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: November 4, 2018

Paper available at `https://ssrn.com/abstract=3046077`
Internet Appendix available at `https://ssrn.com/abstract=3138303`

Keywords: mutual fund 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, Daniel Hemel, Charles Jones, David Laibson, Brett Myers, Emi Nakamura, 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, ESSFM 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 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) – Under some simplifying assumptions, both individuals and the government are indifferent between traditional tax-deferred retirement accounts and "front-loaded" (Roth) accounts. We add investment fees to this standard benchmark and show that individuals are still indifferent but the government is not. We estimate that, by deferring tax revenue, the U.S. government pays $16.1 billion in annual fees, representing an implicit subsidy to asset managers. We also show that this result continues to hold when asset managers choose the fee level competitively in equilibrium. 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 indifference 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. We estimate this to cost the U.S. government $16.1 billion per year, representing an implicit subsidy to the asset management industry. We then ask whether this result holds in general equilibrium when asset managers choose the fee level competitively. The answer depends both on the nature of the cost function for asset management services, and on the nature of market competition, but we find that the result will in general continue to hold: back-loaded taxation is more expensive for the government, and it produces a larger asset management industry. We also find that back-loaded taxation reduces social welfare in our model.

1 Introduction

Retirement savings systems around the world incorporate tax incentives designed to increase saving and enhance retirement security. One dimension along which these incentive schemes vary is the timing of taxation. The traditional way to structure these incentives is through tax deferral—exempting contributions to retirement accounts from current income taxation and then taxing the principal and returns upon withdrawal. This "back-loading" of taxation benefits the investor, because asset returns (interest, dividends, and capital gains) can be earned on the deferred taxes, yielding a higher amount of resources during retirement than would occur in the absence of the tax deferral.

Although a large majority of assets is still held in traditional accounts, "front-loaded" schemes are becoming increasingly widespread. Under these schemes, contributions are made with after-tax income, but then neither the principal nor returns are taxed at any point in the future. In the U.S., the defined contribution retirement system began with the introduction of traditional accounts with back-loaded taxation, including both employer-based accounts ("401(k)s") and individual accounts ("IRAs"). Accounts with front-loaded taxation (named "Roth" accounts 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.[1] In 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 text (Public Law 115-97 of 12/22/2017). A similar pattern occurred in Canada and the U.K., who started with traditional 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 scheme is preferable (Osborne, 2015; Buttonwood, 2015).

Much of the recent U.S. debate focused on the political economy aspects of the choice, i.e., whether front-loading tax revenue with a Roth scheme 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: record-

[1] 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 as well.

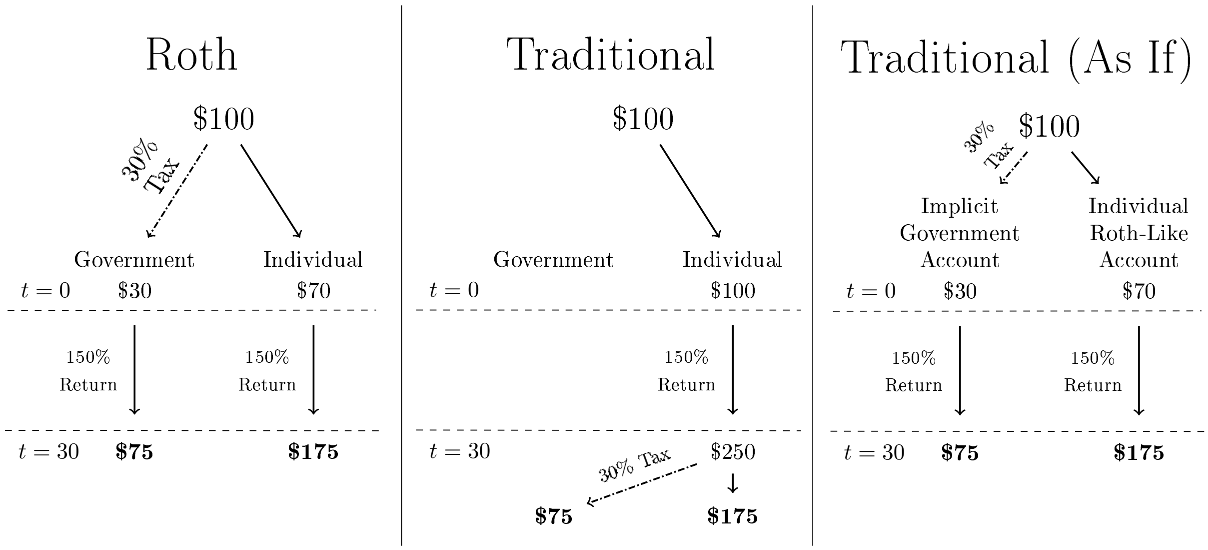

Figure 1: **Benchmark indifference result.** Traditional is equivalent to a Roth account plus an implicit government account equal to the balance of deferred taxes. Without asset management fees, the two accounts yield the same cash flows for individuals, and the same future value of cash flows for the government. A 150% return is approximately equal to the total return on a 30-year Treasury bond (3.10% for 30 years).

keepers, asset managers, and financial advisors charge fees for running retirement plans, managing assets, and advising clients. These fees are typically charged as a percentage of assets under management (AUM). In addition, asset managers incur trading costs on behalf of clients. (For the sake of exposition, we lump these costs together with "investment fees" except where we explicitly treat them as separate items). By deferring tax revenue with a traditional scheme, the government pays investment fees on the substantial amount of assets that sit in retirement accounts waiting to pay future taxes.

Under a few simplifying assumptions, including the constancy of the tax rate across working and retirement years, some basic math shows a standard benchmark indifference result: in the absence of fees, (i) individuals are indifferent between Traditional and Roth accounts, which yield identical cash flow streams; and (ii) the timing of government cash flows differs across accounts, but the future (or present) value of tax revenue is the same.[2] The first two panels of of Fig. 1 provide a simple numerical example of this. The result can be intuitively understood 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

[2]This benchmark result abstracts from differences that exist across Traditional and Roth in features such as contribution limits, withdrawal penalties and required minimum distributions. We briefly discuss these features in Section 2.

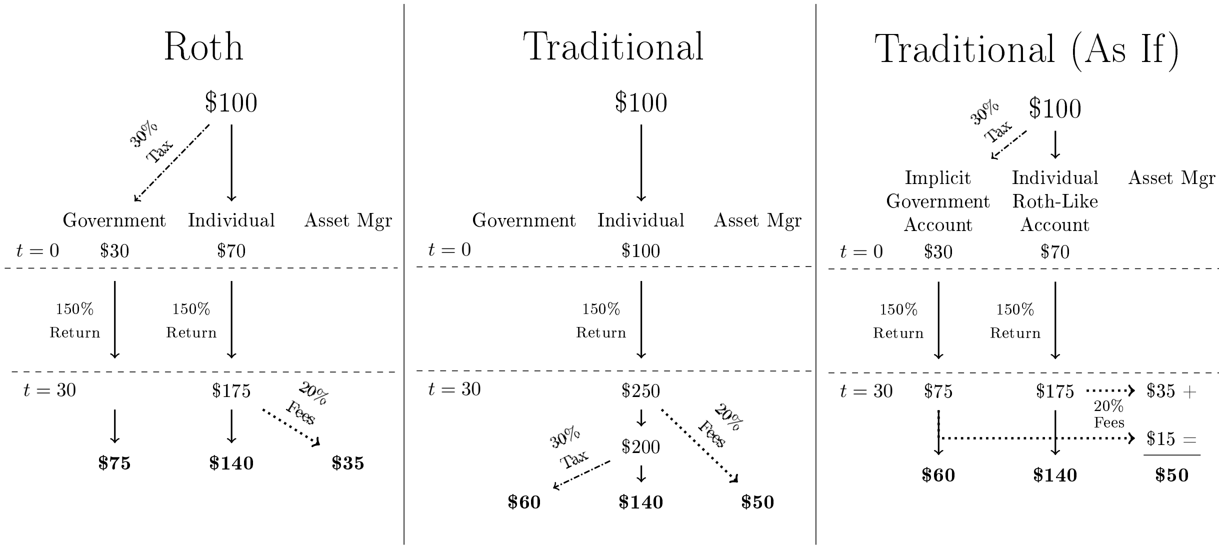

Figure 2: **Fee nonneutrality.** Traditional is equivalent to a Roth account plus an implicit government account equal to the balance of deferred taxes. With asset management fees, the two accounts yield the same cash flows for individuals, but the Roth account yields a superior future value for the 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 0.66%/year (the value we calibrate in this paper) for 30 years.

$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 as in the third panel of Fig. 1.

We add fees to this benchmark model and show that, while individuals remain indifferent, the timing of taxation does affect the future value of government revenues. This is shown in Fig. 2. Because under Roth the government levies taxes upfront, the account size is smaller, and therefore total fees are also smaller. The third panel shows again the decomposition into a Roth-like account and an implicit government account, which also pays fees. In essence, the government is paying investment fees on its share of Traditional defined-contribution retirement accounts, something it does not incur with Roth accounts.

It is possible, of course, that the additional fees that the government pays on its implicit portfolio are compensation for services provided by asset managers, possibly including better performance. However, we deem it unlikely that the government captures these benefits because it implicitly holds a fraction of all retirement portfolios, and many of the potential benefits will cancel out in the aggregate. For example, some of the higher costs might be associated with creating customized funds or asset allocations, such as target-date retirement funds or industry funds. While this might create value for individuals, holding *all* target-date funds or industry funds does not create value for the government. A similar argument applies to active asset management, at least with regard to equity funds. 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."

The U.S. government's implicit account is large. We estimate its size as the total amount of tax-deferred assets in DC plans and IRAs ($15.4 trillion) times 20%, a reasonable approximation of the average marginal tax rate in retirement, 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 66 basis points (bps). We assume that 21% of fees paid by the government are recovered via corporate taxation of the asset managers. Multiplying $3 trillion by $.66\% \times (1 - .21)$, we reach our estimate of $16.1 billion per year—a cost for the government, and an annual subsidy to the asset management industry.

Our estimate of investment fees is itself a contribution, as it is to our knowledge the first estimate that is both comprehensive and asset-weighted. Most published estimates of fees and overall net-of-fees performance are fund-weighted, usually for the purpose of assessing fund manager skill. Similarly, most published estimates of trading costs are volume-weighted, usually for the purpose of assessing trader execution skill or overall market liquidity. However, lower-cost funds attract more AUM. We estimate the total investment fees incurred by the average dollar invested in DC plans and IRAs, a large fraction of all financial assets.

Our main result and the above back-of-the-envelope calculation rely on the partial equilibrium assumption that fees as a percentage of AUM are constant as AUM vary, i.e., that marginal fees equal average fees. Under this assumption, if all Traditional accounts are converted to Roth, total dollar fees collected by asset managers drop proportionally to aggregate assets. This assumption is consistent with recent evidence that an increase in the scale of the financial sector has not resulted in economies of scale, or at least these economies of scale have not been passed along to the consumers of financial services (Greenwood and Scharfstein, 2013; Malkiel, 2013; Philippon, 2015).

Whether such a scenario emerges endogenously in a general equilibrium model depends on

[3]Our estimate of assets excludes the federal government's Thrift Savings Program (TSP), whose fees are negligible. It also excludes 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%.

both the underlying cost structure and the competitive landscape of the asset management industry. 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, or in a small open economy in which an increase in AUM has a negligible effect on the scale at which asset managers operate.

