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

Mattia Landoni† and Stephen P. Zeldes‡

December 4, 2016

*The authors would like to acknowledge the helpful comments of Dan Bergstresser, Kent Daniel, Charles Jones, David Laibson, Brett Myers, Emi Nakamura, seminar participants at the Columbia Business School Finance Free Lunch, the Columbia Macro Lunch, CEIBS, Baruch College, Texas A&M Mays Business School, and conference participants at the Red Rock Finance Conference 2016. We thank Abdullah Al-Sabah for outstanding 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, Cox School of Business, Southern Methodist University

‡Benjamin M. Rosen Professor of Finance and Economics, Columbia Business School, Columbia University; and NBER

Abstract

Governments incentivize retirement saving by allowing individuals to contribute to tax-advantaged accounts where 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 In accounts with "back-loaded" taxation, the individual contributes pretax money and pays taxes when the money is withdrawn.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 both the individual and the government are indifferent between the two types of accounts. We add investment management fees to the benchmark model and show that the neutrality result breaks down. Assuming fees are fixed as a percent of assets under management (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 government $14 billion per year. We then ask whether this result holds in general equilibrium, where fees as a percent of AUM are allowed to vary. 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 produces a larger asset-management industry. Finally, we use the general equilibrium model to examine welfare implications. In a rough calibration of the model, we find that this increase in the size of the asset management industry reduces consumer welfare.

1 Introduction

Retirement savings systems around the world incorporate tax incentives designed to increase saving and enhance retirement security. These incentive schemes differ with respect to 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 provides a benefit, in that 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 deferral. An alternative system is one in which taxation is "front-loaded", i.e., contributions are made with after-tax income, but then neither the principal nor returns are taxed at any point in the future.[1]

In the U.S., almost all assets in defined contribution retirement accounts such as 401(k) and 403(b) plans as well as 93 percent of assets in individual retirement accounts (IRAs) use the "traditional" form in which taxes are back-loaded. Front-loaded taxation was introduced in the U.S. in 1997 with the creation of "Roth" accounts, named after the legislation's sponsor. For clarity, in the rest of the paper we refer to accounts with back-loaded taxation as Traditional, and accounts with front-loaded taxation as Roth.[2]

Under a few simplifying assumptions, including the constancy of the tax rate across working and retirement years, some basic math shows a benchmark neutrality result: Roth and Traditional accounts yield identical consumption for individuals, both in the working years and in retirement, and individuals are indifferent between the two types of plans. Under the additional standard assumption that the discount rate for the government is the same as the expected return on the underlying assets, the present value of cash flows to and from the government is also identical under the two accounts.

While Roth and Traditional are equivalent on a present-value basis, they differ in terms of the timing of cash flows to the government, yielding different short-run budget impacts. The

[1] In addition to taxing money flows when contributed to the account or when withdrawn, there is a wide range of other possibilities. For instance, in Australia income contributed to the superannuation scheme is taxed at all three possible stages (contribution, returns, and withdrawal) but at favored rates (Whitehouse, 2007).

[2] Following Beshears et al. (forthcoming) and the World Bank (Holzmann and Hinz, 2005), we use the terms "front-loaded taxation" and "back-loaded taxation" to refer to the timing of the *taxation*. Income contributed to Roth accounts is taxed upfront, hence the term "front-loaded taxation". Income contributed to Traditional accounts is deductible upfront and taxed upon withdrawal, hence the term "back-loaded taxation". A source of potential confusion is others' 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, those involved in the discussion of the 1997 law that introduced Roth accounts referred to Roth accounts 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.

front-loaded taxation in Roth accounts generates more government revenue in the individual's working years and less revenue in the retirement years relative to Traditional accounts. Thus, a shift from Traditional to Roth accounts generates a short-run budget improvement for the government.[3]

To facilitate comparison of the two types of accounts, we decompose the Traditional account into two virtual accounts: i) a Roth account and ii) a separate implicit government account. The government account contains the assets earmarked to pay future taxes when the investor takes distributions from the account. The investor is indifferent between owning an actual Roth account and owning a virtual Roth account as part of a Traditional account. From the government's perspective, whether it collects its revenue now or later is irrelevant to the present-value calculation.

While this equivalence result is quite general, it does not survive the addition of one crucial bit of realism: running retirement plans and managing assets involves fees paid to record keepers, asset managers, and financial advisors, as well as trading costs. These fees are typically paid as a percentage of assets under management (AUM), and thus reduce the net returns paid on retirement accounts. They thereby introduce a wedge between the net returns and the government discount rate and break the present-value government equivalence result described above. Under back-loaded taxation, the government is effectively paying investment fees on its substantial implicit portfolio, something it avoids under front-loaded (Roth) taxation.

It is possible, of course, that the additional fees that the government pays on its virtual accounts are compensation for better performance, or other services provided to the government by asset managers. We think this is unlikely for two reasons. First, it is not clear that individuals who pay higher fees on retirement accounts receive any additional benefits as a result. But even if some or all of them do, it is unlikely that the government captures this benefit on its virtual account, because it is implicitly holding a fraction of all retirement portfolios, and many of the potential benefits of higher fees will cancel out in the aggregate. For example, some of the higher costs might be associated with creating funds or asset allocations that are customized to a particular group of individuals, such as target date funds that adjust asset allocations as individuals age and get closer to the target retirement date. While this might create value for individuals, holding target date funds of *all* target dates will not create value to the government. Similarly, the government will not benefit from paying higher fees to invest in *all* the funds that focus on style (conservative/aggressive, value/growth), bond maturity (long/short), sector (small cap/large cap,

[3] In fact, one of the motivations for introducing Roth accounts was to help "fund" cuts in the capital gains tax (Pine, 1989).

junk/investment grade) or industry. Finally, with regard to equity funds, the government is unlikely to benefit, in aggregate, from active asset management. 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."

We perform a back-of-the-envelope calculation assuming that the fees as a percent of AUM remain the same under Roth and Traditional. To estimate the size of the government's implicit account, we multiply the total amount of tax-deferred assets in DC plans and IRAs (\$13.9 trillion) by 20%, a reasonable estimate of the marginal tax rate in retirement, leading to our estimate of \$2.8 trillion of retirement assets. Our estimate of assets ignores DB plans. Including corporate and state and local government DB plans would add \$6.5 trillion of tax-deferred money, and thus another \$1.3 trillion of an implicit government account. We conservatively estimate asset-weighted fees to be about 80 basis points (bps) based on the lowest asset-weighted estimates available. We assume that 35% of fees paid by the government are recovered via corporate taxation of the asset managers. Multiplying \$2.8 trillion by $.80\% \times (1 - .35)$, we reach our estimate of the annual costs of about \$14 billion per year. In other words, the government could achieve savings equivalent to \$14 billion per year by forcing the conversion of all existing tax-deferred retirement accounts into Roth accounts.

This calculation takes the supply side as given: that is, we rely on the partial equilibrium assumption that investment management fees as a percentage of assets under management are independent of whether retirement accounts are structured as Traditional or Roth. The extent to which this is true in general equilibrium depends on the underlying cost structure and competitive landscape of the asset management industry. If aggregate asset management costs are proportional to aggregate assets under management, then our partial equilibrium result will also hold exactly in general equilibrium.

In order to examine what happens if economies of scale are present, we then examine an extreme alternative by assuming that all asset management costs are fixed and there are no variable costs. Intuitively, these economic assumptions would lead to a monopoly, something obviously at odds 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, and second, entry is free. Specifically, we model competition among funds as spatial competition (Salop, 1979; Tirole, 1988, Ch.7) in a two-period, general equilibrium model in which each fund only needs a fixed amount of labor to operate. A switch from Roth to Traditional continues to increase assets under management. Despite the assumption that any fund could costlessly expand to manage the additional assets, the model generates increases in (i) the equilibrium number of funds, (ii) employment in the asset management industry, and (iii) the aggregate dollar fees collected.

