ldots N$. Equation (35) then implies that

$$N = \frac{\omega}{\phi} \cdot \frac{2\gamma}{1 + 2\gamma + \delta} (1 - c) \frac{1 - \tau_L}{1 - \tau_S}.$$

(44)

Total employment in the asset management industry ($L$) is found by substituting (44) and (33) into (10):

$$L = \frac{\phi N + cS}{\omega} = \frac{c\delta + 2\gamma}{2\gamma + \delta + 1} \frac{1 - \tau_L}{1 - \tau_S}.$$

(45)

5.6.5 Taxes with entry

With $\Pi = 0$, equation (41) implies $\tau_L^{Roth} = c/\omega$. As described above, under Traditional the government has one degree of freedom. To simplify comparison between the two account types, we assume that $\tau_L^{Trad} = \tau_R^{Trad}$ so that $\tau_M^{Trad} = 0$, i.e. the implicit extra government match on Traditional due to differences in tax rates is zero. Under this additional assumption, using the closed-form expressions for $F$, $f^v$, and $S$ we obtained above, and setting $\Pi$ to zero, the government’s budget
constraint (40) implies

$$\tau_{L}^{\text{Trad}} = \frac{G}{\omega} \cdot \frac{1 + \delta + 2\gamma}{1 + \delta (1 - c)} > \frac{G}{\omega} = \tau_{L}^{\text{Roth}},$$

so that

$$\frac{\tau_{L}^{\text{Trad}}}{\tau_{L}^{\text{Roth}}} = \frac{1 + \delta + 2\gamma}{1 + \delta (1 - c)} > 1.$$  (47)

Thus, tax rates on labor income need to be higher under Traditional than under Roth, in order to pay for the extra asset management fees the government is incurring under Traditional.

5.6.6 Comparison of equilibrium quantities with entry

The equations above imply that in the equilibria with free entry and zero profits,

$$\frac{S^{\text{Trad}}}{S^{\text{Roth}}} = \frac{F^{\text{Trad}}}{F^{\text{Roth}}} = \frac{F^{\text{Trad}} + f S^{\text{Trad}}}{F^{\text{Roth}} + f S^{\text{Roth}}} = \frac{N^{\text{Trad}}}{N^{\text{Roth}}} = \frac{S^{\text{Trad}}}{S^{\text{Roth}}} = \frac{L^{\text{Trad}}}{L^{\text{Roth}}} = \frac{1}{1 - \tau_{L}^{\text{Roth}}} > 1.$$  (48)

Under Traditional, the equilibrium saving, fixed fees, total fees, number of firms, saving, and employment in the asset management industry are all equal to their corresponding values under Roth, scaled up by a factor $1/ (1 - \tau_{L}^{\text{Roth}}) > 0$.

Note also that

$$\frac{C_{0}^{\text{Trad}}}{C_{0}^{\text{Roth}}} = \frac{C_{1}^{\text{Trad}}}{C_{1}^{\text{Roth}}} = \frac{1 - \tau_{L}^{\text{Trad}}}{1 - \tau_{L}^{\text{Roth}}} = \frac{1 - \frac{G}{\omega} \cdot \frac{1 + \delta + 2\gamma}{1 + \delta (1 - c)}}{1 - \frac{G}{\omega}} < 1.$$  (49)

i.e. consumption under Traditional is proportionately lower than under Roth due to the higher taxation under Traditional. Note that the higher are the variable costs $c$, the lower is consumption under Traditional relative to that of Roth.

Note that the tax rate is not guaranteed to be less than 100%, i.e. this policy is not necessarily feasible. The resources the government needs to fund Traditional accounts create a practical limit on expenditure $G$. The higher $\gamma$, the higher the fees individuals are willing to pay, the larger the government’s transfer to asset managers, and the fewer resources available for public expenditure.
5.7 Roth vs Traditional: Welfare analysis

We denote welfare, defined in (13), as

\[ U^k = \ln C_0^k + \delta \ln C_1^k + \gamma \ln N^k + \gamma (1 + \ln 2), \]

(50)

where \( k \in \{*, Roth, Trad\} \) indicates the respective quantities under the planner solution \((U^*)\), a Roth market equilibrium \((U^{Roth})\), and a Traditional market equilibrium \((U^{Trad})\), respectively. The last term \((\gamma (1 + \ln 2))\) is a constant which we henceforth omit for brevity.