Less obviously, our result could also hold even with economies of scale in a closed economy. To show this, we formally examine a two-period, general equilibrium model in which the production of asset management services entails fixed costs and therefore increasing returns to scale. Intuitively, absent any frictions, the existence of economies of scale would lead to a monopoly, which is inconsistent with the large observed number of asset management firms. We therefore introduce two additional realistic assumptions. First, consumers are not fully sensitive to the level of fees (see, e.g., Bergstresser et al., 2009; Gil-Bazo and Ruiz-Verdú, 2009; Henderson and Pearson, 2011; Pool et al., 2016; Cremers et al., 2016), and second, entry is free. Specifically, we model competition among asset management firms ("funds") as monopolistic competition with differentiated products (Salop, 1979). In this model, a switch from Roth to Traditional continues to increase assets under management. Moreover, equilibrium percent fees remain constant even though our cost function specification nests 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. Constant percent fees results in an increase in (i) the aggregate dollar fees collected, (ii) the equilibrium number of funds, and (iii) employment in the asset management industry. This result contributes to the recent literature on the size of the financial services industry (Philippon and Reshef, 2012; Greenwood and Scharfstein, 2013; Malkiel, 2013; Gârleanu and Pedersen, 2018; Bolton et al., 2016) by showing one possible source of the growth of the financial sector—an implicit government subsidy.

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. As a thought experiment, consider a shift from Roth to Traditional. If a market equilibrium under Roth were the social optimum, the shift would result in excess resources devoted to asset management. The total welfare loss would be relatively small, however, because we are departing from the optimum where the social welfare function is flat. However, a market equilibrium under Roth departs from this hypothetical optimum for two reasons. First, monopolistic competition à la Salop results in excess product differentiation, boosting overall demand for asset management relative to the social optimum. Second, the fact that fees are charged as a percent of total assets causes individuals to consume too much at time 0 and thus save too little, because if they save more they pay higher dollar fees. In turn, lower savings curb demand for asset management. Depending on which effect prevails, a market equilibrium under Roth may entail too many or too few resources devoted to asset management compared to a hypothetical optimum.[4] However, regardless of which effect prevails, a switch from Roth to Traditional is never beneficial in our model because it does not create a bigger saving subsidy; it merely exacerbates the individuals' price insensitivity. A switch increases the resources devoted to asset management at the expense of time-1 consumption (which is already too low), while it does nothing to reduce time-0 consumption (which is too high), bringing society even farther from the hypothetical optimum, starting at a point where the social welfare function is already steep.

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. We see our main contribution as pointing out an important difference between Roth and Traditional retirement accounts: Roth accounts have higher present value of tax revenue because Traditional accounts, as currently structured, contain an implicit subsidy to the financial industry—the fees on \$3 trillion of retirement savings that would not exist under Roth. However, there are potentially other important factors not captured by our model that could affect the relative desirability of the two types of account. One possibility is that active asset managers create a positive externality because they make prices more efficient ad do not fully capture the value they create by trading. In this case, a subsidy may result in more price discovery.[5] Other factors include progressive taxation, behavioral biases, and political economy considerations. Potential policies include attempting to reduce the overall level of fees, either directly, by leveraging the bargaining power arising from its large implicit account, or indirectly, via stricter fiduciary standards for retirement savings accounts. We address these public policy issues in the conclusion.

[4]Welfare analysis of a model with lump-sum fees would be more straightforward: Roth would have too many resources devoted to asset management, and Traditional even more. However, we choose to have percent fees in our model because in reality fees *are* charged as a percent of assets under management. This too is a simplification, because asset managers do charge lower percentage fees to larger clients, but percent fees are a much better representation of reality than lump-sum fees. This choice enables us to study the interaction between fees, retirement saving subsidies, and the consumption/saving decision.

[5]For a discussion, see Wurgler (2011). However, note that Traditional accounts are a blunt tool to subsidize price discovery because the extra assets benefit active and passive managers indiscriminately. Moreover, 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.

Finally, our result is also related to an extensive literature showing that the government may be able to borrow more efficiently than private agents (e.g., Barro, 1974; Woodford, 1990; Aiyagari and McGrattan, 1998; Holmström and Tirole, 1998; Heathcote, 2005). Here we provide an example in which the government can *save* more efficiently by simply not borrowing.

Our paper is structured as follows. Section 2 derives the basic result that the investor and the government are indifferent between Traditional and Roth in a benchmark partial-equilibrium model. We then introduce investment fees, and show that the basic indifference result still holds for the investor, but not for the government. In Section 3 we construct an asset-weighted estimate of total investment fees applicable to retirement accounts and use it to estimate the difference between total fees under Traditional and under Roth. Section 4 examines a simple general equilibrium model in which the size of the asset management industry and fees as a percent of AUM are determined endogenously. Section 5 examines the implications for public policy and concludes.

2 The benchmark model and the impact of fees

In this section we begin by describing the standard indifference 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. We show that, together, these results imply that the economic equilibrium is the same if only Roth or only Traditional accounts are allowed, even though assets under management are larger under Traditional.

Next, we add fees to the model, assuming that asset managers provide their services in exchange for a fee, charged as a fixed proportion of account size, f. Under this assumption, we show that the indifference result breaks down and the economic equilibrium is no longer the same because higher assets under management translate to higher total fees.

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 4, we relax this assumption and

[6]For instance, suppose that the same asset managers operate in countries A and B, but A's retirement

Account type	Label	Type of taxation	Tax on initial contribution	Tax rate on investment returns	Tax on retirement payouts
Taxable	TTE	Front-loaded	τ_L	τ_I	0
Traditional	EET	Back-loaded	0	0	τ_R
Roth	TEE	Front-loaded	τ_L	0	0

Table 1: **Different tax treatment of retirement savings.** Money earned and saved for retirement can be taxed at three points: when earned, when it earns returns on investment, and when paid out of the account in retirement. Each type of account is represented by a three-letter abbreviation. For instance, a common taxable account is "TTE" because earned income is taxable, investment returns are taxable, but account distributions in retirement are exempt.

allow asset managers to optimally choose the level of fees.

2.1 Base assumptions and notation

We assume two constant proportional tax rates, one for labor income (τ_L) and one for retirement income (τ_R). Neither tax rate varies with the level of income.[7] Table 1 compares three possible account types: Taxable, Traditional and Roth. Following convention, we label these accounts TTE, EET, and TEE, respectively.[8]

- *Taxable account* (TTE): all labor income is taxed at rate τ_L when earned. Intermediate investment returns are taxed at a rate $\tau_I > 0$. This scheme is referred to as TTE, because the earned income is taxable, investment returns are taxed, and account distributions in retirement are exempt.

- *Traditional retirement account* (EET): income tax on retirement account contributions is deferred until retirement, when the account is assumed to be paid out as 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 to \$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. We briefly address these complications in the conclusion.

[8] This notation is standard in publications by the World Bank (Holzmann and Hinz, 2005; Whitehouse, 2007) and the OECD (Antolín et al., 2004; OECD, 2015a).

income, taxed at a rate τ_R. Intermediate investment returns are not taxed. This scheme is referred to as EET because the earned income put into the account is exempt, the returns are exempt, and the full amount of the retirement account is taxed on withdrawal.

- *Roth retirement account* (TEE): all labor income is taxed at rate τ_L when earned. Intermediate investment returns are not taxed. This scheme is referred to as TEE because the earned income is taxable, the returns on investment are exempt, and account distributions in retirement are exempt.

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

We assume that an individual's savings are held in an account and invested in the only one asset in positive supply, government bonds, paying a known return of r. We also assume that the individual has the same pretax labor earnings under either account type, and we abstract for now from details such as contribution limits, withdrawal penalties, and required minimum distributions.[9] Finally, we initially guess (and later verify in the resulting equilibrium) that individuals' initial consumption and the interest rate they earn on the assets in their retirement account are the same under Roth and Traditional. Together, these assumptions imply that the individual directs the same amount of pretax labor earnings to account contributions regardless of account type. We normalize this amount to $1.

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 τ_R of the balance, and the individual receives the remaining fraction $1-\tau_R$. With a Roth account, the government taxes the money upfront and receives τ_L, so the individual's starting balance is $1-\tau_L$. No additional taxation happens, and therefore at time T the individual keeps the entire balance, $(1-\tau_L)(1+r)^T$.

[9] For clarity of exposition we abstract from features that may make one or the other account more attractive to the individual in practice. Advantages of Traditional include the individual's ability to take advantage of low marginal tax rates by timing withdrawals, converting to Roth, and until recently recharacterizing a Roth conversion. Advantages of Roth include less restrictive contribution limits (Burman et al., 2001), and fewer restrictions on withdrawals.

Account	Individual			Government		
	Initial balance	Future balance	Final payout	Initial revenue	Future revenue	PV @ r
Traditional	1	$(1+r)^T$	$(1+r)^T \cdot (1-\tau_R)$	0	$(1+r)^T \cdot \tau_R$	τ_R
Roth	$1-\tau_L$	$(1-\tau_L) \cdot (1+r)^T$	$(1-\tau_L) \cdot (1+r)^T$	τ_L	0	τ_L
Traditional - Roth	τ_L	$\tau_L \cdot (1+r)^T$	$(1+r)^T \cdot (\tau_L-\tau_R)$	$-\tau_L$	$(1+r)^T \cdot \tau_R$	$\tau_R-\tau_L$
If $\tau_R=\tau_L$			0			0

Table 2: **Benchmark cash flows under Traditional and Roth.** With flat taxes, and assuming that the tax rate on retirement income (τ_R) is the same as the tax rate on labor income (τ_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).

It is immediate to see that if $\tau_R = \tau_L$, the individual's ending wealth is the same under both account types, and therefore a consumption plan that is feasible under Traditional is also feasible under Roth, and vice versa. Moreover, we assumed that pretax labor earnings are the same under both accounts, and we guessed that the only price in the economy, the interest rate, is also the same. Given the same endowment and prices, the individual's optimal consumption plan would be the same under Traditional as under Roth.[10]

The government's cash flow differs across plans—with Roth accounts revenue τ_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 government bonds, the time-0 present value of revenue under Traditional is $(1+r)^{-T}(1+r)^T \tau_L = \tau_L$, i.e., the same as the immediate revenue under Roth. The government will therefore be indifferent in a present value sense between the accounts.

Up until now we have assumed that the interest rate is the same under the two systems, but we now show the equilibrium result that a shift from Roth to Traditional does not affect the equilibrium interest rate. Under Traditional, the account balance is τ_L larger than the

[10]We ignore for now any behavior 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.

balance under Roth. Since the account is invested in government bonds, this creates additional demand for government bonds equal to τ_L. At the same time, the government faces a revenue shortfall (relative to Roth) of τ_L. Assuming that government expenditure is exogenous, the government must issue an amount τ_L of new bonds, adding to the existing supply. Thus, the increase in desired private saving is exactly offset by the decrease in government saving, leaving desired national savings and the equilibrium interest rate unchanged.

Next, consider the case in which $\tau_R \neq \tau_L$. In this case, the indifference result does not hold. However, we can still decompose a Traditional balance into three virtual accounts as follows:

$$V_t^{Trad} = (1+r)^t \begin{bmatrix} \underbrace{(1-\tau_L)} & + & \underbrace{(\tau_L - \tau_R)} & + & \underbrace{\tau_R} \\ \text{Individual} & & \text{Government} & & \text{Implicit} \\ \text{Roth-like} & & \text{Matching} & & \text{Government} \\ \text{Account} & & \text{Account} & & \text{Account} \end{bmatrix}. \tag{1}$$

The first term is a "Roth-like" account of size $1-\tau_L$, as if the individual had contributed to a Roth account. The second term is 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 with a grant $\tau_L - \tau_R$. The size of this grant corresponds to the difference in tax rates between labor income and retirement income, which depends on separate policy choices. (Although often in practice $\tau_L - \tau_R > 0$, specific individuals may experience $\tau_L - \tau_R < 0$, resulting in a negative "match"). The last term is the "implicit government account" of size τ_R. This account represents the claim the government has on the account due to future taxes when the investor takes distributions from the account.