We next examine whether the larger asset management industry under Traditional accounts can be optimal. Consider a Roth-based system as the starting point. If that were the social optimum, a shift to Traditional would cause there to be too many funds. Because we are departing from the optimum, however, the total loss would be relatively small, as the social welfare function is flat at the optimum. Only if the equilibrium number of funds increased significantly would social welfare suffer serious consequences. Alternatively, if a Roth-based system started with too few funds, then a shift to Traditional would get society closer to the optimum, and it would be beneficial. Finally, if a Roth-based system started with too many funds, the shift to Traditional would be especially pernicious, as it would bring society even farther from the optimum, starting at a point where the social welfare function is already steep. Under a rough calibration, we find that the equilibrium number of funds when all retirement plans are Roth turns out to be about twice as large as a rational planner would set. Since the number of funds in a Roth system is already higher than the social optimum, shifting to a Traditional system (and the higher number of funds) generates a substantial welfare loss.

Our results have implications for public policies related to retirement saving. Policy options include mandating the use Roth accounts for new contributions, converting existing tax-deferred accounts to Roth, and explicitly segregating the virtual government account (with a mechanism similar to withholdings) so that the government is able to negotiate lower fees. In addition, since the government is paying investment management fees on its implicit portion of retirement accounts, our results suggest an additional rationale for requiring retirement advisers to abide by a fiduciary standard, as established by the new U.S. Department of Labor fiduciary rule set to go into effect by April 2017. Not only is the welfare of retirees at stake, but that of taxpayers as well.

Our paper is related to a recent literature on the optimal size of the financial services industry. Greenwood and Scharfstein (2013) note that the financial services industry (encompassing insurance, securities and credit intermediation) as a share of GDP doubled in size in the last 50 years, going from 4% to 8%. Half of this 4 percentage point increase (1.5–2 percentage points) is due to growth of the asset management industry, which has managed to keep its revenue a relatively stable fraction of the stock market. Malkiel (2013) argues that in spite of a more than 100-fold increase in assets under management, the benefits from the vast economies of scale inherent to the asset management industry have accrued to industry insiders, because fees (as a percentage of assets under management) have not fallen proportionately. French (2008) and Fama and French (2010) attempt to quantify the amount of resources spent in the zero-sum game of attempting to beat the market. Philippon and Reshef (2012) find empirically that financial deregulation is associated with greater skill intensity, increased job complexity, and higher wages for finance employees. Bolton et al. (2016) show in theory that it is possible for the financial industry to extract excessively high rents for the provision of financial services, thus attracting too much talent. Our study features another mechanism by which, because of search frictions, the financial industry attracts too much labor, and points out that under reasonable assumptions this mechanism not only exists, but it is magnified by government policy.

We also contribute to the literature on limited price sensitivity in retail financial markets. The existing literature has proposed explanations based on the inability of retail investors to observe shrouded prices of complex financial products (Gabaix and Laibson, 2006; Carlin, 2009; Henderson and Pearson, 2011), the inability to precisely observe the quality of fund management (Gil-Bazo and Ruiz-Verdú, 2008, 2009), or the unwillingness to sever relationships with brokers (Bergstresser et al., 2009) or trusted advisors (Gennaioli et al., 2015). We propose a different explanation: a government subsidy. Our proposed explanation is not mutually exclusive with the existing ones; in fact, it is complementary. Our model of spatial competition is very general and it stands in for any form of limited price sensitivity that results in too many firms. Conditional on having too many firms, the effect of a subsidy on social welfare is particularly deleterious.

Our paper is structured as follows. Section 2 derives the basic result that the investor and the government are indifferent between Roth and Traditional in a benchmark partial-equilibrium model. In section 3, we introduce a realistic assumption: investment fees. We show that the basic indifference result still holds for the investor, but not for the government: the present value of government revenue is higher under Roth. In Section 4 we provide an asset-weighted estimate of total investment fees applicable to retirement accounts. Section 5 examines and calibrates a simple general equilibrium model in which the size of the asset management industry is determined in equilibrium. We show that the non-neutrality result still holds. Section 6 briefly examines the implications for public policy and concludes.

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

In this section we describe the classic result that under flat taxation and no time variation in the tax rate, both the individual and the government are indifferent between a front-loaded tax scheme ("Roth") and a back-loaded one ("Traditional"). This result is discussed in detail

Account type i	Abbreviation	Type of taxation	Tax on initial contribution	Tax rate on investment returns τ^i	Tax on retirement payouts
Taxable	TTE	Immediate	τ_L	$\tau^{Tax} = \tau_I > 0$	0
Traditional	EET	Deferred	0	$\tau^{Trad} = 0$	τ_R
Roth	TEE	Immediate	τ_L	$\tau^{Roth} = 0$	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.

in Brady (2013).

2.1 Base assumptions

To begin, we assume a model with no uncertainty. Individuals earn one dollar of pretax labor income at time 0 and put it into an account, where their savings earn additional investment income. Income tax rates are flat, i.e., they do not vary with the level of income. However, they may differ across the life cycle and across different types of income:[4]

- labor income during the working years is taxed at a rate τ_L;
- retirement income (including principal and returns from Traditional retirement accounts) is taxed at a rate τ_R;
- investment income is taxed at a rate τ^i that varies depending on the type of account i.

Table 1 presents three possible way of taxing retirement savings:

- *Taxable account*: all labor income is taxed at rate τ_L when earned. Intermediate investment returns are taxed at a rate $\tau^{Tax} = \tau_I > 0$. This 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*: income tax on retirement account contributions is deferred until the time of retirement T, when the account is assumed to be liquidated

[4] In practice of course, the tax system does not have flat rates, but is instead progressive, with marginal tax rates increasing with income. When coupled with uncertain labor income and/or asset returns, marginal tax rates become stochastic, introducing complications to the analysis that we do not address in this paper.

and all the money is paid out as retirement income, taxed at a rate τ_R. Intermediate investment returns are not taxed, i.e., $\tau^{Trad} = 0$. 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*: all labor income is taxed at rateτ_L when earned. Intermediate investment returns are not taxed, i.e., $\tau^{Roth} = 0$. 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.

Money in the account is invested in the only one asset in positive supply, government bonds, paying a certain return of r.

2.2 Basic neutrality result: investor final wealth and present value of government revenue

Table 2 shows the initial and future cash flows for both the individual and the government, under the assumption that the individual chooses to contribute \$100 of pretax money to a retirement account. With a Traditional account, the government has no revenue upfront, and the individual's account balance is 100. At time T, when the individual retires and the account is liquidated, the balance ($100 \cdot e^{rT}$) is paid out and taxed. The individual receives $100 \cdot e^{rT} \cdot (1 - \tau_R)$ and the government receives $100 \cdot e^{rT} \cdot \tau_R$. Conversely, with a Roth account, the government taxes the money upfront receiving $100 \cdot \tau_L$. The individual's starting balance is thus $100 \cdot (1 - \tau_L)$. No additional taxation happens, and therefore at time T the individual can keep the entire balance $100 \cdot (1 - \tau_L) \cdot e^{rT}$.

It is immediate to see that if $\tau_L = \tau_R$, the individual's beginning and ending wealth are the same. With constant wealth and constant prices, the individual would choose the same consumption plan with both Roth and Traditional, and would therefore be indifferent between the two. The only price in the economy is the interest rate, i.e., the price of shifting consumption from work life (time 0) to retirement (time T). We begin by assuming that the interest rate is unchanged.

The government's cash flow differs across plans—with Roth accounts revenue is received up front, whereas with Traditional accounts the revenue is deferred until the future. But assuming that the government discount rate is equal to the interest rate on government bonds, the present value as of time 0 of the revenue streams are equal: under Traditional it is simply $e^{-rT} \cdot 100 \cdot e^{rT} \cdot \tau_R = 100 \cdot \tau_R = 100 \cdot \tau_L$, i.e., the same as the immediate revenue under Roth. The government will therefore be indifferent (in a present value sense) between the accounts.

Account	Individual			Government		
	Initial balance	Future balance	Final payout	Initial revenue	Future revenue	PV @ r
Traditional	100	$100e^{rT}$	$100e^{rT} \cdot (1-\tau_R)$	0	$100e^{rT}\tau_R$	$100\tau_R$
Roth	$100 \cdot (1-\tau_L)$	$100 \cdot (1-\tau_L)\, e^{rT}$	$100 \cdot (1-\tau_L)\, e^{rT}$	$100\tau_L$	0	$100\tau_L$
Traditional - Roth	$100\tau_L$	$100\tau_L e^{rT}$	$100e^{rT} \cdot (\tau_L-\tau_R)$	$-100\tau_L$	$100e^{rT}\tau_R$	$-100 \cdot (\tau_L-\tau_R)$
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 labor income (τ_L) is the same as the tax rate on retirement income (τ_R), 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).