We begin by comparing social welfare under Roth and Traditional:

\[ U^{Trad} - U^{Roth} = \ln \left( \frac{C_0^{Trad}}{C_0^{Roth}} \right) + \delta \ln \left( \frac{C_1^{Trad}}{C_1^{Roth}} \right) + \gamma \ln \left( \frac{N^{Trad}}{N^{Roth}} \right) \]

(51)

The two terms of (51) have different signs. The first term is negative because higher total asset management costs under Traditional relative to Roth lead to higher taxes in the model, resulting in proportionally lower consumption and welfare under Traditional. However, the second term is positive because Traditional also generates a larger number of asset management firms, decreasing average distance and contributing to higher welfare.

In the Internet Appendix, we prove that the sum of the two terms is always negative, i.e. the first term always dominates the second so that social welfare in the model is always lower under Traditional than Roth.\(^{56}\)

To understand this result, note that

\[ \frac{N^{Roth}}{N^*} = 2 \left( \frac{1}{1 + \frac{\gamma}{1 + \delta + \gamma}} \right) > 1 \]

(52)

\(^{56}\)In Section 1 of the Internet Appendix we show that Roth has higher social welfare than Traditional even if the government can choose \( \tau_L^{Trad} \) and \( \tau_R^{Trad} \) optimally, resulting in a slightly increasing time path of tax rates \((\tau_R \geq \tau_L, \text{ holding with equality if } c = 0)\).
i.e. the number of firms under the Roth equilibrium is strictly greater than under the planner, and welfare in the model is correspondingly lower. If \( N^{Roth} \) were instead exogenously set equal to \( N^* \), funds would make positive profits,\(^{57}\) which would induce additional firms to enter. This is the well-known “business-stealing” effect (see, e.g., Mankiw and Whinston, 1986): if an entrant causes incumbents to reduce output, entry is more desirable to the entrant than it is to society, resulting in a tendency toward excessive entry. The market equilibrium under Traditional has even more firms than under Roth. This effect, together with higher variable costs devoted to asset management (if \( c > 0 \)) and an unaltered Euler equation for consumption, generates lower social welfare in the model than under Roth.

To provide further intuition, we take a log approximation of (51) and substitute in the values for \( T_{L}^{Trad} \) and \( T_{L}^{Roth} \) to yield

\[
U^{Trad} - U^{Roth} \approx -T_{L}^{Roth} \left\{ \frac{(1 + \delta)}{1 + \delta (1 - c)} \right\} \cdot \left( \delta c + 2\gamma - \gamma \right).
\]

Consider first the case of \( c = 0 \), (no variable costs), in which case \( U^{Trad} - U^{Roth} \approx -T_{L}^{Roth}\gamma \) which is clearly negative.\(^{58}\) One might be tempted to think that with all costs fixed, each firm could simply absorb the extra assets under management associated with Traditional, incur no additional costs, and keep fixed and variable fees the same (thereby reducing total fees as a percentage of assets under management), thus leaving profits unchanged at zero. In this case, there would be no additional resources devoted to asset management, and thus no welfare loss under Traditional. However, this would not be a market equilibrium, because given \( N \) it is profit-maximizing for firms to charge higher fixed fees \( F \) (see section 5.6.3) and the resulting positive profits would generate entry and correspondingly lower welfare.

If we assume that variable costs are greater than zero, the welfare loss under Traditional is even greater because the larger assets under Traditional cause additional real resources to be devoted to asset management. Mathematically, comparing an equilibrium with \( c > 0 \) to one with \( c = 0 \), the ratio of the number of funds \( (N^{Trad} / N^{Roth}) \) is the same, but the ratios of time-0 and time-1 consumption are lower when \( c > 0 \) because taxes are higher.