2.3 Extended indifference result: adding a risky asset

So far we have assumed that there is only one asset in the economy: government bonds. When individuals contribute to a Traditional account, the government issues bonds to replace the delayed tax revenue and acquires an implicit account invested in the same bonds. In reality, however, roughly two-thirds of the government's $3 trillion implicit account is invested in a mix of actively and passively managed stocks. A casual observer may argue that levered holdings of equities benefit the government, which pays a low interest rate on the debt issued to finance its implicit position while earning a high expected return on stocks. However, future government cash flows have both higher expected value *and higher risk* than in the benchmark case, because they are proportional to realized stock returns.

To analyze this, consider the base case above with $\tau_R = \tau_L$, and suppose that there are now two assets: the government bond yielding r and a risky asset (stocks) with expected return $r_s > r$. Individuals choose the share a of the portfolio that is held in stocks.

Holding tax and interest rates constant, 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. The individual would therefore be indifferent between the two. For the government, however, the timing of the cash flows would differ (as before) between Roth and Traditional.[11] Only if the government discounts the cash flows from its implicit stockholdings at rate r_s does the present value of tax revenues continue to be the same under the two systems.

Under what circumstances will r_s be the appropriate discount rate, yielding the benchmark equilibrium result? This depends on whether the government is able to make changes in its portfolio. In the absence of other changes, 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, *and* a portfolio swap in the government's implicit account that raises the demand for stocks and lowers the demand for bonds. If the government offsets the added demand for stocks under Traditional by selling stocks in other parts of its portfolio, then the benchmark indifference result above will continue to hold. However, if the government is constrained in some way from offsetting the change in the overall demand for stocks, then this may not be the case. In the absence of a "super" Ricardian equivalence (in which future taxpayers under Traditional see through the veil of government and reduce stockholdings in today's portfolios to account for the fact that future taxes will depend to a greater extent on the performance of the stock market) Roth and Traditional will not yield the same equilibrium.[12]

In the absence of an indifference result, which system would yield higher welfare? This depends on the nature of the government constraint on holding equities, and in particular whether the government is constrained to hold too large or too small a position in stocks,

[11] Under Roth, the tax revenue is τ_L at time 0. Under Traditional, the *expected* tax revenue at time T is $[a(1+r_s)^T + (1-a)(1+r)^T]\tau_L$, which is increasing in a.

[12] Infinitely lived and perfectly forward-looking individuals may realize that the government will have to change taxes in the future based on realized stock returns up to then. Under some assumptions about how the government adjusts taxes in response to stock market returns, this amounts to an implicit long position in stocks. To offset this implicit position, individuals 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.

relative to the optimum. For example, there may be political constraints on the holding of equities, because democracies are usually averse to direct government holdings of productive assets (see, e.g., Che and Qian, 1998). These constraints are not universal, however, as some governments do hold trillions of dollars of professionally managed risky investments via sovereign wealth funds.[13] Also, it is not clear that it is optimal for the government to hold positive amounts of equity, even in the absence of a political constraint, because it 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 (Auerbach, 2004). If, under a Roth system, the government is constrained from holding more stocks, then Traditional could ease that constraint and potentially improve welfare.[14] However, if the government already has too large an exposure to the stock market and is constrained from reducing it, Traditional accounts worsen the binding constraint by forcing it to hold even more stocks. The issues that arise here are similar to those in the literature on whether the government should run a sovereign wealth fund that holds stocks or whether the Social Security Trust Fund should invest in equities.[15] In the next section, we will examine an additional factor that needs to be incorporated—the fees paid on retirement savings accounts.

2.4 Fee nonneutrality in our benchmark model

We now examine the effects of introducing asset management fees. At time 0, we assume an asset management firm charges fees equal to a fixed proportion f of account size. As in the base case, we assume that the interest rate is the same under Roth and Traditional. We begin by assuming $\tau_R = \tau_L$, a necessary condition for the benchmark indifference result. We also assume that tax rates remain the same under Traditional and Roth (i.e., $\tau_L^{Roth} = \tau_L^{Trad} = \tau_L$); this is an important assumption that we discuss below. For simplicity, we assume initially that the government does not tax the corporate income of asset managers ($\tau_C = 0$), and

[13]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 end 2017 Norway's Government Pension Fund Global, the world's largest sovereign wealth fund, held roughly two-thirds of its assets in equities (`https://www.nbim.no/en/the-fund/`).

[14]A complete shift to Traditional from Roth could go too far and make the government's stock position too large, potentially reducing welfare if there were also a constraint on reducing government stock exposure through other channels.

[15]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).

that it has a nonnegative amount of debt.[16] As in subsection 2.2, we guess that initial consumption is the same and then we verify whether this assumption holds in the resulting equilibrium.

Table 3 shows our results. The individual's final retirement wealth is lower with fees than it was without, but it is still the same across Traditional and Roth. Both accounts 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 individual. Under Traditional, the initial balance is 1, and the final after-tax distribution from the account is $(1+r)^T(1-f)^T(1-\tau_R) = (1+r)^T(1-f)^T(1-\tau_L)$. Under Roth, the initial balance is $1-\tau_L$, and the final distribution from the account is $(1-\tau_L)(1+r)^T(1-f)^T$.

The right panel of Table 3 calculates the present value of tax revenue for the government. Even with fees, the government discounts cash flows at r, the interest rate it pays on outstanding debt. 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 $\tau_L^{Roth} = \tau_L^{Trad} = \tau_L$, the government has unambiguously lower present value of tax revenue under Traditional:

$$\mathbf{PV}\left(\mathbf{Tax\ Revenue}^{Trad}\right) - \mathbf{PV}\left(\mathbf{Tax\ Revenue}^{Roth}\right) = -\tau_L[1-(1-f)^T] < 0. \qquad (2)$$

This formula has an intuitive interpretation: τ_L 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.

We now examine the resulting equilibrium. For a given initial endowment and consumption, the individual's final wealth is the same under Traditional and Roth, and therefore a consumption plan that is feasible under Traditional is also feasible under Roth. Moreover, given the same wealth and prices, a consumption plan that is optimal under Traditional is also optimal under Roth. Note that since the rate of return on the account is lower, the optimal choice of savings and consumption may be different from the no-fee case. However, the individual is still indifferent between Roth and Traditional because a fraction $1-(1-f)^T$ of her final wealth is eroded by fees regardless of account type.

Generally speaking, in an equilibrium with fees, the interest rate and the individual's

[16] We make this assumption for simplicity of exposition. If the government had positive net assets, delaying revenues would not increase the debt. Instead, it would reduce the amount of assets in the government's account and increase the amount of assets in retirement accounts. To the extent that the government account were managed more efficiently than the average retirement account, our results would continue to hold.

Account	Individual			Government		
	Initial balance	Future balance	Final payout	Initial revenue	Future revenue	PV @ r
Traditional	1	$(1+r)^T \cdot (1-f)^T$	$(1+r)^T \cdot (1-f)^T \cdot (1-\tau_R)$	0	$(1+r)^T \cdot (1-f)^T \cdot \tau_R$	$(1-f)^T \cdot \tau_R$
Roth	$1-\tau_L$	$(1-\tau_L) \cdot (1+r)^T \cdot (1-f)^T$	$(1-\tau_L) \cdot (1+r)^T \cdot (1-f)^T$	τ_L	0	τ_L
Traditional - Roth	τ_L	$\tau_L \cdot (1+r)^T \cdot (1-f)^T$	$(1+r)^T \cdot (1-f)^T \cdot (\tau_L - \tau_R)$	$-\tau_L$	$(1+r)^T \cdot (1-f)^T \cdot \tau_R$	$-[\tau_L + - (1-f)^T \tau_R]$
If $\tau_R = \tau_L$			0			$-\tau_L \cdot [1 - (1-f)^T]$

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 (τ_R) is the same as the tax rate on labor income (τ_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).

final wealth need *not* remain the same under Traditional and Roth. If government expenditure follows an exogenous path, then under Traditional tax rates would have to be raised to cover the loss of revenue, reducing the individuals' wealth. Thus, a consumption plan that is feasible and optimal under Roth may not be feasible under Traditional. For the sake of exposition, we do not solve here for the equilibrium consumption plan, and instead we continue to assume that the individuals' initial consumption stays constant. Later in Section 4 we solve a general equilibrium model with an endogenously determined optimal consumption path in which nonetheless time-0 consumption remains the same.

Up to this point we have not considered corporate taxation. We now show that taxing the income of asset managers at a rate τ_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 asset management is a pure fixed-cost business. In this case, the additional assets under Traditional result in additional fee revenues without additional costs, so that every dollar of asset manager revenue turns into a dollar of pretax income and τ_C dollars of tax revenue.[17] 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 τ_C of fees:[18]

$$\begin{aligned}&\mathbf{PV}\left(\mathbf{Tax\ Revenue}^{Trad}\right) - \mathbf{PV}\left(\mathbf{Tax\ Revenue}^{Roth}\right) = \\ &= -\tau_L[1-(1-f)^T]\,(1-\tau_C) < 0.\end{aligned} \tag{3}$$

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 its expected return. If the government were to shift from Traditional to Roth, it would receive the revenue upfront and take on lower debt. By 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

[17] 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 τ_C, or escape taxation altogether. We abstract from these complexities for the sake of brevity.

[18] 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 τ_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).

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 3 of the Internet Appendix we consider two alternative hypothetical accounts: a "deductible Roth", in which individuals receive a deduction for fees, and a "nondeductible Traditional", in which the government taxes the individual based on the gross-of-fees balance.[19] 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.

As in the benchmark case, if $\tau_R \neq \tau_L$, the individual is not indifferent between Traditional and Roth, 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. However, the existence of a government matching account does not create any additional fee revenue. Asset managers receive additional fee revenue because the government leaves an amount τ_L in the account at time 0, and how this amount is ultimately split between the government and the individual is inconsequential. The present value of the individual's retirement wealth simply increases by $(1-f)^T(\tau_L-\tau_R)$, and the present value of the government's tax revenue drops by the same amount:

$$\begin{aligned}&\mathbf{PV}\left(\mathbf{Tax\ Revenue}^{Trad}\right)-\mathbf{PV}\left(\mathbf{Tax\ Revenue}^{Roth}\right)=\\&=-(1-f)^T(\tau_L-\tau_R)-\tau_L[1-(1-f)^T](1-\tau_C).\end{aligned} \tag{4}$$

Thus, from the individual's point of view, the relative preference for one or the other account is determined by the sign of $\tau_L-\tau_R$, as in the case with no fees. From the government's point of view, the cost of the matching is unaffected by fees, because the amount of matching received by the individual is equal to the drop in present value of tax revenue.

Thus, the government's *virtual* account pays *real* fees. The fees are real because we assume that the same percentage fees are charged on a larger account size. Two natural questions arise. First, how important are fees? Second, would percentage fees really remain the same if the dollar size of accounts varied? We address these questions in Sections 3 and 4 respectively.

[19]We thank Mariacristina DeNardi for bringing this possibility to our attention.

2.5 Fee nonneutrality with a risky asset

In section 2.3 above, we argued that if the government is constrained in some way from holding equities, and if it is optimal for it to hold more than it does, Traditional accounts provide an indirect way to gain exposure to equities that could increase welfare. Here we discuss how this analysis would be affected by the presence of investment fees on retirement assets. In the absence of constraints on government ownership, the government could, for example, pay an investment management firm to run a largely passive portfolio that mimics the holdings of aggregate retirement accounts, something that would likely only cost a few basis points. If this is not politically possible, then one would need to compare the benefit from added stock holdings to the loss from the higher fees paid on the implicit account under a Traditional system.

If the government were to acquire an additional dollar of stock, it would value it using a discount rate r'_s lower than the equilibrium discount rate r_s, for a benefit equal to approximately $r_s - r'_s$. Moreover, if the government acquires indirect exposure to equities via a claim on Traditional accounts—as opposed to a direct investment via a sovereign wealth fund or an index fund—the entire balance 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.[20] As the government increases its equity holdings, demand for equity increases (r_s drops) and the government's own exposure to equity increases (r'_s increases), resulting in diminishing marginal benefits.