At any time $t \in [0, T]$, the balance in a Traditional account can be decomposed into three separate virtual accounts to reflect this indifference result:

$$V_t^{Trad} = e^{rt}\left[(1-\tau_R)+\tau_R\right] = e^{rt}\left[(1-\tau_L)+(\tau_L-\tau_R)+\tau_R\right]. \tag{1}$$

This decomposition is depicted in Fig. 1. The first term is a "Roth equivalent" account of size $1-\tau_L$, belonging to the individual. This account has exactly the same size as the entire account would have under a Roth system. The second term is a "transfer" account of size $\tau_L-\tau_R$, which constitutes the true difference in final wealth between Roth and Traditional. This account represents a future transfer from the government to the individual (or vice versa, if $\tau_R > \tau_L$). This transfer is not a consequence of inherent differences between Roth and Traditional, but rather of the difference in tax rates between labor income and retirement income, a separate policy choice.

Finally, the third term is a "deferred tax asset" account of size τ_R. This account represents wealth that is already earmarked to pay future taxes when the money is paid out of the account. Throughout the paper, we refer to this term as the government's virtual account or implicit account. If $\tau_L = \tau_R$, the present value of this account that the government acquires under Traditional is equal to the revenue that the government receives under Roth, even though the government receives different cash flows at different times under the two schemes.

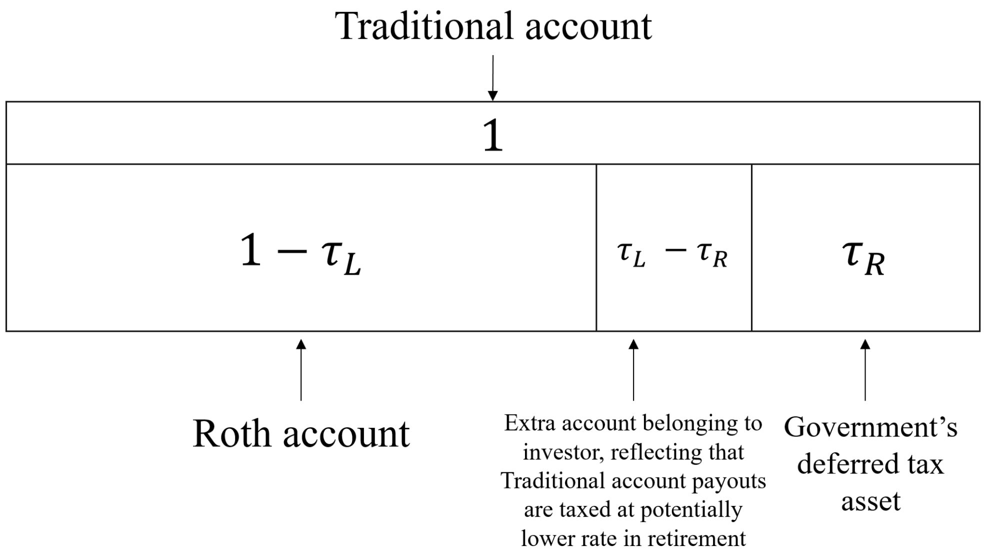

Figure 1: **Decomposition of Traditional tax-deferred account into three virtual accounts.** Under some simplifying assumptions, the savings in a Traditional account can be decomposed into a Roth equivalent account (i.e., the money the investor would have, had he contributed the money to a Roth account); a transfer account (i.e., the investor's difference in final wealth due to different tax rates between work life and retirement); and a deferred tax asset account (i.e., money earmarked to pay future taxes).

Up until now we have assumed that interest rates are the same under the two systems, but we now show the Ricardian result that a shift from Roth to Traditional does not affect equilibrium interest rates. Assuming that individuals act rationally, at any given interest rate they consume the same amount under Roth and Traditional. Thus, if the balance in Traditional is S, the balance in Roth will be $S(1-\tau_L)$. The extra balance in Traditional is invested in government bonds, creating additional demand for government bonds equal to $S \cdot \tau_L$. On the other hand, under Traditional, the government faces a revenue shortfall (relative to Roth) of $S \cdot \tau_L$. Assuming for simplicity that $\tau_R = \tau_L$ and that government expenditure is exogenous, the government must issue an amount $S \cdot \tau_L$ of new bonds, adding to the existing supply. Thus, demand for government bonds still equals supply, and the equilibrium interest rate will remain unchanged.

2.3 Extended neutrality result: adding a risky asset

Now suppose that there are two assets: the government bond yielding r and a risky asset (stocks) with expected return r_s. The appropriate discount rate for this expected risky return is also r_s. To show the neutrality result in the presence of stocks, let us assume first that all retirement accounts are held in bonds, as in the example above. Now allow investors to switch from holding bonds to holding a percentage α of their retirement accounts in stocks. From the investors' perspective, the Roth account would yield an identical outcome to the Traditional. However, the demand for risky assets would be greater under Traditional than under Roth, because a percentage α of the government account would also be held in stocks. In addition, under Traditional the government would now hold an unhedged position, because it has issued riskless bonds, but has a future claim on risky assets. The Ricardian result requires the government to take additional action and sell stocks from an existing portfolio, or else short stocks if it doesn't have such a portfolio. The increase in supply of stocks will exactly offset the increase in demand, and Ricardian equivalence will hold: interest rates and stock returns will remain unchanged, as will household consumption.

Looking at the present value of cash flows, we can also derive the indifference result for the government. If the government borrows an extra dollar by issuing an additional riskless bond, its cost of capital will be r. If the government receives future revenue that is proportional to stock returns, the appropriate discount rate for those cash flows will be r_s. It is then easy to see that the present value of government tax revenue is τ_L under a Roth account and τ_R under a Traditional account. Thus, if $\tau_L = \tau_R$, both the individual and the government are still indifferent between traditional and Roth.

The Ricardian result above assumed that the government was not initially constrained in its holdings of stocks, i.e. there was nothing preventing it from holding either more or less stocks. In practice, it could be that the government would like to hold more stocks, but is constrained from doing so. In this case, having Traditional accounts would ease the binding constraint and improve welfare. On the other hand, it could be that the government already had too large an exposure to the stock market (e.g. because future income taxes are tied directly or indirectly to future stock market performance), but was constrained from reducing its exposure. In this case, having Traditional accounts worsens the binding constraint by forcing the government to hold even more stocks than it otherwise would under Roth. The arguments here parallel those in the literature on the costs and benefits of the Social Security Trust Fund holding equities. See, e.g. (Geanakoplos et al., 1999; Abel, 2001; Diamond and Geanakoplos, 2003).

3 The effect of asset management fees

In this Section, we return to the world with just a risk-free asset and add one crucial bit of realism: asset management fees. Under this additional assumption, the individual is still indifferent, but the government is no longer indifferent, preferring the front-loaded scheme ("Roth") to the back-loaded scheme ("Traditional").

Suppose that all accounts are managed by an asset management firm. At time t, the asset management firm levies proportional fees $f > 0$. f is a fixed fraction of account size. Under these assumptions, the individual's final retirement wealth is lower, but still the same across Traditional and Roth. Both accounts grow at a net-of-fee rate of $r - f$. The left panel of Table 3 calculates the final payouts for the individual. Under Traditional, the initial balance is 100, and the final aftertax distribution from the account is $100 \cdot e^{(r-f)T} \cdot (1 - \tau_R)$. Under Roth, the initial balance is $100 \cdot (1 - \tau_L)$, and the final distribution from the account is $100 \cdot (1 - \tau_L) \cdot e^{(r-f)T}$. It is easy to see that if $\tau_R = \tau_L$ the indifference result for the individual is preserved. In present value, the individual's retirement wealth is simply $e^{-fT} \cdot 100 \cdot (1 - \tau_L)$. The first term ($e^{-fT}$) measures the extent to which fees erode retirement wealth.