\(^{57}\) \( \Pi^{Roth, N=N^*} = (\omega - G) \frac{(1 - c) \gamma}{(1 + \delta + \gamma)} > 0 \)

\(^{58}\) As suggested by the approximation, welfare is lower under Traditional than Roth even for some \( c < 0 \), e.g. if funds have positive net alpha after variable costs, as is assumed for instance in Berk and Green (2004) (ignoring for simplicity the possibility that either gross alpha or costs vary with fund size). We thank Michael Halling for bringing to our attention the relationship between that paper and ours.
5.7.1 Alternative specifications of costs and fees

So far we have discussed the case in which firms face both fixed and variable costs and are free to set both fixed and variable fees (with zero variable costs a special case) and showed that firms choose to set variable fees equal to marginal costs \( f = c \). What if firms are instead restricted to charge either variable fees or fixed fees, but not both? In the Internet Appendix, we solve the model under these restrictions (and also consider the other possible combinations of fixed and variable costs and fixed and variable fees). The business-stealing effect exists independent of fee and cost structure, and in each case it steers society towards an overly large asset management sector \( N^{Roth} > N^{*} \). However, if there is a mismatch between the fee structure and the cost structure, another friction arises. For example, with fees restricted to be fixed-only \( f = 0 < c \), individuals do not internalize the cost of managing assets in their saving decision and save “too much” (the Euler equation is now different from the planner’s, because \( 1 - f \neq 1 - c \), and the optimal \( C_0 \) under Roth is lower than what the planner would choose.) These additional assets further increase the number of equilibrium funds, resulting in additional hypertrophy and even greater welfare losses. Conversely, if there are fixed costs but fees are variable-only, funds must set \( f > c \) in order to cover their fixed costs, and individuals save “too little,” starving the asset management sector of assets and offsetting the business-stealing effect. In this case, it may appear beneficial to increase aggregate saving and the number of funds. However, a direct comparison of maximized utilities shows that under a reasonable calibration a switch from Roth to Traditional is still welfare-destroying in the model.\(^{59}\)

The intuition is that Traditional does not create a bigger saving subsidy than Roth in our model, and therefore it does not solve the undersaving problem; it merely exacerbates the individuals’ price insensitivity because it subsidizes fees while leaving the individual’s consumption/saving tradeoff intact.

\(^{59}\)Namely, even assuming no variable costs \( c = 0 \), maximized utility under Roth is higher than under Traditional unless \( \gamma \) is larger than roughly 0.4, implying that over one-third of total resources in the economy are devoted to asset management. This result is independent of the size of firm-level fixed costs \( \phi \).
6 Conclusion

Under some simplifying assumptions about tax rates, a standard benchmark model yields an equivalence result between front-loaded (Roth) and back-loaded (Traditional) taxation of retirement savings. Individuals’ consumption in each period is the same under Roth and as it is under Traditional, and the present value of government tax revenues is the same under the two systems. However, the timing of taxation is different, and as a result back-loaded taxation leads to higher outstanding government debt and a correspondingly greater amount of retirement assets. These additional assets represent an implicit government portfolio, i.e., resources earmarked to pay future taxes when the money is distributed from the account. In this paper, we add one crucial bit of realism to the benchmark model: asset management fees. We show that the equivalence result breaks down because the government is paying an estimated $19.5 billion a year in fees on its $3 trillion implicit portfolio. This represents an added demand for asset management services. To examine how this added demand affects equilibrium fees, real resources devoted to asset management, and asset management profits, we develop a simple, stylized general equilibrium model. We show that welfare in that model is lower under Traditional than under Roth because back-loading taxation inefficiently increases the amount of resources spent on asset management. The size of the welfare loss depends on the degree to which asset management costs are fixed or variable. With only variable costs, the partial equilibrium cost results apply in general equilibrium, and all of the extra costs generate welfare losses. At the other extreme, we consider the case when all costs are fixed. Holding the number of firms constant, because the government effectively pays part of the fees under Traditional and because firms have some market power, firms charge higher total dollar fees under Traditional than under Roth, resulting in higher profits for the asset management industry. If entry is allowed, the higher profits under Traditional lead to a greater number of firms, and thus inefficiently high aggregate resources devoted to asset management.

Our results raise the policy issue of whether governments should encourage or possibly mandate wider adoption of Roth. In the U.S., a switch to Roth for new contributions (“Rothification”) was discussed during the debate leading up to the 2017 tax reform, although it was ultimately not
enacted. Our model, taken literally, would imply that a switch is beneficial. However, both the benchmark model and our model abstract from other potential drivers of the policy choice between front-loaded and back-loaded taxation of retirement savings.