3 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 τ_R is the effective tax rate on retirement payouts. This implicit account incurs investment fees at an annual percentage rate f, creating an implicit subsidy, i.e., a transfer from the government to asset managers that would not exist under Roth. We define f to include both explicit fees and trading costs, i.e., all costs that create a wedge between the return paid by the issuer of the securities

[20] For instance, with fees equal to 0.66% (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 required return on the marginal dollar of stocks should be lower than 7% in order for Traditional accounts to be welfare-improving.

$ billion	Total	Roth (TEE)	Traditional (EET)
Total retirement assets	26,623	1,053	23,454
Individual retirement arrangements (IRAs)	9,263	834	8,429
Defined contribution (DC) plans	7,780	219	7,561
401(k) and 403(b)	6,338	199	6,139
Other private-sector DC	540	0	540
Thrift Savings Plan (TSP)	580	9	571
457	323	10	313
Annuities	2,116	N/A	
Defined benefit (DB) plans	7,464	0	7,464

Table 4: 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.

held in a retirement account and the actual return received by the individual investor who owns that account. We assume that the government recovers a fraction τ_C of these fees via corporate taxation of the profits of asset managers and other intermediaries. Thus, a simple estimate of the annual subsidy can be calculated as:

$$\text{Annual subsidy} = S \cdot \tau_R \cdot f \cdot (1 - \tau_C) \,. \tag{5}$$

In the rest of this section, we calibrate the inputs to Equation (5). More detail on our estimates is provided in Section 4 of the Internet Appendix.

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

Table 4 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. We exclude for now DB plans ($7.5 trillion), which are also tax-deferred, because they are likely to be a more efficient investment vehicle. 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, DC plans and IRAs are both important components of overall retirement assets. A large majority of IRA assets were initially accumulated in DC plans: according to the ICI Retirement Market Statistics (Tables 9 and 10), 90% of IRA inflows from 1995 to 2015 were rollovers, primarily from DC plans. From these assets we further remove $1.0 trillion of Roth assets[21] and $0.6 trillion of assets in the federal government's Thrift Savings Plan (TSP), whose fees are negligible. This results in an estimated amount $S = \$15.4$ trillion.

3.2 Investment fees (f)

In order to estimate the magnitude of f, we aim to obtain an asset-weighted estimate of total investment fees actually incurred in U.S. self-directed retirement accounts. Asset-weighted estimates are the relevant estimates for the purpose of assessing total fees paid by the government on its implicit account. They are typically substantially lower than equal-weighted estimates, because low-fee funds have larger assets under management (Hubbard et al., 2010).

3.2.1 Explicit fees

An individual saving for retirement faces at least three types of fees: asset-level fees, account-level fees and individual-level fees. Asset-level fees are charged based on what financial products the account money is invested in, and include both ongoing fees like mutual fund expense ratios and one-time fees such as sales loads. These fees cover the operating costs and profits of asset managers (e.g., mutual fund sponsors, insurance companies, and issuers of structured notes) and distribution channels (e.g., mutual fund brokers and securities brokers). Asset-level fees exist both in IRAs and DC plans, although distribution fees in DC plans are usually minimal. Account-level fees include account maintenance fees in IRAs and administrative costs in DC plans. Individual-level fees are charged primarily for financial

[21] The ICI Retirement Market Statistics report that $834 billion or 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 and in 2012 for TSP. We are not aware of Roth options for other private-sector DC plans.

advice on one or multiple accounts. These fees, often referred to as advisory fees or "wrap" fees, cover the costs of providing expertise in asset allocation, estate and tax planning, and other services.

In practice, fees are not always easy to categorize. Individual-level fees are difficult to separate from account-level fees because some account providers include a basic level of advice in account fees. Especially in IRAs, account-level fees are difficult to separate from asset-level fees because a substantial fraction of asset-level distribution fees (sales loads and so-called 12b-1 fees) are rebated to the account provider who is either part of the distribution network or is the broker itself.

For DC plans, 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, which represent 74% of non-TSP DC plan assets. Deloitte (Rosshirt et al., 2014) estimates the "all-in fee" at 58 bps, while BrightScope (2018) estimates "total plan costs" at 37 bps. The Deloitte estimate is survey-based, providing a less precise but more representative estimate. The survey excludes plans with less than $1 million in assets and oversamples large plans, claiming representation of roughly 97% of the assets (36% of the plans) within the universe of plans filing Form 5500 with the Department of Labor. The BrightScope estimate is based on filings by audited plans, which generally means plans with 100 or more participants. As a consequence, the BrightScope study excludes about $1 trillion or 27% of total assets held in the smallest, and likely most expensive, plans. We use 50 bps as our estimate of total account and asset-management costs based on a rough average of these two estimates.

For IRAs, we estimate asset-level fees alone to be at least 38 bps. Although one-third of IRA assets is invested in individual securities without explicit fees other than trade commissions, more than half is held in mutual funds with asset-weighted average fees of 56 bps, and the remainder is invested in other products like real estate funds, commodity funds, etc. whose fees are estimated at 65 bps.[22] The fees of mutual funds held in IRAs (56 bps) are

[22]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 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 the difficulty of understanding the price of a financial product does not necessarily translate to a high price, a growing

substantially higher than those of collective investment products held in DC plans (37 bps), reflecting higher distribution fees due to the dispersed nature of individual accounts. We expect IRAs to be more expensive also with respect to account-level fees and individual-level advice fees. However, since to the best of our knowledge no comprehensive estimates of these fees are available, we conservatively assume total fees of 50 bps, the same as for DC plans.

3.2.2 Trading costs

The above "all-in" estimates do not include trading costs, which nonetheless affect net returns. Individual investors who directly trade individual securities incur explicit trading commissions and implicit trading costs in the form of bid-ask spreads, defined as the difference between the buy price and the sell price. Mutual funds trading on behalf of individual shareholders also incur the same costs. Trading commissions are not included in the expense ratio and they constitute implicit costs for mutual fund shareholders. Another important implicit cost for mutual funds is 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.[23]

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

$$\text{Annual trading costs} = \text{Trading costs per unit of volume} \times 2 \times \text{Annual turnover}.$$

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

literature on shrouded prices (Gabaix and Laibson, 2006; Carlin, 2009; Henderson and Pearson, 2011) suggests that this is typically the case.

[23]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). These arrangements reduce the funds' explicit expenses to the detriment of execution quality, and thus muddy the distinction between explicit fees and trading costs.

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 $2 \times$ 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). Assuming this lower figure applies to both DC plans and IRAs, we obtain annual trading costs for equity mutual funds of $26 \times 2 \times 32\% = 17$ bps. This estimate is lower than typical pronouncements by industry insiders,[24] and it is driven by the low turnover of index funds.

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. This estimate is consistent with other recent works (Choi and Huh, 2017; Goldstein and Hotchkiss, 2018).

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.[25] Thus, our estimate of

[24]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.

[25]When selling a security and purchasing another, a fund trades twice; when receiving 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

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.[26] These numbers reflect a large asset allocation to equity.

We combine our estimates of explicit fees and trading costs using the total amount of assets in DC plans and IRAs from Table 4 to obtain asset-weighted fees of 50 bps and trading costs of 16 bps, resulting in a conservative, asset-weighted estimate of "all-in" average fees of 66 bps, or $f = 0.66\%$.

3.3 Tax rates (τ_R and τ_C)

The effective tax rate on retirement payouts, τ_R, is challenging to estimate because the effect of progressive personal income taxes is not negligible. We aim to estimate the "average marginal" tax rate, i.e., the tax rate on the average dollar paid out of a tax-deferred retirement account. This 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.

We estimate τ_R in two ways. First, an average marginal tax rate range of 20%–30% is deduced by reverse-engineering present-value tax expenditure estimates published by the federal government (Office of Management and Budget, 2014) or its employees (Lurie and Ramnath, 2011). Second, using data on retirement wealth reported in the Survey of Consumer Finances (SCF), we independently estimate an average marginal tax rate of 26%.[27]

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 ($2 \times 4/5 + 1 \times 1/5 = 1.8$). 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%.

[26]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).

[27]For each SCF observation we calculate W^T, taxable wealth, and W^{Trad}, tax-deferred retirement wealth. We assume baseline retirement income to be equal to taxable wealth times a constant nominal rate of

This estimate could be slightly overstated because it is obtained assuming individuals do not time their withdrawals to minimize tax liabilities. For this reason, as a conservative estimate we use 20%, the lowest justifiable number. The size of the implicit government account is therefore $S \cdot \tau_R$ =\$3 trillion.

For τ_C, the corporate tax rate, we simply use the top statutory corporate tax rate of 21%.

3.4 Calibration

Using the calibrated inputs in Eq. (5), we obtain our estimate:

$$\begin{aligned} \text{Annual subsidy} = S \cdot \tau_R \cdot f \cdot (1 - \tau_C) = \\ = \$15.4 \times 20\% \times 0.66\% \times (1 - 21\%) = \$16.1 \text{ billion.} \end{aligned} \tag{6}$$

As discussed above in Section 3.1, 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, even defined-benefit plan assets can 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 66 bps),[28] the estimated subsidy rises to \$21.6 billion.

In Table 5, we carry out the same back-of-the envelope calculation for the seven countries with the largest dollar amounts of tax-deferred assets. For each country, we obtain information on all existing types of tax-advantaged retirement plans and their tax treatment

return $r = 3\%$ ($Y_{Baseline} = W^T r$). In addition to this baseline income, we assume that the individual uses W^{Trad} to purchase a fixed annuity lasting $T = 20$ years, discounted using the same rate $r = 3\%$ ($Y_{Combined} = W^T r + W^{Trad}/\{[1-(1+r)^{-T}]/r\}$). We then calculate total dollar tax on the baseline income ($T_{Baseline}$) and on the combined total income ($T_{Combined}$) applying the U.S. tax brackets that were in force at the time of the survey, and calculate the average tax rate as $(T_{Combined} - T_{Baseline})/(Y_{Combined} - Y_{Baseline})$. Using both sampling weights and value weights, we arrive at an average tax rate of 25.8%.

[28] This rough estimate is implied by Table II of Dyck and Pomorski (2011). We define total plan fees as "Overall 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.

from the Organization for Economic Cooperation and Development (OECD, 2015a,b). For Traditional plans and other plans with a tax-deferral feature, we obtain an estimate of the total assets in each type of plan from various sources. We then estimate the size of the implicit government account by multiplying total tax-deferred assets by the average tax rate on retirement payouts (τ_R). We obtain information on average retirement income from each country's statistical office, and information on basic deductions and tax brackets from each country's tax authority. With this information, we estimate a lower bound to τ_R as the average tax rate faced by a person earning the average retirement income with no other income. Fees are estimated as the asset-weighted average of money market, equity and fixed-income mutual fund fees based on overall (not retirement-only) asset allocation in each country. Information on fees is collected from Morningstar (Alpert et al., 2013) and other sources. As before, τ_C, the corporate tax rate, is simply the top statutory corporate tax rate as reported by each country's tax authority.[29] We then aggregate the different types of plans up to the country level and convert to U.S. dollars using current exchange rates.

For consistency with our U.S. estimates, we exclude defined benefit (DB) pension plans from the calculation. With or without DB plans, the 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.

The average tax rate on retirement payouts (τ_R) is another important factor. 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.

Finally, the level of fees is obviously an important determinant of the size of the subsidy. Although our non-U.S. fee estimates are not as precise as the U.S. estimates, greater accuracy is unlikely to change rankings. For instance, within the seven countries in Table 5, Canada has the second-largest subsidy in dollar terms ($5.2 billion) and the largest as a fraction of

[29] Unlike other countries, Australia's Superannuation Guarantee is taxed under a TTE scheme: contributions are taxed at favored flat rates (usually 15%) and returns are also taxed at favored rates (15% for interest and dividends, 10% for capital gains), while usually payouts are tax-exempt. Thus, compared to a pure Roth scheme in which all taxes are levied upfront, the Australian system entails some degree of tax deferral. In this case, we define τ^{Roth} as the tax rate that under a pure Roth scheme would raise the same present value of tax revenue as the actual revenue raised under the current scheme. Then, since the Australian government does levy a 15% upfront tax on contributions, the size of the government's implicit account is simply the remainder, $\tau^{Roth} - 15\%$.