The right panel Table 3 calculates the present value of tax revenue for the government with fees. To begin, we assume for simplicity that the government does not tax the asset manager's income. Clearly, the stream of tax revenue cash flows is different between Traditional and Roth. Unlike in the benchmark case, the present value of these cash flows is also different. Under the assumption that $\tau_R = \tau_L$, the individual is still indifferent because his or her final wealth is equally eroded by fees regardless of account type. On the other hand, the government has unambiguously lower present value of tax revenue under Traditional:

$$\textbf{Tax Revenue}^{Trad} - \textbf{Tax Revenue}^{Roth} = -100\tau_L \cdot \left(1 - e^{-fT}\right) < 0. \qquad (2)$$

This formula has an intuitive interpretation: $100\tau_L$ is the initial size of the government's virtual account of Traditional, and $\left(1 - e^{-fT}\right)$ is the fraction of the account that gets eroded by fees.

Now assume that the government levies a corporate tax τ_C on the asset manager's profits. For simplicity, conservatively assume that every additional dollar of fee revenue equals profit for the asset manager. Now the government has not only the initial and final revenue, but also a stream of corporate tax revenues that grows at the same rate as the account balance. The algebra is slightly more convoluted, but the end result is still amenable to an intuitive interpretation:[5]

$$\textbf{Tax Revenue}^{Trad} - \textbf{Tax Revenue}^{Roth} = -100\tau_L \cdot \left(1 - e^{-fT}\right) \cdot (1 - \tau_C) < 0, \qquad (3)$$

[5] For a Traditional account, the present value of corporate tax revenues is equal to $100 \cdot f \cdot \tau_C \cdot A(r, r - f, T)$, i.e. the initial account balance, times the percentage fees f (to obtain the asset manager's instantaneous revenue flow) times the corporate tax rate τ_C (to obtain the government's instantaneous corporate tax revenue flow) times a growing annuity term $A(r, r - f, T)$. For a given growth rate g, $A(r, g, T) = \left[1 - e^{-(r-g)T}\right] / (r - g)$ is the present value of a unit flow growing at a rate g until time T discounted at rate r. Similarly, the present value of corporate tax revenues for a Roth is and $100(1 - \tau_L) \cdot f \cdot \tau_C \cdot A(r, r - f, T)$, and therefore a Traditional yield an additional $100\tau_L \cdot f \cdot \tau_C \cdot A(r, r - f, T)$ in corporate tax revenues. Adding this term to (2), we obtain (3).

Account	Individual			Government		
	Initial balance	Future balance	Final payout	Initial revenue	Future revenue	PV @ r
Traditional	100	$100e^{(r-f)T}$	$100e^{(r-f)T} \cdot (1-\tau_R)$	0	$100e^{(r-f)T} \cdot \tau_R$	$100\tau_R$
Roth	$100 \cdot (1-\tau_L)$	$100e^{(r-f)T} \cdot (1-\tau_L)$	$100e^{(r-f)T} \cdot (1-\tau_L)$	$100\tau_L$	0	$100\tau_L$
Traditional - Roth	$100\tau_L$	$100\tau_L \cdot e^{(r-f)T}$	$100e^{(r-f)T} \cdot (\tau_L - \tau_R)$	$-100\tau_L$	$100e^{(r-f)T} \cdot \tau_R$	$-100\left(\tau_L + -e^{-fT}\tau_R\right)$
If $\tau_R = \tau_L$			0			$-100\tau_L \cdot \left(1-e^{-fT}\right)$

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 labor income (τ_L) is the same as the tax rate on retirement income (τ_R), 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).

where the first two terms are the same as in the case with no corporate tax, and $(1-\tau_C)$ is the fraction of fee income that is *not* recaptured by the government via taxation of the asset manager.

If $\tau_R \neq \tau_L$, the individual is not indifferent between Roth and Traditional, just as in the benchmark case. The Traditional account can be still decomposed into three virtual accounts, as shown in Figure 2: a Roth equivalent, a transfer account, and a deferred tax asset belonging to the government. However, the existence of a virtual transfer account due to a difference in the tax treatment of labor income and retirement income does not create any additional inefficiency. The inefficiency is created by the government's leaving an amount $100 \cdot \tau_L$ in the account at time 0. How this amount is ultimately split between the government and the individual does not matter. At time T, the individual simply receives an additional $100 \cdot e^{-fT}(\tau_L - \tau_R)$, and the government's tax revenue drops by the same amount:

$$\textbf{Tax Revenue}^{Trad} - \textbf{Tax Revenue}^{Roth} =$$

$$= -100 \cdot e^{-fT}(\tau_L - \tau_R) - 100 \cdot \tau_L \left(1 - e^{-fT}\right)(1-\tau_C). \quad (4)$$

Summarizing, if $\tau_R = \tau_L$, the investor is still indifferent between Roth and Traditional;

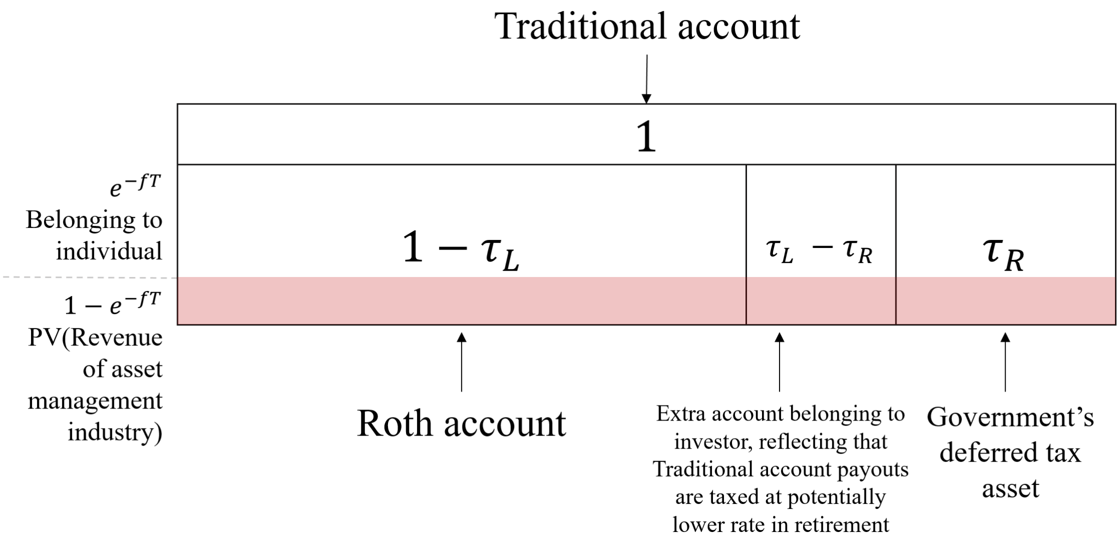

Figure 2: **Decomposition of Traditional tax-deferred account into three virtual accounts with asset management fees.** The decomposition is the same as in Fig. 1. All three virtual accounts (including the government's deferred tax asset) are eroded by asset management fees equally.

if $\tau_R \neq \tau_L$, the individual's relative preference for one or the other account does not change vis-à-vis the case with no fees, because the sign of the transfer depends only on $\tau_L - \tau_R$ regardless of fees. For the individual, both Roth and Traditional are eroded by fees in equal proportions. For the government, however, things are different. With a Roth account, the government receives the tax revenue upfront. With a Traditional account, the government receives the tax revenue when the money is paid out. This "consumption-based" recognition of income can seem natural, but as a consequence, a fraction τ_R of the Traditional account actually constitutes a virtual account owned by the government.

If the government were to receive the revenue upfront, like in a Roth, it would pay down some debt. By leaving the money in the account, the government keeps paying an interest rate r on the outstanding debt, but receives a net-of-fees return $r - f$. In other words, the government keeps money in a *virtual* account, which pays *real* fees. The fees are real because we assume that the same percentage fees are charged on a larger account size. Two things are left to clarify. First, how important are fees? Second, would fees really remain the same if the size of accounts varied? We try to answer these questions in the next two Sections.

All the results obtained so far 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. (forthcoming) provide evidence that individuals do not adjust their retirement savings in this way, but instead find that contributions under a Roth 401(k) are similar to those under a Traditional 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.[6]

4 Investment fees in retirement accounts

Tax-deferred retirement accounts include two main components: employer-sponsored defined contribution retirement accounts such as 401(k) and 403(b) plans (DC plans), and individual retirement accounts (IRAs). Most of the dollars in IRAs were initially accumulated in DC plans and later "rolled over" to an individual account. With roughly $7 trillion of assets each, DC plans and IRAs are both important components of overall retirement assets.