First, most real-world tax systems have a progressive schedule of tax rates with marginal tax rates that are increasing in income. This means that $\tau_L$ and $\tau_R$ are not fixed flat rates but determined by uncertain future tax schedules and individual taxable incomes during working and retirement years, complicating the analysis in a number of ways (Brown et al., 2017). With progressive taxation, lifetime taxes are more aligned with lifetime income under Traditional than under Roth.\textsuperscript{60} In addition, if a worker’s lifetime income is not known in advance, Traditional coupled with progressive taxation may work as insurance because the average tax rate on distributions is higher when the account balance is higher. Finally, under progressive taxation a switch to Roth would eliminate the government “match” that arises for those individuals who have lower taxable income (and thus lower tax rates) in retirement than in working years.\textsuperscript{61} But a match through Traditional involves costs beyond the match itself: to offer a match of $\tau_L - \tau_R$ through Traditional requires that an extra $\tau_L$ be put into an implicit government account, leading to extra fees compounded for the average life of the account. If a government “match” is deemed an important policy goal, the government could instead offer a direct match on Roth retirement accounts (i.e., grant an extra $m$ cents for every dollar contributed to the account, as is done for Lifetime ISAs in the United Kingdom) and avoid the extra fees associated with the added implicit government account.

Second, behavioral biases that cause people to save too little are a frequently-cited motive for the provision of retirement saving incentives and could affect the relative desirability of the two systems. Behavioral arguments cut both ways. During an individual’s work life, if individuals ignore or underestimate the future “tax bite” under back-loaded taxation, this would lead to lower

\textsuperscript{60}For example, consider two workers with the same lifetime income: one with high annual earning and a short work life (e.g. a firefighter), and another with lower annual earnings but a longer work life (e.g. a librarian). Under Roth, the firefighter would pay more lifetime taxes than the clerk. Under Traditional, the gap between the lifetime taxes paid by the two workers will shrink and potentially disappear.

\textsuperscript{61}There would be distributional consequences of such a shift. Individuals with high labor income but little expected income in retirement would be disproportionately affected, as their current marginal tax rate is high and their tax rate in retirement is low.
saving than front-loaded taxation (Iwry and John, 2009). However, back-loaded taxation could also provide a more powerful behavioral response because of the “instant gratification” of an immediate tax benefit (Feenberg and Skinner, 1989; Bernheim, 2002). Beshears et al. (2017) find empirically that Roth induces individuals to save more, and argue that this is because individuals underweight future taxes and focus on nominal contributions and savings rather than on consumption. In addition, during the individual’s retirement, under progressive taxation a Traditional system penalizes large withdrawals with higher tax rates. As part of the recent British debate, an Economist editorial claims that this feature “is actually quite useful in that it stops people blowing their pension pot in a spending spree at 65” (Buttonwood, 2015). Of course, the other side of the coin is that Traditional penalizes individuals who withdraw funds in bulk for legitimate reasons such as hardship or investment. We are not aware of any systematic study of this tradeoff.

Third, there are political economy considerations that are important to the debate over a shift from Traditional to Roth. U.S. budget rules make it more cumbersome to pass bills that increase the total budget deficit over a five- or ten-year window. A transition from Traditional to Roth generates more cash flow upfront and less when the relevant workers retire, thus bringing more revenue into the budget window, resulting in a temporary deficit reduction which could ease the passage of other legislation that involved lower taxes or higher spending. This additional short-run fiscal flexibility may or may not be considered desirable, but it certainly makes Roth attractive to many real-world policymakers. Indeed, one of the purported motivations for originally proposing Roth accounts in the U.S. was to help “fund” cuts in the capital gains tax (Pine, 1989).