	Retirement Assets		Govt. Acct. Size				Subsidy	
Country	$b	% Deferred	τ_R	$b	Fees	τ_C	$b	% GDP
United States	16,464	94%	20%	3,084	0.66%	21%	16.1	0.09%
Canada	2,082	95%	15%	295	2.06%	15%	5.2	0.34%
United Kingdom	950	32%	20%	41	1.45%	20%	0.7	0.02%
Netherlands	108	100%	39%	41	1.41%	25%	0.4	0.06%
Switzerland	945	100%	4.0%	38	1.29%	18%	0.4	0.06%
Australia	1,797	55%	3.4%	34	1.10%	30%	0.3	0.02%
Japan	112	100%	2.6%	3	1.47%	30%	0.0	0.00%

Table 5: **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, τ_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. τ_C, the corporate tax rate, is simply the top statutory tax rate. Sources: see text.

GDP (0.34%). This is in part driven by the surprisingly large fees charged by Canadian funds (2.06%). We have no evidence that this figure is exaggerated, as two independent sources report numbers in the neighborhood of 2% (Alpert et al., 2013; Investor Economics, 2012). However, Canada would retain its rank—both in dollars, and in percent—even using a fee estimate of 1%.

3.5 The value received in exchange for fees

Investors do receive value in exchange for fees. If the government obtains any benefits in exchange for the fees paid on its virtual account, these benefits should be subtracted from the fees to arrive at a net cost. However, we argue the government's implicit account does not benefit from most services provided by the asset management industry, and our estimate should not be adjusted.

3.5.1 Services

One important service that individual investors receive is basic portfolio management, i.e., asset managers build and manage large portfolios of securities. However, basic portfolio management services are an inexpensive commodity. The U.S. federal government's Thrift

Savings Plan (TSP) and at least one private-sector fund family offer passive index funds charging no expense ratio.[30]

Other services obtained in exchange for fees have no value to the government. First, as discussed above, a substantial fraction of the cost of investing in mutual funds and other structured products consists of distribution fees.[31] The government, however, does not benefit from distribution services.

Second, 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.[32]

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.

[30]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. The private-sector family offers two equity funds with zero expense ratio, seemingly as loss leaders.

[31]For 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).

[32]To 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.

3.5.2 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 indifferent between active and passive funds in equilibrium. On its face, this would imply that the overall cost of active management should not be measured but rather assumed to be equal to the cost of passive management.

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 *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, although many are based on fund-level factor regressions and thus essentially equal-weighted. Two recent studies provide figures comparable to our asset-weighted estimate. 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).[33] 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.

[33] According to Petajisto (2009), the S&P 500 index itself has implicit trading costs of 20-28 bps.

Upon a switch from Roth to Traditional, the additional government assets would create little additional value.

To conclude, it is important to point out that we provide 66 bps as our best estimate of total investment fees—a contribution in and of itself. However, our conclusion does not depend qualitatively on the exact number. Even if the true number were half as large, Traditional accounts would still entail a large subsidy to the asset management industry.

4 A simple general equilibrium model of retirement savings and asset management

In the previous sections we have shown that, under the assumption of constant percentage fees f, Traditional accounts result in greater assets under management and a subsidy to asset managers. In this section, we examine whether these results continue to hold in general equilibrium and if so, how the magnitudes are affected. We also examine the implications of the choice between Roth and Traditional for employment in the asset management sector and for social welfare.

4.1 Choosing a general equilibrium model

To study these questions, we write down and solve a simple, two-sector model that includes an asset management industry and is compatible with the basic empirical evidence about cost structure, market structure, and competition in the asset management industry. The model yields predictions on several equilibrium quantities, including total assets under management S (i.e., the quantity of services provided), the level of asset management fees f (i.e., the price), and total employment in the asset management industry L (the resources used). For simplicity, we assume free entry, which results in zero equilibrium profits under both Roth and Traditional. Below, we describe the model and the resulting equilibrium. A full derivation of the equilibrium prices and quantities is provided in Section 1 of the Internet Appendix.

4.1.1 Cost structure

In analyzing the equilibrium effects of a switch from Roth to Traditional, and the associated increase in assets under management, perhaps the most important issue is the existence (and sign) of economies of scale. 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 seems uncontroversial: most empirical studies agree that larger funds and larger plans have lower (percent) administrative costs (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 Vanguard 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 because of cost economies and access to a better investment opportunity set. On the other hand, for 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 with a fixed component and a variable component,

$$\text{Cost}_j = \gamma_1 + \gamma_2 \cdot q_j S, \tag{7}$$

where $q_j S$ represents the AUM of fund j (aggregate AUM $S \times$ market share q_j). We choose, conservatively, to rule out the possibility of diseconomies of scale by requiring $\gamma_1 \geq 0$. This specification nests a few interesting cases.

Consider first the special case in which costs are proportional to AUM ($\gamma_1 = 0$, no economies or diseconomies of scale). This case is similar to the fixed-supply case we examined in Section 2, and under weak assumptions the results we obtained there would hold exactly, i.e., an increase in AUM does not affect the equilibrium level of fees, f.

Alternatively, consider the extreme case in which all asset management costs are fixed costs ($\gamma_2 = 0$, positive economies of scale). In this case, for a given fund, an increase in AUM entails no new costs. We believe this case is interesting and, as a stylized representation of economies of scale on the cost side, intuitively plausible. Although, in reality, asset manager costs likely have both fixed and variable components, we focus most of our discussion on this special case.

4.1.2 Market structure and competition

If all costs were fixed, a model in which firms compete on price alone would result in a natural monopoly. However, there is abundant evidence that individual demand for financial services is imperfectly sensitive to the price. Hortacsu and Syverson (2004) show evidence of substantial price dispersion, noting 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 (the most recent year of complete CRSP data) there existed 79 share classes (46 distinct funds) with a fee interquartile range of 102 bps.[34] 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" (Cremers et al., 2016).

All proposed explanations for the survival of dominated funds point to information frictions or outright inertia: 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; Gârleanu and Pedersen, 2018), persistent reputation (Hubbard et al., 2010), investors' unwillingness to sever relationships with brokers or trusted advisors (Bergstresser et al., 2009; Gennaioli et al., 2015), or even irrationality (Elton et al., 2004).

Consistent with the empirical evidence, we specify a model with limited price sensitivity in which funds face a downward-sloping demand function, i.e. if they raise their fees, their demand falls, but it does not fall to zero. However, with respect to the driving force behind limited price sensitivity, we take a view that is more conservative given what we aim to show. Instead of a model with information frictions,[35] we use a model of monopolistic competition

[34] 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. Christoffersen and Musto (2002) show evidence of fee dispersion for money market mutual funds.

[35] For instance, we could have based our analysis on a model like the one in Anderson and Renault (1999). Their model, like almost every model of firm competition based on search frictions, results in high prices and excessive entry. Similarly it would have been possible to write a model with captive demand and shrouded fees. In such a model, back-loaded taxation would still cause an increase in the resources devoted to asset management, and welfare consequences would be undoubtedly more severe.

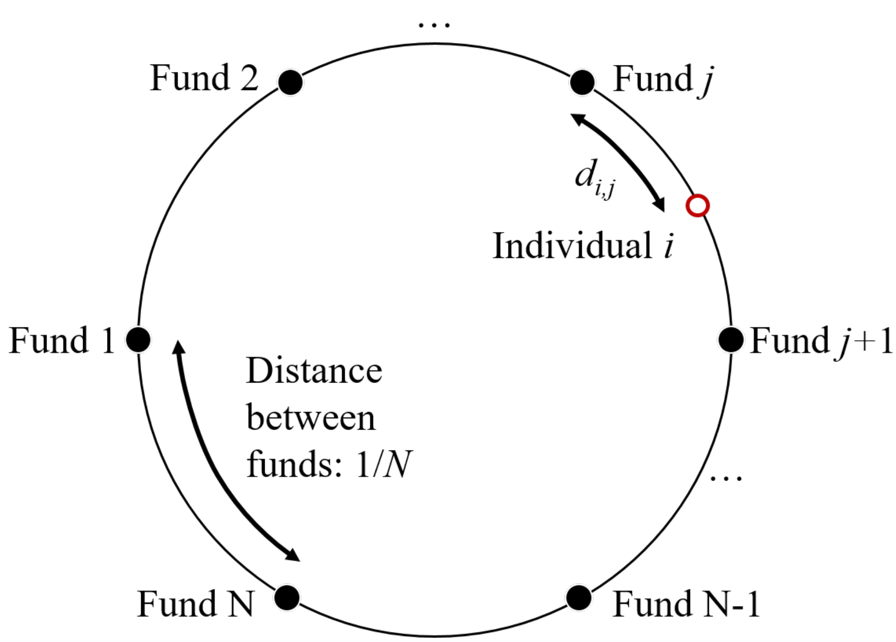

Figure 3: **Spatial competition**. N funds are uniformly spaced along a circle of circumference 1, as in Salop (1979). Individuals are distributed uniformly along the circle, and prefer funds located at a closer distance.

with differentiated products similar to Salop (1979).

4.1.3 Entry

Because our model allows for product differentiation, the assumption of a dominant fixed cost component does not result in a monopoly. Moreover, empirical evidence and casual observation suggest low barriers to entry in the mutual fund industry (Hubbard et al., 2010; Baumol et al., 1990). In 2017 alone, 464 new mutual funds, 241 new exchange-traded funds, and 41 new fund sponsors entered the industry (ICI, 2018). A similar situation is reflected in the non-mutual fund segments of the asset management industry. For instance, in a 2016 survey sent by a leading industry publication to 1,070 known third-party retirement plan administrators, the majority of respondents were established in the past 25 years. To reflect 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.

4.2 Demand: individuals

The model is a version of Salop's (1979) circular city embedded 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 they are old ($t = 1$), individuals retire and deplete all the savings, 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. For a given wage, individuals are indifferent between sectors.

Saving yields an exogenous return equal to r. Individual savings, S, must be allocated to one of N firms producing asset management services ("funds") who invest in a mix of government bonds and financial assets.[36] Funds are also situated on the circle at locations ι_1, ι_2, ...ι_N.

In our model, each individual's utility depends negatively on the "distance" from their chosen fund, and every additional fund entering the market increases utility for the average individual because it reduces the average distance. A low distance can be thought of as literally low physical distance from the nearest branch, but also ease of finding (e.g., the fund is recommended by the account administrator), 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

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

Individuals derive diminishing utility from current consumption (C_0) and from future consumption (C_1) discounted by a factor δ. In addition, individuals derive diminishing disutility from the distance between their own location i and the location of the nearest fund j, $d_{i,j} \equiv |i - \iota_j|$. On the circle, the distance between two points is defined as the shortest possible distance—for instance, the distance between 0.1 and 0.9 is 0.2, not 0.8.

We choose this functional form because it is economically sensible and tractable, yielding

[36] In our simple model we ignore the existence of multiple layers of financial intermediation (recordkeepers, asset managers, fund families, subadvisors, securities brokers, etc.).

easy-to-interpret expressions for the quantities of interest.[37] One potential problem with our specification 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 because of them. However, if one were to find this setup philosophically problematic, the same results could be obtained by assuming that no individuals live in an ε-neighborhood centered around each fund and examining the limit as $\varepsilon \to 0$.