Assets in DC plans are invested in a menu of investment products chosen by the employer. The menu typically includes mutual funds and other pooled investment products. These products include funds that are not available to the general public, such as collective investment trusts (CITs), and stable value products, such as guaranteed investment certificates (GICs). Henceforth we refer to collective investment products as "funds". Where applicable, the menu often includes employer stock. Individual participants choose how to allocate contributions across funds in the menu. Individuals may also shift existing balances across funds.

Assets in IRAs are invested in a menu of investment products chosen by the account provider. This menu is usually much broader, including publicly available funds, products that are unique to the account provider, and individual securities. Account providers include brokerage firms, mutual fund sponsors, banks, and insurance companies.

An individual saving for retirement faces at least three types of costs: account fees, asset management fees, and implicit trading costs. Account fees cover the cost of account maintenance as well as any financial advice that comes with the account. In DC plans, account fees are usually assessed as a function of assets or of number of participants. On average, roughly one-tenth of these fees is covered by the employer, and the rest is charged directly to the participants (Rosshirt et al., 2014). In what follows we include employer-paid fees as part of total fees. In IRAs, account fees include fixed annual fees, transaction fees and sales loads (one-time fees charged upon buying or selling a product). We exclude fees

[6] Roth is more cost effective than Traditional. If the total amount of assets is constant under Roth and Traditional, 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 deliver the same savings subsidy for a lower cost to the government.

for voluntary additional services such as robo-advisors or premium human advisors.

Asset management fees are charged based on what financial products the account money is invested in. These fees cover the operating costs and profits of mutual funds sponsors and other providers. These asset management fees are hard to separate from account fees, because they often contain distribution-related fees (such as so-called 12b-1 fees, sub-transfer agent fees, shareholder servicing fees) that are rebated by the asset manager to the account provider.

Finally, implicit trading costs are incurred while buying and selling securities. Both funds and individuals trading on their own account incur bid-ask spreads, defined as the difference between the buy price and the sell price. For funds, implicit trading costs also include market impact (i.e. adverse price moves caused by the fund's trades) and even trading commissions, which are explicit for the fund, but are not included in the expense ratio and therefore are implicit costs for the individual. Investors in DC plans are sometimes able to access a brokerage window to buy and sell individual securities outside of the employer-chosen menu, but funds make up a large majority of total assets. Investors in IRAs are more likely to invest in individual securities, but they may also have higher bid-ask spreads and transaction fees.

4.1 Account fees and asset management fees

We choose 50 bps as our estimate of total account costs and asset management costs for tax-deferred retirement savings. Account fees and asset management fees are difficult to separate because the providers of funds and other financial products often rebate some of the fees paid to the account provider. As a result, two recent industry publications have attempted to provide an asset-weighted estimate of overall fees for 401(k) accounts, a large fraction of overall DC plans. Deloitte (Rosshirt et al., 2014) estimates the "all-in fee" for 401(k) accounts at 58 bps of assets under management. BrightScope (Brightscope and Investment Company Institute, 2014) estimates "total plan costs" at 39 bps. Both estimates are done in partnership with the industry trade association, the Investment Company Institute. The BrightScope estimate is based on filings by audited plans, which generally means plans with 100 or more participants. The result is to exclude about $1 trillion or 27% of total assets held in the smallest (and likely most expensive) plans. 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 universe of plans filing Form 5500 with the Department of Labor. Our 50 bps estimate is a rough average of these two estimates.

These estimates imply that fees on funds held in DC plans are substantially lower than the industry average. However, the fees charged by mutual funds in defined contributions accounts have been the object of scrutiny recently, as plan managers appear to choose sub-optimal menus of funds (Ayres and Curtis, 2015; Pool et al., Forthcoming). This contributes to increase investment costs for plan participants and, in the case of Traditional accounts, reduce future government revenues.

Fees incurred in IRA accounts are likely to be similar, although there are no equivalent published estimates of overall account costs. On the one hand, the fees on mutual funds held in IRAs are likely closer to the asset-weighted industry average of 67 bps (Collins et al., 2016). On the other hand, a different asset composition is likely to reduce overall fees incurred. About half of IRA assets are held in mutual funds (ICI, 2015). Roughly a third of IRA assets are held in nonfund money market products and individual securities, resulting in low or no fees. Finally, other investment products constitute an important share of total IRA assets, but it is not easy to assess the level of fees on these other products. Often, their payoff is not directly linked 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 literature on shrouded prices (Gabaix and Laibson, 2006; Carlin, 2009; Henderson and Pearson, 2011) is unanimous in reporting this finding.

4.2 Implicit trading costs

Our estimate of implicit trading costs is 30 bps. Asset management fees do not include the trading costs incurred by mutual funds and other funds. Some of these costs take the form of explicit trading commissions (on which information is often available, although not included in mutual funds' net expense ratio) but others are altogether invisible, as they take the form of bid/ask spreads and adverse market impact of trades, and are therefore implicit in the prices that mutual funds pay and receive for securities when they trade. While index funds do not trade much, active funds have considerable asset turnover. An asset management startup estimates that equal-weighted average portfolio commissions alone are in the order of 20 bps (Wealthfront, 2016). Livingston and Zhou (2015) find a very similar number (18bps). The effect of bid/ask spreads and adverse market impact is even greater: because of their unique disclosure requirements and liquidity needs, mutual funds' trades are more predictable than those of other investors; mutual funds are "sitting ducks" liable to be front-run (Shive and Yun, 2013) and to trade against short-sellers (Arif et al., 2015). Wermers (2000) estimates that commissions, transaction costs and cash drag due to liquidity cause a 230 bps wedge between the average equity mutual fund's returns and the return of the stocks they hold. Edelen et al. (2013) estimate average total trading costs of 144 bps using a sample of over

3,000 U.S. domestic equity funds. Although this estimate is equal-weighted, the authors also report an estimate of the expense ratio in the same sample (119 bps) suggesting that implicit costs may be larger than explicit costs.

The lowest estimate in the literature is Bogle (2014), who estimates the overall impact of commissions and market impact around 50 bps for active equity funds, and next to nothing for passive equity funds. According to Morningstar's Fund Flows (Jan 2016), about 70% of assets are held in active funds. Combining these two figures, asset-weighted average trading costs for equity funds are likely to be close to 35 bps. This figure is likely to be closer to 25 bps for bond funds and zero for money market funds.[7] Based on the overall asset allocation in DC plans (Collins et al., 2016) and IRAs (Copeland, 2016), we estimate total implicit trading costs as 30 bps, the weighted average of these three asset classes, a number that reflects a large asset allocation to equity funds.

4.3 The value of services received in exchange for fees

Investors receive services in exchange for fees. In a Traditional account, the government owns a virtual account that pays the same fees. If the government obtains any benefit in exchange for the fees paid on the virtual account, these benefits should be subtracted from the fee to arrive at a net cost. However, the government does not benefit from most services provided by the asset management industry.

We discussed above that if the government's investment portfolio is constrained, Traditional accounts may well increase the government's total exposure to stocks. If the increase is welfare-improving, the government may therefore benefit from basic, passive portfolio management services. However, the cost of these services is negligible compared to the magnitude of fees. Collins (2005) shows that portfolio management services for S&P 500 index funds are essentially a commodity, with costs between 1 and 5 bps and average of 3 bps across all reporting funds. Similarly, the U.S. federal government's Thrift Savings Plan (TSP) outsources basic index funds at a cost of roughly 3 bps. Any additional fees must be charged for services that are provided in addition to basic portfolio management.

[7] We are not aware of any published estimate of asset-weighted average trading costs for bond funds. However, Bessembinder et al. (2016) estimate that transaction costs on the largest corporate bond trades are roughly 0.20% of trading volume. The asset-weighted average portfolio turnover of bond funds is between 90% and 193%, depending on fund type (Rowley and Dickson, 2012). We use a rough average of 125%. Multiplying these two numbers together, we obtain 25bps as a lower bound estimate of annual transaction costs incurred as a percentage of assets. For reference, Malkiel (2013) estimates that the average fund underperforms the reference index by 82bps, suggesting that this magnitude is reasonable.