---

62 Our results rely on the assumption that individuals are rational savers and therefore under our benchmark model contribute enough extra dollars into a Traditional plan relative to what they would contribute under a Roth plan to ensure that retirement consumption would be the same under the two plans. Beshears et al. (2017) provide evidence that individuals do not adjust their retirement savings in this way, but instead find that contributions under a Roth 401(k) plan are similar to those under a Traditional 401(k) plan, implying a higher retirement consumption under a Roth plan. If these findings generalized to the policy experiments we consider, they would complicate our welfare analysis, but the gist of our argument would still be valid. Roth is more cost effective than Traditional. If the total amount of assets is constant under Traditional and Roth, then Roth delivers a larger savings subsidy for the same cost to the government. At the other extreme, if, as in our paper, the total amount of retirement consumption is constant, then Roth delivers the same savings subsidy for a lower cost to the government.

63 The effectiveness of this approach is complicated by the “Byrd rule,” which requires a supermajority to approve any deficit increases that could occur beyond the period covered by the budget resolution (Committee for a Responsible Federal Budget, 2016).
Finally, as discussed in Section 2, Traditional accounts lead to a proportional increase in retirement assets via the creation of an implicit government account. This account represents an increase in the overall demand for asset management services, including actively managed funds and equity funds. The added demand for actively managed funds could be beneficial if a market equilibrium under Roth were to feature too little price discovery relative to the social optimum, and if actively managed retail funds were a cost-effective vehicle to do price discovery. The added equity exposure could be beneficial if the government desired it and were unable to obtain it in a more direct and cost-effective way, but it could lead to excess equity exposure as well. Part of our contribution is to highlight that currently, the U.S. government indirectly owns about $2 trillion in equities via tax-deferred retirement accounts. We have no evidence that this is a conscious policy choice, and certainly there has not been a public debate about it.

Our analysis raises some additional policy issues. The $19.5 billion cost to the government exists because the government owns an implicit account that incurs substantial investment fees. One way to reduce this cost is to shrink or eliminate this account by switching to Roth, as discussed throughout the paper. An alternative approach would be to leave the size of the account unchanged, but to reduce the level of investment fees. An example of a policy aimed at this goal was the U.S. Department of Labor fiduciary rule implemented in June 2017 but then subsequently struck down by the courts. One of the stated motivations for the rule was protecting individuals with retirement accounts from aggressive marketing of high-fee products—especially senior investors as they consider rolling over their employer plan savings into an individual retirement account. If a fiduciary rule had the effect of reducing fees incurred by individual investors, our results suggest that it would also protect the government’s future revenue from being eroded by high fees, providing a possible additional rationale for implementing such a rule.
References

Abel, A.B., 2001. The effects of investing Social Security funds in the stock market when fixed
nomics 42, 447–469.
tax-favoured retirement plans. doi:10.1787/138080145732. OECD Economics Department
94, 155–160.
Ayres, I., Curtis, Q., 2015. Beyond diversification: The pervasive problem of excessive fees and
io/indices/sites/2/2017/03/Index-Methodology-2017-03-17-FINAL-FINAL.pdf on
Markets: Competition Versus Regulation. Rochester Studies in Managerial Economics and Pol-
Economics.
Political Economy 112, 1269–1295.
and Savings. pp. 1173–1249.
Beshears, J., Choi, J.J., Laibson, D., Madrian, B.C., 2017. Does front-loading taxation increase


Buttonwood, 2015. EET your TEE, George. The Economist.


Gil-Bazo, J., Ruiz-Verdú, P., 2008. When cheaper is better: Fee determination in the market for


search 22, 331–339.
Lurie, I.Z., Ramnath, S.P., 2011. Long-run changes in tax expenditures on 401(k)-type retirement
Perspectives 27, 97–108.
Mankiw, N.G., Whinston, M.D., 1986. Free entry and social inefficiency. RAND Journal of Eco-
nomics 17, 48–58.
Novick, B., Medero, J., Rosenblum, A., Barry, R., 2016. Breaking down the data: A closer look at
bond fund AUM. Blackrock white paper.
OECD, 2015a. Stocktaking of the tax treatment of funded private pension plans in OECD and EU
countries.
OECD, 2015b. The tax treatment of funded private pension plans - OECD and EU country profiles.
Osborne, G., 2015. Strengthening the incentive to save: a consultation on pensions tax relief. Her
Majesty’s Treasury Cm 9102.
Political Economy.
74, 21–36.
Quantitative Analysis 44, 1013–1044.
Philippon, T., 2015. Has the U.S. finance industry become less efficient? on the theory and mea-