4.3 Supply: Funds

The unit of production of asset management services is the "fund", a generalist firm that produces asset management services. We ignore for simplicity the existence of distinct layers of financial intermediation. We believe the basic intuition of our model would continue to hold in a more complicated model. 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.[38]

As anticipated in Section 4.1.1, we assume that, for every fund $j \in \{1, 2, \ldots N\}$, operating costs have a fixed component (each fund needs a fixed amount of labor φ to be able to operate, regardless of its assets under management) and a variable component (a fraction c of assets

[37] To the best of our knowledge, 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. Then, for distance disutility, we choose a logarithmic specification for tractability. A linear specification would yield the same economic intuition but with much more cumbersome expressions for the quantities of interest.

[38] These assumptions lighten the exposition considerably and are consistent with the preexisting 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. The resulting positive equilibrium profits must then be assigned to individuals.

under management net of fixed costs), resulting in the cost function[39]

$$\text{Cost}_j = c\left(q_j S - \varphi\omega\right) + \varphi\omega. \tag{9}$$

Since aggregate labor supply is 1, the amounts of labor going to the asset management sector and the consumption goods sector are $L \equiv \varphi N$ and $1-L$ respectively. The production technology for the consumption goods is linear: aggregate output of consumption goods is $\mathcal{F}(1-\varphi N) = \omega(1-\varphi N)$, where ω is the marginal (and average) product of labor. Since labor markets are frictionless, ω also equals the wage of each worker.

4.4 The planner's problem

After describing our assumptions about demand and supply, we solve the model. In this subsection, we characterize the social optimum in the model by assuming a planner directly chooses individual consumption $\{C_{0,i}\}$ and $\{C_{1,i}\}$ and the number of funds N to maximize aggregate utility. In the following subsections, we compute the market equilibrium, and then compare the two. The planner maximizes

$$U = \max_{\{C_{0,i}\},\{C_{1,i}\},N} \int_0^1 \ln C_{0,i} + \delta \ln C_{1,i} - \gamma \ln d_{i,j}\, di, \tag{10}$$

subject to the resource constraint

$$C_1 = \left[\omega\left(1-\varphi N\right) - C_0 - G\right]\left(1+r\right)\left(1-c\right), \tag{11}$$

where $C_t \equiv \int_0^1 C_{t,i} di$ is aggregate consumption at time $t \in \{0,1\}$, φN is the amount of labor absorbed by the N funds, G is an exogenously given amount of public expenditure, and $(1+r)(1-c)$ is the exogenously given gross return on a storage technology, net of variable costs.

Note that 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 the utility function is concave and separable in its arguments. Then, using the assumption that the N funds are located equidistantly along

[39]This specification is equivalent to Equation (7) with $\gamma_1 = \varphi\omega(1-c)$ and $\gamma_2 = c$. The reason for subtracting fixed costs from AUM before computing variable costs is that an amount $\varphi\omega$ is taken out at time 0 to pay fixed costs (e.g., staff salaries) and never invested, and variable costs (e.g., trading costs) are proportional to the assets actually being managed. Other specifications would have produced essentially identical results, but this specification yields a definition of S that is consistent between the planner problem and the market equilibrium and lends itself to the most straightforward interpretation.

the circle, the planner's objective function simplifies to

$$U = \max_{C_0, N} \ln C_0 + \delta \ln C_1 + \gamma \ln N - \gamma (1 + \ln 2) \tag{12}$$

(see the Internet Appendix for a derivation). The first-order conditions for C_0 and N are, respectively,

$$C_0 = \frac{1}{\delta (1+r)(1-c)} C_1 = \frac{1}{1+\delta} [\omega (1 - \varphi N) - G] \tag{13}$$

and

$$\frac{1+\delta}{\omega (1 - \varphi N) - G} \varphi \omega = \frac{\gamma}{N}. \tag{14}$$

The Euler equation (13) says that the resources devoted to C_0 and C_1 should be fixed proportions of disposable income, defined as consumption good output $[\omega (1 - \varphi N)]$ minus government expenditure (G). Equation (14) says that the marginal utility of consumption times the cost of operating a fund (the left-hand side) should be equal to the marginal benefit (the right-hand side).

Simplifying the first-order conditions yields the following optimal quantities:

$$C_0^* = \frac{1}{1+\delta+\gamma} (\omega - G), \tag{15}$$

$$C_1^* = \frac{\delta}{1+\delta+\gamma} (\omega - G)(1+r)(1-c), \tag{16}$$

$$N^* = \frac{\gamma}{1+\delta+\gamma} (\omega - G) \frac{1}{\varphi \omega}. \tag{17}$$

Thus, at time 0, the planner optimally allocates fractions $1/(1+\delta+\gamma)$, $\delta/(1+\delta+\gamma)$, and $\gamma/(1+\delta+\gamma)$ of available resources $(\omega - G)$ to current consumption, next-period consumption, and asset management, respectively. For the purpose of comparing the planner solution with the market equilibrium, it is also useful to define aggregate savings S^* as the amount of available resources minus current consumption, or equivalently, the resources devoted to next-period consumption plus asset management:

$$S^* = \omega - G - C_0^* = \frac{\delta+\gamma}{1+\delta+\gamma} (\omega - G). \tag{18}$$

We now turn our attention to market outcomes.

4.5 The market economy

In the market equilibrium, individuals choose consumption (C_0 and C_1) and saving (S) and allocate their savings to a fund (j) to maximize utility. N funds choose to enter the market, pay a competitive wage (ω) to attract labor, and set their fees (f) to maximize profits taking competitors' choices as given. The government spends an exogenously given amount G. To balance the budget, the government taxes individuals' labor income at a rate τ_L, retirement income at a rate τ_R, and grants the individual a deduction for retirement savings at a rate τ_S whose value depends on the specific incentive scheme (Roth or Traditional).

4.5.1 The individual's problem

Individuals earn labor income ω regardless of whether they work for a fund or produce consumption goods. At time 0, the individual consumes C_0 and saves and invests an amount S to finance retirement consumption C_1. We assume that the individual must allocate all savings to one fund, j, which charges a proportional fee $f_j S$. The individual's savings are invested in either government bonds or in a storage technology, both yielding an exogenously given return r.[40] Thus, the individual chooses C_0 and j to maximize utility:

$$U_i = \max_{C_0, j} \ln C_0 + \delta \ln C_1 - \gamma \ln d_{i,j}, \tag{19}$$

subject to the following budget constraints:

$$C_0 + S = \omega (1 - \tau_L) + \tau_S S. \tag{20}$$

$$C_1 = S (1 - f_j)(1 + r)(1 - \tau_R). \tag{21}$$

The individual's first-order condition with respect to consumption is

$$\frac{1}{C_0} = \frac{\delta}{\omega (1 - \tau_L) - C_0}. \tag{22}$$

[40]Note that the storage technology is the only way to finance time-1 consumption, and therefore it is always in positive demand ($S^* > B$, where B is the stock of government bonds). In the Internet Appendix we formally allow the return on bonds to be determined endogenously, but since an unlimited amount of assets can be invested in the storage technology, we obtain the same result. We omit this step for brevity.

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:

$$j^* = \arg \max_{j \in \{1,2,...N\}} \ln C_0 + \delta \ln C_1 - \gamma \ln d_{i,j}. \tag{23}$$

4.5.2 Competition between funds and the fund's problem

We assume that funds $j \in \{1, 2, \ldots N\}$ choose to enter the market until profits are zero:

$$\pi_j = 0 \qquad \forall j \in \{1, 2, \ldots N\}. \tag{24}$$

Funds are assumed to locate at equal distances along the circle. Thus, once the entry decision is made, fund j's only remaining decision is f_j, the percentage fee it charges. f_j is chosen to maximize profit π_j, defined as fee revenue, $(f_j q_j S)$ minus variable costs, $(c(q_j S - \varphi\omega))$ minus the fixed labor cost $\varphi\omega$. The fund takes into account that these choices will affect its market share q_j, but it takes competitors' choices as given. Thus, the fund's problem is (rearranging in a more convenient form):

$$\max_{f_j} \pi_j = (f_j - c)\, q_j S - \varphi\omega\,(1 - c). \tag{25}$$

Fund profits are taxed at the same rate as other income. We do not keep track of fund profits, however, as they are zero because of free entry.

4.5.3 Government and tax policy

At time 0, the government spends an exogenously given amount G, taxes income at a rate τ_L, and grants individuals a deduction for savings at a rate τ_S. In order to satisfy its time-0 budget constraint, the government borrows an amount B at the market interest rate r:

$$G = \omega\tau_L - S\tau_S + B. \tag{26}$$

At time 1, the government taxes retirement income at a rate τ_R to satisfy the time-1 budget constraint:

$$B\,(1 + r) = S\,(1 - f)\,(1 + r)\,\tau_R, \tag{27}$$

where f is the equilibrium level of fees.

We consider two policy options corresponding to Roth (TEE) and Traditional (EET) incentive schemes. Under Roth, the government grants no deduction ($\tau_S^{Roth} = 0$) and does not tax retirement income ($\tau_R^{Roth} = 0$). Thus, it has no revenue at time 1, and therefore it cannot borrow ($B^{Roth} = 0$). This leaves no choice but to balance the budget at time 0 by setting the tax rate equal to the ratio between government expenditure and output:

$$\tau_L^{Roth} = G/\omega. \tag{28}$$

Under Traditional, the government grants a deduction for savings at the same rate as labor income ($\tau_S = \tau_L$). With three policy variables (the labor income tax rate τ_L, the retirement income tax τ_R, and the amount of borrowing B) and two budget constraints (at times 0 and 1), the government has one degree of freedom. To eliminate this source of indeterminacy and to facilitate comparison between the two account types, we initially assume that the tax rate on labor is the same under Traditional as it is under Roth, so that

$$\tau_L^{Trad} = \tau_L^{Roth} = G/\omega. \tag{29}$$

Then, the government's time-1 budget constraint requires

$$\tau_R^{Trad} = \frac{1}{1-f} {}^{G}\!/_{\omega} > {}^{G}\!/_{\omega} = \tau_L^{Trad}. \tag{30}$$

In the Internet Appendix we examine the possibility that the government chooses τ_L^{Trad} optimally.

4.6 Consumption, entry, and fees in the market equilibrium

In Section 1 of the Internet Appendix we prove that the values in the market equilibrium of the key variables of interest (C_0 and C_1, aggregate consumption; S, aggregate assets under management; f, the equilibrium level of fees; N, the number of funds; and $L \equiv \varphi N$, total employment in the asset management industry) are:

$$C_0 = \frac{1}{1+\delta}\omega\left(1-\tau_L\right), \tag{31}$$

$$C_1 = \frac{\delta^2}{\left(1+\delta\right)\left(2\gamma+\delta\right)}\omega\left(1-\tau_R\right)\frac{1-\tau_L}{1-\tau_S}\left(1-c\right)\left(1+r\right), \tag{32}$$

$$S = \frac{\delta}{1+\delta}\omega\frac{1-\tau_L}{1-\tau_S}, \tag{33}$$

$$f = \frac{2\gamma + c\delta}{2\gamma + \delta}, \text{ and} \tag{34}$$

$$N = \frac{1}{\varphi} \cdot \frac{2\gamma\delta}{(1+\delta)(2\gamma+\delta)} \cdot \frac{1-\tau_L}{1-\tau_S}. \tag{35}$$

If $\varphi = 0$ (i.e. if entry is costless), the zero-profit condition simplifies to $f = c$, N is infinite, and time-1 consumption is

$$C_1 = \frac{\delta}{1+\delta}(\omega - G)\frac{1-\tau_R}{1-\tau_S}(1-c)(1+r). \tag{36}$$

We compare Roth and Traditional using their respective definitions ($\tau_L^{Roth} = G/\omega$, $\tau_S^{Roth} = 0$, $\tau_R^{Roth} = 0$ and $\tau_L^{Trad} = G/\omega$, $\tau_S^{Trad} = \tau_L^{Trad}$, $\tau_R^{Trad} > G/\omega$) and equations (31)–(35). The comparison yields the same results as in our fixed-supply model of Section 2: Roth and Traditional have the same time-0 consumption and the same level of percent fees, and Traditional has higher AUM and higher total dollar fees. In addition, we obtain some results specific to our general equilibrium setting: Traditional has higher number of funds, higher employment in the asset management industry, and lower retirement consumption. The existence of a subsidy for asset management results in greater resources devoted to asset management under Traditional. Because the government's expenditure is exogenous and the government's budget constraint is binding, these resources must be provided at the expense of individuals' consumption. Thus, unlike in the fixed-supply model of Section 2, the individual is no longer indifferent between Roth and Traditional.[41]

Importantly, these results do not depend on specific assumptions about the cost function, as long as asset management has positive costs, i.e., they hold both in the case with pure variable costs ($\varphi = 0$, $c > 0$) and with pure fixed costs ($\varphi > 0$, $c = 0$).