4.3.1 Advice

Collins (2005) proposes that part of the large dispersion in the actual fees of S&P 500 index funds may be attributable of different levels of financial advice bundled with the basic fund management services. For instance, a mutual fund's net expense ratio may include so-called 12b-1 fees. These fees, together with any front- or back-loads, are rebated to the account provider and may help pay for services that benefits the investor, such as administrative costs as well as any complimentary advice that the account provider may offer. Compared to an investor with a Roth account, an investor with a Traditional account pays higher dollar fees because the same percentage fee is applied to a larger amount of assets; the difference goes to erode the government's virtual account. This account does not necessitate of any advice, and therefore the government does not benefit from bundled financial advice.

4.3.2 Asset allocation

Higher fees may be caused by costs associated with creating funds or asset allocations that are customized to a particular group of individuals. For instance, some funds may focus on styles (conservative/aggressive, value/growth), maturity (long/short), sector (small cap/large cap, junk/investment grade) or industry. These funds cater to individuals with particular preferences (e.g., low or high risk tolerance, preference for skewness, etc.), or beliefs (e.g., that the health care industry is about to experience massive growth), or personal situations (e.g., health care industry workers who would like to have no health care stocks in their portfolio). Although individual investors may experience real benefits from holding these funds, these benefits largely cancel out in aggregate, because the government holds a fraction of all these funds. Target date funds are a particularly fitting example. Target funds adjust asset allocations as individuals age and get closer to the target retirement date. Clearly, holding target date funds of *all* target dates does not create value to the government.

In general, because individuals' allocations to specialized funds largely cancel out in aggregate, the fees paid to obtain these allocations in the government's virtual accounts entail zero benefit for the government. However, 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. If this departure from the market portfolio exists and is optimal for the government, even this modest benefit could probably be obtained at a much lower expense by running a sovereign wealth fund.

	Performance (bps)		
Source	Net	Gross	Benchmark
Berk and van Binsbergen (2015)	-12 *		Investable Vanguard funds
Malkiel (2013)	-64		Large cap active vs. SP500 Index
Fama and French (2010)	-100	-5	3- and 4-factor benchmarks
Wermers (2000)	-100	130	Own stock holdings
Carhart (1997)	-154 per 100 **		1-, 3- and 4-factor benchmarks
Jensen (1968)	-40	~0	1-factor benchmark (CAPM)
Fama (1965)	-60	+20	Market

Table 4: **Estimates of average equity mutual fund underperformance.** "Net" and "Gross" refers to expenses. The definition of "expenses" is typically the expense ratio, but in the case of Wermers (2000) it includes everything including cash drag and trading costs (see text) — Footnotes: [*] Underperformance with respect to the Vanguard benchmark, which charges fees of 18bps [**] 100 bps of expense ratio are associated with underperformance of 154 bps.

4.3.3 Fees and alpha

Actively managed funds have significantly higher fees than passive index funds. However, it is possible that actively managed funds also have higher expected returns. On the one hand, skilled managers may be able to generate enough excess returns that even after paying higher fees shareholders are able to come out ahead. On the other hand, active management is at least in part a zero-sum game. Even if some mutual funds show evidence of excess returns, at least some of the gain comes at the expense of other mutual funds. As Fama and French (2010) point out, "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". Unless funds held in tax-deferred accounts are systematically winning the zero-sum game against funds in nonretirement accounts or against non-funds, this argument constitutes a further reason not to subsidize the asset management industry.

Measuring mutual fund performance is difficult. First, actual performance net of the benchmark has a large random component, and a reliable estimate of performance requires a long time series. Second, unlike direct estimates of fees, every benchmark-based estimate implies and depends on an asset-pricing model. As a result, the literature on mutual fund performance contains numerous estimates done using different methodologies and benchmarks, a few of which are summarized Table 4.

The literature begins with classics such as Fama (1965) and Jensen (1968). Both studies show no evidence of managers predictably beating the market; on average, mutual funds show a small underperformance with respect to the market benchmark; consistent with market efficiency, this underperformance is of the same magnitude of fees and cash drag. More recently, Carhart (1997) compiles a mutual fund database that is comprehensive and free of survivorship bias,[8] and uses it to replicate the basic result that there is no evidence of skilled or informed mutual fund managers. Using four-factor and three-factor benchmarks, Carhart finds that there is return predictability that is not explained by fees, but only for the worst-performing funds. He estimates that 100bps of expense ratio are associated with a 154bps underperformance with respect to the market. Wermers (2000) decomposes mutual fund returns into stock-picking talent, style, transaction costs and expenses, concluding that mutual funds hold stocks that beat the market by 1.3%, but the funds' returns underperform the market by 1%. He attributes the large discrepancy to cash drag (0.7%) and expenses and transaction costs (1.6%). Based on four-factor and three-factor benchmarks, Fama and French (2010) estimate net-of-fees underperformance of about 1% per year. Malkiel (2013) compares several categories of funds with their indices, finding that that large-cap equity funds underperform the S&P 500 Index by 64 bps, and bond funds underperform the Barclay US Aggregate Bond Index by about 84 bps.

Some recent studies have focused on investable benchmarks. French (2008) estimates a broad measure of the annual cost of *active* management, including not only costs faced by individual investors but also costs faced by institutions and market-making gains by financial intermediaries over 1980-2006. The cost of active management is 0.67% of the aggregate value of the market, in addition to approximately 0.10% cost of passive management. As a passive benchmark, French uses the Vanguard Total Stock Market Index. Berk and van Binsbergen (2015) compare active funds' dollar returns (as opposed to percent returns) against the relevant Vanguard benchmarks. They estimate a value weighted net alpha of -12 bps (not statistically different from zero) in addition to the cost of investing in the Vanguard benchmark (18 bps), implying a total cost of active money management of about 30 bps.

4.4 Calibration: excess investment costs under a Traditional scheme

Based on asset management and account fees of 50 bps, implicit trading costs of 30 bps, and zero benefit, a conservative, asset-weighted estimate of "all-in" average fees is about 80 bps, or 0.8 percentage points. A simple estimate of the ongoing flow of excess investment fees is obtained by multiplying the size of the government's virtual account times the net fees paid

[8] Malkiel (1995) also addresses survivorship bias and extends the sample period of previous studies which claimed to find persistence in returns. Carhartt also addresses those studies, explaining their findings as the result of momentum investing.

on the account in one year. The government's virtual account (\$2.8 billion) is obtained by multiplying the total amount of assets in DC plans and IRAs (S = \$13.9 trillion) times a reasonable tax rate on retirement income ($\tau_R = 20\%$). Net fees are equal to fees estimated in this section (f = 80 bps) multiplied by one minus the corporate tax rate ($\tau_C = 35$), to account for the revenue recovered through taxation of the asset managers (under the conservative assumption that a dollar of fee revenue translates to a dollar of taxable profits for the asset manager):

$$\text{Excess investment fees} = S \cdot \tau_R \cdot f \cdot (1 - \tau_C) =$$

$$= \$13.9 \text{ trillion} \times 20\% \times 0.8\% \times (1 - 35\%) = \$14.5 \text{ billion.}$$

Our estimate of assets under management is also conservative, as it ignores another \$6.5 trillion of tax-deferred assets in state and local government and corporate defined-benefit pension plans. Accounting for these assets would increase our estimate of the size of the government's virtual account by another \$1.3 trillion. 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 not when the employee becomes entitled to them, but when they are actually 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 incur lower asset management costs than defined contribution plans or IRAs, they still incur a positive cost of managing the assets held in the government's virtual account.[9]

In Table 5, we carry out the same back-of-the envelope calculation for selected countries. The U.S. has the world's largest retirement assets, and therefore leads the list. However, other countries have substantial amounts of tax-deferred retirement assets and fees that are significantly higher than those paid by U.S. investors. Thus, the implicit subsidy is still very large, and often larger than the U.S. as a percentage of GDP.

5 A general equilibrium model of retirement savings

In this section we examine the conditions under which a larger amount of retirement assets would truly result in a larger amount of resources dedicated to asset management. Consider first the case in which there are no fixed costs and all asset management costs are a linear function the amount of assets under management. Under the assumption of proportional

[9] In a future draft of this paper, we plan to produce our own estimate of asset-weighted average costs incurred by defined contribution and defined benefit plan participants, as detailed in Appendix A.

Country	Ret. Assets (US$ Billion)	Income Tax Rates Personal	Corporate	Fees	Subsidy (US$ Billion)
United States	13,900	20%	35%	0.80%	14.5
Canada	2,462	15%	15%	2.10%	6.5
U.K.	2,684	14%	20%	1.40%	4.2
Japan	1,221	15%	23%	1.47%	2.0
Netherlands	1,282	15%	25%	1.40%	2.0
Switzerland	788	19%	9%	1.30%	1.8
Germany	236	23%	16%	1.70%	0.8
South Africa	306	15%	20%	1.10%	0.4
Italy	163	20%	28%	1.59%	0.4
China	231	15%	20%	1.00%	0.3

Table 5: **International estimates of subsidy to the asset management industry in countries with back-loaded taxation.** 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. Tax rates are calculated as the average tax rate faced by a person earning the average wage with no other income. Where unavailable, fees were set to 1%, and asset allocation and tax rates were set to the international average. Sources of non-U.S. values: OECD (tax rates and retirement assets), Morningstar and others (fees).

costs, there is no need to solve an equilibrium model. Upon a switch from Roth to Traditional, the government postpones the receipt of an amount B of revenue until later. This revenue shortfall is covered by issuing bonds by an amount B. Thus, aggregate retirement savings and aggregate assets increase (compared to what they were under Roth) by exactly the same amount B, causing aggregate asset management costs to increase. An increase in assets under management directly causes an increase in real resources devoted to asset management. In addition, if investors' demand for the asset management services of a given firm is not perfectly responsive to price, the switch from Roth to Traditional would further reduce investors' sensitivity to fees, as the same dollar fees become a smaller percentage of asset, with the end result of increasing asset managers' equilibrium profits.

Of course, the assumption that asset management costs are proportional to assets under management is extreme. The asset management business is likely subject to substantial economies of scale. Therefore, consider the other extreme in which all asset management costs are fixed costs and none are variable. In this case, a switch from Roth to Traditional can still indirectly bring about an increase in resources devoted to asset management by increasing the number of firms that are viable in equilibrium, and with it, the total number of times fixed costs are incurred. If consumers are not perfectly sensitive to prices, a switch from Roth to Traditional and the associated increase in assets likely enables asset managers to charge higher dollar fees. If entry is free, higher fees cause an increase in the number of firms, as new competitors enter the market until equilibrium profits are zero again. If no new firms can enter, on the other hand, then the total resources devoted to asset management are constant by assumption. Even in this case, however, existing firms could charge higher dollar fees, enjoying higher profits and a larger government subsidy. Thus, when all asset management costs are fixed costs, upon a switch from Roth to Traditional the subsidy increases, and the total amount of resources devoted to asset management is likely to increase too.

Finally, suppose that all costs are per-participant. Clearly, a switch from Roth to Traditional would not increase resources used, because the number of investors remains constant. However, unless investors are perfectly sensitive to prices, dollar fees would likely increase because, once again, investors' sensitivity to dollar fees is decreased. Thus, the only case in which an increase in assets does not cause an increase in the transfer to the asset management industry is a case in which all costs are per-participant, *and* investors are perfectly sensitive to prices.

Summarizing, intuition suggests that 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. To be more precise, we need to take a stand on which of the above assumptions most resembles the actual asset management industry. Existing literature suggests that entry is essentially free, investors are not perfectly sensitive to price, and there are substantial economies of scale.

5.1 Empirical evidence on market structure, cost structure and competition

The least controversial finding is the lack of barriers to entry or expansion (Hubbard et al., 2010; Baumol et al., 1990). In 2014 alone, 654 new mutual funds and 69 new mutual fund sponsors entered the industry, for a net increase of 292 funds and 25 fund sponsors (ICI, 2015). 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.

Evidence of economies of scale on the cost side is presented by several studies (Baumol et al., 1990; Latzko, 1999; Coates and Hubbard, 2007; Dyck and Pomorski, 2011). Intuition suggests that the asset management business should have a strong fixed-cost component. For instance, Gao and Livingston (2008) find that the "paperwork" part of mutual funds expenses (custodian fees, recordkeeping fees, etc.) does not vary meaningfully with fund size. Statements by industry insiders also confirm this intuition: Jeffrey S. Molitor, the director for portfolio review at Vanguard, is quoted in Kahn (2002) as stating that the "marginal cost of managing increasing dollars is minimal." This statement refers specifically to *active* funds whose management Vanguard outsources to subadvisers; for passive funds, the economies of scale are obvious.

However, costs are not fees. Lacking perfect competition, firms can charge more than their marginal cost and therefore fees will not necessarily drop as costs do. Malkiel (2013) notes that "academic research has documented substantial economies of scale in mutual fund administration", but between 1980 and 2010, in spite of a more than 100-fold increase in assets under management, percentage fees stayed relatively flat around 70-80 bps. In particular, the fees charged by active managers rose from 66 to 91 bps. Similarly, Philippon (2015) argues that in aggregate the unit cost of financial intermediation (defined as the ratio of the income of financial intermediaries to the quantity of intermediated assets) is very stable in the long run and it has not dropped over the last century in spite of a large increase in total assets under management—both in absolute terms and as percent of output. The fact alone that fees are charged as a percentage of assets under management (as opposed to absolute dollar prices) suggests that either there is some variable cost component, or that clients with more assets have lower price sensitivity and asset managers are engaging in price discrimination.[10]

The available empirical evidence supports both explanations. On the one hand, although all empirical studies of costs support substantial economies of scale, all studies also find that costs do increase as assets increase (both overall assets, and assets per account). On the other hand, there is also abundant evidence of investors' imperfect sensitivity to price. Although funds with lower fees tend to have higher market shares (Hubbard et al., 2010), many studies point out the continued existence of dominated funds; for instance, Hortacsu and Syverson (2004) document the existence of 82 distinct S&P 500 Index funds with large dispersion in fees (an interquartile range of 89 bps). We update their analysis using 2015 data and find that the large dispersion still persists. Hortacsu and Syverson make sense of this phenomenon by

[10] Costs are not fees, and fees are not net-of-fees performance. In Section 4 we discussed the relationship between the level of fees and the level of performance. However, theoretical literature suggests that performance could be subject to diseconomies of scale (Berk and Green 2004; Pástor and Stambaugh 2012). The empirical evidence is mixed. Berk and van Binsbergen (2015) find in favor of this, but Reuter and Zitzewitz (2015) present evidence of insignificant changes in performance using a change in Morningstar rating as a shock to assets under management. Dyck and Pomorski (2011) find that large pension plans are able to obtain superior returns through increased access to alternative investments. Pástor et al. (2015) show strong support of the industry-level decreasing returns to scale hypothesis (Pástor and Stambaugh 2012): as the size of the active mutual fund industry increases, the ability of any given fund to outperform declines. For simplicity, we abstract from performance-related considerations, a conservative assumption.

assuming the existence of search costs. Gil-Bazo and Ruiz-Verdú (2009) find evidence that "underperforming funds and funds faced with less performance-sensitive investors charge higher marketing and nonmarketing fees," as in a theory initially proposed proposed by Christoffersen and Musto (2002). Bergstresser et al. (2009) find that broker-sold funds are more costly *and* underperform, implying that the broker channel enables the survival of otherwise dominated funds. A report by the Executive Office of the President of the United States (2015) summarizes the academic literature on investment advice, finding that fund distribution channels are able to charge one or two percentage points from investors who seek for investment advice. Finally, Gârleanu and Pedersen (2015) features a financial market where finding information is costly (a la Grossman and Stiglitz, 1980), with the additional feature that finding money managers is also costly. Agents can pay a cost and allocate money to a well-managed fund, or decide to stay uninformed and become "noise allocators". Only by introducing the concept of noise allocators are the authors able to reproduce several otherwise puzzling empirical facts, such as the existence of a sizable minority of mutual fund managers who can consistently pick stocks well enough to cover their costs.

5.2 A model

In order to formalize the intuition laid out at the beginning of this section, we examine a general equilibrium model of the asset management industry. In this model we make the most conservative assumption possible that are compatible with the empirical evidence summarized above. First, we assume that the asset management business is essentially a fixed-cost business. Specifically, we assume that each firm needs a fixed amount of labor to operate. Second, we assume no barriers to entry at all. Fixed costs and free entry imply that the equilibrium profits of this industry are zero. Third and last, we assume that individuals' demand for the services of a given firm is not perfectly elastic.