4.7 Welfare in the model

We now compare social welfare in the model under Traditional and under Roth. We first compare the market equilibrium under Roth ("Roth equilibrium") with the planner solution, and then compare the market equilibrium under Traditional ("Traditional equilibrium") with the Roth equilibrium. The socially optimal quantities under the planner solution are denoted

[41] This statement refers to the comparison we have operated so far between a Roth-only world and a Traditional-only world. If both accounts are allowed to coexist in the model, the individual may or may not be indifferent depending on the choice of tax rates. (Note that this would require setting the same tax rates for both accounts). Setting $\tau_L = G/\omega$ would require $\tau_R > \tau_L$, causing individuals to internalize the cost of Traditional and prefer Roth.

by an asterisk (*), whereas the market quantities are denoted by their respective labels (*Roth* and *Trad*).

First, we observe that compared to the optimal level of time-0 consumption (equation 15), in a Roth equilibrium individuals consume too much (equation 31):

$$C_0^{Roth} = \frac{1}{1+\delta}(\omega - G) > C_0^* = \frac{1}{1+\delta+\gamma}(\omega - G), \tag{37}$$

and thus save too little:

$$S^{Roth} = \frac{\delta}{1+\delta}(\omega - G) < S^* = \frac{\delta+\gamma}{1+\delta+\gamma}(\omega - G). \tag{38}$$

Since $C_0^{Trad} = C_0^{Roth}$ and $S^{Trad}\left(1 - \tau_R^{Trad}\right) < S^{Roth}$, the phenomenon is even more acute under Traditional. Undersaving in the model is a consequence of the fact that fees are charged as a percent of total assets, as opposed to a lump sum: individuals have an implicit disincentive to save, because if they save more they pay higher dollar fees.

Next, we ask whether a Roth equilibrium results in excess resources devoted to asset management. The answer depends on the model parameters. Namely, the equilibrium number of funds (equation 35) under Roth is higher than the optimal number (equation 17) if:

$$N^{Roth} > N^* \iff \delta + \delta^2 > 2\gamma. \tag{39}$$

The intuition is that the size of the asset management industry departs from the optimum because of two effects. First, undersaving due to percent fees results in too few assets and therefore too few funds. Second, spatial competition à la Salop results in too many funds ("excess product differentiation"). If (39) is verified, the second effect prevails, and a Roth equilibrium has too many funds.[42] In practice, the empirical evidence for the continued existence of dominated funds and price dispersion suggests that this condition does in fact hold in reality.

[42] To the extent that it makes sense to calibrate our simple model, a Roth equilibrium has too many funds for reasonable parameter values. γ (or, to be precise, $\gamma/(1+\delta+\gamma)$) represents the fraction of resources devoted to asset management. Greenwood and Scharfstein (2013) show that the securities industry (including investment banking and asset management) contributes about 2% of gross domestic product—a large amount, but a small fraction of the economy. δ represents the discount factor applied to retirement consumption. Assuming a time horizon of 30 years and an annual discount rate of 5% implies $\delta = .215$ and $\gamma = 0.024$. Under this calibration, annualized equilibrium fees are 0.56%. In the Internet Appendix we show that Roth is guaranteed to have a higher number of funds than the optimum if the first effect is shut down (i.e., if fees are required to be charged as a lump sum F instead of a proportional amount fS).

If a Roth equilibrium has too many funds, too many resources expended on time-0 consumption and on asset management must result in too few resources available for time-1 consumption. Then, a switch from Roth to Traditional results in lower aggregate utility in the model because Traditional has even more funds ($N^{Trad} > N^{Roth} > N^*$) and even less time-1 consumption than Roth ($C_1^{Trad} < C_1^{Roth} < C_1^*$).

If, on the other hand, a Roth equilibrium entails undersaving and underprovision of asset management services, it may appear beneficial to increase aggregate saving and the number of funds. However, even in this case a switch from Roth to Traditional has negative consequences on aggregate welfare in the model. In the Internet Appendix we show that aggregate utility under Roth is always higher than under Traditional, regardless of condition (39). The reason is that a switch from Roth to Traditional does not affect C_0 (which is too high), so that the increase in the number of funds happens at the expense of C_1 (which was already too low). In other words, Traditional does not create a bigger saving subsidy than Roth in our model; it merely exacerbates the individuals' price insensitivity.

The model's main insight is that Traditional accounts result in more assets, and that this increase in assets causes an increase in resources devoted to asset management. This is clearly true if the cost of producing asset management services at the firm level is increasing in the amount of assets under management ($c > 0$). Less obviously, as our model shows, it is also sufficient that the firm-level cost function has a fixed component ($\varphi > 0$), as more assets make more firms viable in equilibrium. Because in practice the cost function likely has both variable and fixed components, in general equilibrium the subsidy would almost certainly result in a transfer to the asset management industry, and this transfer is likely to result in excess real resources devoted to asset management. Moreover, as discussed above, it is likely that resources devoted to the asset management industry are already too high under a Roth-only system, and therefore the excess resources diverted to asset management under a Traditional system would be damaging for social welfare as defined in our model.

Our model also contributes to a recent literature on the optimal size of finance. Our study shows one plausible mechanism (price insensitivity driven by product differentiation) by which the financial industry attracts too much labor (Philippon and Reshef, 2012; Bolton et al., 2016), and economies of scale fail to result in lower fees (Greenwood and Scharfstein, 2013; Malkiel, 2013; Philippon, 2015). Within this setting, we examine the effect of a government subsidy and show that it magnifies individuals' price insensitivity. Although the subsidy does not affect investors' sensitivity to *percent* fees, which remain constant in equilibrium, it reduces their sensitivity to *dollar* fees, which rise in equilibrium upon a switch from Roth to Traditional. As a result, in the presence of fixed costs, investors are subsidized in seeking more variety.

5 Conclusion

Under some simplifying assumptions about tax rates, a standard benchmark model yields an indifference result between front-loaded (Roth) and back-loaded (Traditional) taxation of retirement savings. Under either system, individuals obtain the same consumption in every period and the present value of government tax revenues is the same. The timing of taxation is different: 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 indifference result breaks down because the government is paying an estimated $16.1 billion a year in fees on its large implicit portfolio. In a stylized general equilibrium model, we also show that back-loading taxation inefficiently increases the amount of resources spent on asset management, thus reducing welfare.

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") has been discussed as part of 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 are progressive. With progressive taxation, lifetime taxes are more aligned with lifetime income under Traditional than under Roth. 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 municipal clerk). 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. 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 not affect all individuals equally. Individuals with high labor income but no expected income in retirement would be disproportionately affected, as their current marginal tax rate is high and their tax rate in retirement is low.[43]

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 exist in favor of either taxation scheme. During an individual's work life, back-loaded taxation could induce individuals to underestimate the future "tax bite" and lead to lower 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 focus on nominal contributions and savings and underweight future taxes.[44] 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

[43]As part of the debate, industry sources argued that a shift to Roth would endanger retirement security because Traditional provides more resources for retirement under progressive taxation. That argument, which focuses on the overall level of taxation, is distinct from ours, which focuses on distribution. If that argument is correct, under a Roth system the government could use the additional tax revenue to provide additional resources for retirees. Our model allows for differential tax rates during work life and retirement, which we consider to be a separate policy choice.

[44]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 may 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.

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 involves lower taxes or higher spending.[45] 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).

Finally, as discussed in Section 2, Traditional accounts cause the government to indirectly invest in equities. This equity exposure could be a benefit, if equity exposure were desired, and if the government were unable to obtain it in a more direct and cost-efficient way. 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, but 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 $16.1 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 and recently struck down by the courts. One of the stated motivations for the rule was protecting retail investors from aggressive marketing of high-fee products—especially senior investors that prepare to roll over their employer plan savings into an individual retirement accounts. 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 the rule itself.

[45]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).

References

Abel, A.B., 2001. The effects of investing Social Security funds in the stock market when fixed costs prevent some households from holding stocks. American Economic Review 91, 128–148.

Aiyagari, S.R., McGrattan, E.R., 1998. The optimum quantity of debt. Journal of Monetary Economics 42, 447–469.

Alpert, B.N., Rekenthaler, J., Suh, S., 2013. Global Fund Investor Experience 2013 Report. Technical Report. Morningstar Fund Research.

Anand, A., Irvine, P., Puckett, A., Venkataraman, K., 2012. Performance of institutional trading desks: An analysis of persistence in trading costs. Review of Financial Studies 25, 557–598.

Anderson, S.P., Renault, R., 1999. Pricing, product diversity, and search costs: A bertrand-chamberlin-diamond model. RAND Journal of Economics 30, 719–735.

Antolín, P., de Serres, A., de la Maisonneuve, C., 2004. Long-term budgetary implications of tax- favoured retirement plans. doi:10.1787/138080145732. OECD Economics Department Working Papers, No. 393, OECD Publishing.

Auerbach, A., 2004. How much equity does the government hold? American Economic Review 94, 155–160.

Ayres, I., Curtis, Q., 2015. Beyond diversification: The pervasive problem of excessive fees and 'dominated funds' in 401(k) plans. Yale Law Journal 124, 1346–1835.

Barclays, 2017. Bloomberg barclays index methodology. Retrieved at https://www.bbhub.io/indices/sites/2/2017/03/Index-Methodology-2017-03-17-FINAL-FINAL.pdf on 11/15/2017.

Barro, R.J., 1974. Are government bonds net wealth? Journal of Political Economy 82, 1095–1117.

Baumol, W.J., Goldfeld, S.M., Gordon, L.A., Koehn, M.F., 1990. The Economics of Mutual Fund Markets: Competition Versus Regulation. Rochester Studies in Managerial Economics and Policy, Kluwer Academic Publishers.

Ben-Rephael, A., Kandel, S., Wohl, A., 2011. The price pressure of aggregate mutual fund flows. Journal of Financial and Quantitative Analysis 46, 585–503.

Bergstresser, D., Chalmers, J.M.R., Tufano, P., 2009. Assessing the costs and benefits of brokers in the mutual fund industry. Review of Financial Studies 22, 4129–4156.

Berk, J., van Binsbergen, J., 2015. Measuring skill in the mutual fund industry. Journal of Financial Economics .

Berk, J.B., Green, R.C., 2004. Mutual fund flows and performance in rational markets. Journal of Political Economy 112, 1269–1295.

Bernheim, D., 2002. Handbook of Public Economics. Elsevier Science B.V.. chapter 18. Taxation and Savings. pp. 1173–1249.

Beshears, J., Choi, J.J., Laibson, D., Madrian, B.C., 2017. Does front-loading taxation increase savings? evidence from Roth 401(k) introductions. Journal of Public Economics 151, 84–95.

Bessembinder, H., Jacobsen, S.E., Maxwell, W.F., Venkataraman, K., 2018. Capital Commitment and Illiquidity in Corporate Bonds. Journal of Finance 73, 1615–1661.

Bogle, J.C., 2014. The arithmetic of 'all-in' investment expenses. Financial Analysts Journal 70, 13–21.

Boguth, O., Simutin, M., 2018. Leverage constraints and asset prices: Insights from mutual fund risk taking. Journal of Financial Economics 127, 325–341.

Bohn, H., 1990. Tax smoothing with financial instruments. The American Economic Review 80, 1217–1230.

Bolton, P., Santos, T., Scheinkman, J.A., 2016. Cream-skimming in financial markets. Journal of Finance 71, 709–736.

Brady, P., 2012. The tax benefits and revenue costs of tax deferral. Investment Company Institute, Washington, DC.

BrightScope, ICI, 2018. The BrightScope/ICI Defined Contribution Plan Profile: A Close Look at 401(k) Plans, 2015. Technical Report.

Burman, L., Gale, W.G., Weiner, D., 2001. The taxation of retirement saving: Choosing between front-loaded and back-loaded options. National Tax Journal 54, 689–702.

Busse, J.A., Chordia, T., Jiang, L., Tang, Y., 2018. Mutual fund trading costs. Available at SSRN: https://ssrn.com/abstract=2350583.

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

Carlin, B.I., 2009. Strategic price complexity in retail financial markets. Journal of Financial Economics 91, 278–287.

Che, J., Qian, Y., 1998. Insecure property rights and government ownership of firms. The Quarterly Journal of Economics 113, 467–496.

Chen, J., Hong, H., Huang, M., Kubik, J.D., 2004. Does fund size erode mutual fund performance? the role of liquidity and organization. American Economic Review 94, 1276–1302.

Choi, J., Huh, Y., 2017. Customer liquidity provision: Implications for corporate bond transaction costs. SSRN working paper. Retrieved at https://ssrn.com/abstract=2848344 on 11/15/2017.

Christoffersen, S.E., Musto, D.K., 2002. Demand curves and the pricing of money management. Review of Financial Studies 15, 1499–1524.

Coates, J.C., Hubbard, R.G., 2007. Competition in the mutual fund industry: evidence and implications for policy. Harvard John M. Olin Discussion Paper No. 592, available at http://ssrn.com/abstract=1005426.

Collins, S., Holden, S., Duvall, J., Chism, E.B., 2016. The economics of 401(k) plans: Service, fees and expenses, 2015. ICI Research Perspective 24.

Committee for a Responsible Federal Budget, 2016. Reconciliation 101. Retrieved on 17/12/2017 at http://www.crfb.org/papers/reconciliation-101.

Committee on Finance of the U.S. Senate, 1997. Expanding IRA's. U.S. Government Printing Office.

Conrad, J.S., Johnson, K.M., Wahal, S., 2001. Institutional trading and soft dollars. Journal of Finance 56, 397–416.

Cooper, M., Halling, M., Yang, W., 2018. The persistence of fee dispersion among mutual funds. Available at https://ssrn.com/abstract=1456079.

Copeland, C., 2016. 2014 update of the EBRI IRA database: Ira balances, contributions, rollovers, withdrawals, and asset allocation. EBRI Issue Brief 424.

Cremers, M., Ferreira, M.A., Matos, P., Starks, L., 2016. Indexing and active fund management: International evidence. Journal of Financial Economics 120, 539–560.

Diamond, P., Geanakoplos, J., 2003. Social security investment in equities. American Economic Review 93, 1047–1074.

Dyck, A., Pomorski, L., 2011. Is bigger better? size and performance in pension plan management. Rotman School of Management Working Paper No. 1690724, available at http://ssrn.com/abstract=1690724.

Economides, N., 1989. Symmetric equilibrium existence and optimality in differentiated product markets. Journal of Economic Theory 47, 178–194.

Elton, E.J., Gruber, M.J., Busse, J.A., 2004. Are investors rational? choices among index funds. Journal of Finance 59, 261–288.

Fama, E.F., French, K.R., 2010. Luck versus skill in the cross-section of mutual fund returns. The Journal of Finance 65, 1915–1947.

Feenberg, D.R., Skinner, J., 1989. Sources of IRA saving. Tax Policy and the Economy 3, 25–46.

French, K.R., 2008. Presidential address: The cost of investing. The Journal of Finance 53, 1537–1573.

Gabaix, X., Gopikrishnan, P., Plerou, V., Stanley, H.E., 2006. Institutional investors and stock market volatility. Quarterly Journal of Economics 121, 461–504.

Gabaix, X., Laibson, D., 2006. Shrouded attributes, consumer myopia, and information suppression in competitive markets. Quarterly Journal of Economics 121, 505–540.

Gao, X., Livingston, M., 2008. The components of mutual fund fees. Financial Markets, Institutions and Instruments 17, 197–223.

Gârleanu, N.B., Pedersen, L.H., 2018. Efficiently inefficient markets for assets and asset management. Journal of Finance 73, 1663–1712.

Geanakoplos, J., Mitchell, O., Zeldes, S.P., 1999. Social Security Money's Worth. Pension Research Council, University of Pennsylvania Press. chapter 5. Prospects for Social Security Reform, pp. 79–151.

Gennaioli, N., Shleifer, A., Vishny, R., 2015. Money doctors. Journal of Finance 70, 91–114.

Gil-Bazo, J., Ruiz-Verdú, P., 2008. When cheaper is better: Fee determination in the market for equity mutual funds. Journal of Economic Behavior and Organization 67, 871–885.

Gil-Bazo, J., Ruiz-Verdú, P., 2009. The relation between price and performance in the mutual fund industry. The Journal of Finance 64, 2153–2183.

Goldstein, M.A., Hotchkiss, E.S., 2018. Providing Liquidity in an Illiquid Market: Dealer Behavior in U.S. Corporate Bonds. Journal of Financial Economics forthcoming.

Gong, Q., Liu, Q., Zhang, Y., 2016. Optimal product differentiation in a circular model. Journal of Economics 119, 219–252.

Greenwood, R., Scharfstein, D., 2013. The growth of finance. Journal of Economic Perspectives 27, 3–28.

Heathcote, J., 2005. Fiscal policy with heterogeneous agents and incomplete markets. Review of Economic Studies 72, 161–188.

Henderson, B.J., Pearson, N.D., 2011. The dark side of financial innovation: A case study of the pricing of a retail financial product. Journal of Financial Economics 100, 227–247.

Holmström, B., Tirole, J., 1998. Private and public supply of liquidity. Journal of Political Economy 106, 1–40.

Holzmann, R., Hinz, R., 2005. Old Age Income Support in the 21st Century. The World Bank.

Hortacsu, A., Syverson, C., 2004. Product differentiation, search costs, and competition in the mutual fund industry: A case study of S&P 500 index funds. Quarterly Journal of Economics , 403–456.

Hubbard, R.G., Koehn, M.F., Ornstein, S.I., Audenrode, M.V., Royer, J., 2010. The Mutual Fund Industry: Competition and Investor Welfare. Columbia Business School Publishing.

ICI, 2018. 2018 Investment Company Fact Book. Investment Company Institute. Available at http://www.icifactbook.org/deployedfiles/FactBook/Site\%20Properties/pdf/2018/2018_factbook.pdf.

Ince, O., Kadlec, G.B., McKeon, S.B., 2018. Institutional counterparties and performance. Available at SSRN: https://ssrn.com/abstract=3172301.

Investor Economics, 2012. Mutual fund MERs and cost to customer in Canada: Measurement, trends and changing perspectives. Retrieved on 2017/12/17 at https://www.ific.ca/wp-content/uploads/2013/08/Canadian-Study-Mutual-Fund-MERs-and-Cost-to-Customer-in-Canada-September-2012.pdf/1655/.

Iwry, J.M., John, D.C., 2009. Pursuing universal retirement security through automatic IRAs. Brookings Institution, Retirement Security Project, Research Report 2009-3.

Kahn, V.M., 2002. Investing; mutual fund expertise, for rent. The New York Times .

Latzko, D.A., 1999. Economies of scale in mutual fund administration. Journal of Financial Research 22, 331–339.

Lucas, D.J., Zeldes, S.P., 2009. How should public pension plans invest? American Economic Review 99, 527–532.

Lurie, I.Z., Ramnath, S.P., 2011. Long-run changes in tax expenditures on 401(k)-type retirement plans. National Tax Journal 64, 1025–1038.

Malkiel, B.G., 2013. Asset management fees and the growth of finance. Journal of Economic Perspectives 27, 97–108.

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.

Office of Management and Budget, 2014. Budget of the U.S. Government. Analytical Perspectives. Fiscal Year 2015. Technical Report. United States Government.

Osborne, G., 2015. Strengthening the incentive to save: a consultation on pensions tax relief. Her Majesty's Treasury Cm 9102.

Pástor, L., Stambaugh, R.F., 2012. On the size of the active management industry. Journal of Political Economy .

Pástor, L., Stambaugh, R.F., Taylor, L.A., 2015. Scale and skill in active management. Journal of Financial Economics 116, 23–45.

Pedersen, L.H., 2018. Sharpening the arithmetic of active management. Financial Analysts Journal 74, 21–36.

Perold, A.F., Salomon, R.S.J., 1991. The right amount of assets under management. Financial Analysts Journal 47, 31–39.

Petajisto, A., 2009. Why do demand curves for stocks slope down? Journal of Financial and Quantitative Analysis 44, 1013–1044.

Philippon, T., 2015. Has the U.S. finance industry become less efficient? on the theory and measurement of financial intermediation. American Economic Review 105, 1408–1438.

Philippon, T., Reshef, A., 2012. Wages and human capital in the U.S. financial industry: 1909–2006. Quarterly Journal of Economics 127, 1551–1609.

Phillips, D., 2013. Mutual fund urban myths. Morningstar Adviser , 80Retrieved at http://news.morningstar.com/articlenet/article.aspx?id=600657 on 11/15/2017.

Pine, A., 1989. GOP senators offer capital gains cut, new type of IRA. Los Angeles Times . Retrieved online on 10/19/2016.

Pool, V.K., Sialm, C., Stefanescu, I., 2016. It pays to set the menu: Mutual fund investment options in 401(k) plans. Journal of Finance 71, 1779–1812.

Romaniuk, K., 2013. Pension fund taxation and risk-taking: should we switch from the EET to the TEE regime? Annals of Finance 9, 573–588. doi:10.1007/s10436-012-0204-3.

Rosshirt, D.E., Parker, S.A., Pitts, D.A., 2014. Inside the Structure of Defined Contribution/401(k) Plan Fees, 2013: A study assessing the mechanics of the 'all-in' fee. Technical Report. Deloitte Consulting LLP.

Rowley, Jr., J.J., Dickson, J.M., 2012. Mutual funds—like ETFs— have trading volume. Technical Report. Vanguard.

Salop, S.C., 1979. Monopolistic competition with outside goods. Bell Journal of Economics 10, 141–156.

Shive, S., Yun, H., 2013. Are mutual funds sitting ducks? Journal of Financial Economics 107, 220–237.

Tergesen, A., Rubin, R., 2017. Talk of retirement savings cap rattles financial industry. Wall Street Journal 10/21/2017.

Thaler, R.H., 1994. Psychology and savings policies. The American Economic Review, Papers and Proceedings of the Hundred and Sixth Annual Meeting of the American Economic Association 84, 186–192.

Thrift Savings Plan, 2018. Highlights.

Tucker, M., 2011. Q&a on bond funds and churn: Why turnover can be misleading. BlackRock Blog Retrieved at `https://www.blackrockblog.com/2011/09/13/qa-on-bond-funds-and-churn-why-turnover-can-be-misleading/` on 11/15/2017.

Whitehouse, E., 2007. Pensions Panorama—Retirement-Income Systems in 53 Countries. The World Bank.

Woodford, M., 1990. Public debt as private liquidity. American Economic Review 80, 382–388. Papers and Proceedings of the Hundred and Second Annual Meeting of the American Economic Association.

Wurgler, J., 2011. Challenges to Business in the Twenty-First Century: The Way Forward. American Academy of Arts and Sciences. chapter 4.

Yan, X.S., 2008. Liquidity, investment style, and the relation between fund size and fund performance. Journal of Financial and Quantitative Analysis 43, 741–768.