In our simple model we ignore the existence of multiple layers of financial intermediation. The unit of production of asset management services is the "mutual fund", and individuals give their savings directly to mutual funds who charge explicit fees. For this reason, we use the words "fund" and "firm" interchangeably. Funds do not rebate any of the fee revenue to distribution channels, and do not incur any trading costs.

In equilibrium, funds face a downward-sloping demand function, i.e. if they raise their fees, their demand falls, but it does not fall to zero. Although the existence of dominated or duplicate funds and other evidence point to outright inertia as the cause for this low sensitivity, we take a more optimistic (that is, conservative) view, and we assume that individuals have preferences for convenience; more funds means that the distance between an individual and their chosen fund is on average lower, i.e., utility is higher. A low "distance" can be thought of as literally low physical distance from the nearest branch, but also trust (Gennaioli et al., 2015), or ease of finding, convenience (by having one's choice of funds as part of a restricted set all in one place, as it is the case for many tax-deferred retirement plans) or even a preference for non-portfolio characteristics of funds, such as the level of customer service.

In sum, in this model, mutual funds have only fixed costs, they create value, and they face competition that is imperfect but stiff enough to warrant zero equilibrium profits. In spite of that, we aim to (a) show that under a reasonable calibration there will be too many funds in equilibrium, and (b) using comparative statics, show that a switch from Roth to Traditional causes the number of funds to become even higher.

The model features a two-period economy. When individuals are young (time t), they work, produce, consume, and save for retirement by investing via mutual funds in a mix of government bonds and financial assets. When they are old (time t'), they receive passive retirement income, consume, and die without bequest. Mutual funds require a fixed amount of labor to operate, and therefore must pay a competitive salary to individuals who otherwise would work to produce goods.

In this section we present assumptions and the main results. A full discussion of the model is provided in Appendix C.

5.3 Individuals

Individual $i \in \mathcal{I}$ lives two periods and leaves no bequest. In period t the individual works, saves, and allocates the savings. The individual starts with a net worth of 0 and is endowed with one unit of labor to spend either producing consumption goods, or managing a mutual fund. The production technology is linear: if the individual allocates L units of labor to the production of goods, the output is $F(L) = \omega L$. In order to attract labor, therefore, mutual funds must pay a competitive wage ω. The individual's total supply of labor is inelastic. Overall, the individual's pretax income is ω regardless of how labor is allocated. Income from work is taxed at rate τ_L. In period t', the individual retires and depletes all the savings.

The individual draws utility from current consumption (C) and from discounted future consumption ($\delta \cdot C'$). To finance future consumption, the individual saves and invests an amount S. All investment must be carried out via a chosen mutual fund j charging a proportional fee $f_j S$. The individual derives disutility from the distance between the location of chosen fund j (ι_j) and one's own location ($d_{i,j} \equiv |\iota_j - \iota_i|$). A fraction a of the individual's savings is invested in an exogenous storage technology yielding a return ρ, and the remainder in government bonds paying a return r. The government grants the individual a deduction for savings at a rate τ_S. Thus, the utility of individual i is:

$$U_i = \max_{C,C',S,a,j} \ln C + \delta \ln C' - \gamma d_{i,j} \tag{5}$$

subject to the budget constraints:

$$C = \omega \left(1 - \tau_L\right) - S \left(1 - \tau_S\right), \tag{6}$$

$$C' = S \left(1 - f_j\right) \left[1 + a\rho + \left(1 - a\right) r\right] \left(1 - \tau_R\right). \tag{7}$$

5.4 Mutual funds

Mutual funds can differentiate themselves over one qualitative characteristic, ι, defined on the $[0, 1)$ circle. Individuals are uniformly distributed over this circle, and their utility is decreasing in the distance from their chosen fund. Every fund $j \in \mathcal{J}$ needs a fixed amount of labor φ just to be able to operate. Fund profits are equal to revenue minus cost:

$$\pi_j = f_j Q_j - \varphi\omega,$$

where Q_j is the fund's assets under management, and f_j are the percent fees the fund charges, so that $f_j Q_j$ is the fund's total revenue. The fund's problem is to choose f_j and ι_j to maximize π_j, taking into account that Q_j depends on the fund's choices of pricing (f_j) and location (ι_j), and taking competitors' choices as given.

Mutual fund profits accrue to their managers and are taxed as other income. However, because of free entry, profits will be zero, so we do not keep track of who the profit accrues to.

5.5 Government

The government spends an exogenously given amount G. This expenditure is inevitable and it does not affect the utility of agents. At time t, the government taxes income at a rate τ_L, and grants individuals a deduction for savings at a rate τ_S. The government can also borrow an amount B at the market interest rate r to satisfy the government budget constraint:

$$G = \tau_L \omega - S\tau_S + B$$

To pay off the bonds at time t', the government can tax retirement income at a rate τ_R so as to satisfy

$$B\left(1+r\right)=\tau_R S\left(1-f\right)\left[1+a\rho+\left(1-a\right)r\right]$$

where f is the equilibrium level of fees.

The government takes τ_L and τ_R as given. For τ_S, only two policy options are on the table: Traditional retirement accounts ($\tau_S=\tau_L$), and Roth accounts ($\tau_S=0$). Once chosen τ_S, the government chooses B to balance the budget constraint.

5.6 Market equilibrium

The model has one symmetric equilibrium in which N funds distribute themselves over the circle at equal distance from one another and set equal fees (Tirole, 1988). Individuals' equilibrium savings are

$$S^{*}=\omega\frac{1-\tau_L}{1-\tau_S}\frac{\delta}{1+\delta}. \tag{8}$$

Obviously, aggregate savings are higher with a Traditional scheme than they are with a Roth scheme because in the case of Traditional retirement accounts, $\tau_S=\tau_L$, and in the case of Roth accounts, $\tau_S=0$.

The equilibrium number of funds is defined implicitly by the following quadratic equation:

$$\frac{\delta}{\gamma}N^{2}+N=\frac{S}{\varphi\omega}. \tag{9}$$

The explicit solution for N^{*} is ugly, but without solving explicitly, it is evident that N is an increasing function of S. Therefore, the number of funds does increase under a Traditional scheme as compared to a Roth scheme.

Finally, the equilibrium level of fees is a decreasing function of N, i.e., higher competition does translate to lower percent fees:

$$f_j^{*}=f^{*}=\frac{1}{1+\frac{\delta}{\gamma}N}. \tag{10}$$

It is easy to show that, as S increases because of a switch from Roth to Traditional, f does not drop proportionally, so that total fees fS increases, supporting a larger number of funds.

5.7 Planner solution

We compare the market equilibrium with the solution chosen by a benevolent planner who also takes G as exogenously given. The planner chooses savings S, and number of funds N directly to maximize social utility:

$$U = \max_{C,C',S,N} \ln(C) + \delta \ln(C') - \gamma \bar{d}, \tag{11}$$

where $\bar{d}$ indicates the *average* distance between an investor and their fund. The planner's budget constraints are simply

$$C = \omega(1 - \varphi N) - S - G, \tag{12}$$

$$C' = S(1+\rho) = (\omega(1 - \varphi N) - C - G)(1+\rho). \tag{13}$$

Under these assumptions, we obtain an implicit expression for the socially optimal number of funds:

$$4\frac{1+\delta}{\gamma}N^2 + N = \frac{1 - G/\omega}{\varphi}. \tag{14}$$

5.8 Calibration

Next, we turn to the question whether a larger asset management industry (a larger N) is optimal. Consider a Roth-based system as the starting point. If that were the social optimum, a shift to Traditional would imply that there are too many funds. Because we are departing from the optimum, however, the total loss need not be too great, as the social welfare function is not steep at the optimum. Only if such a shift caused a large change in the equilibrium number of funds would social welfare suffer serious consequences. On the other hand, if a Roth-based system were to produce too few funds, then a shift to Traditional would get society closer to the optimum, and it would be very beneficial. Finally, if a Roth-based system were to produce too many funds, the shift to Traditional would be particularly pernicious, as it would bring society even farther from the optimum, starting at a point where the social welfare function is already steep. Thus, if the number of funds under the market solution (N_M) is higher than the number of funds under the planner solution (N_P), Roth is to be preferred because with Traditional N_M would be even higher. Vice versa, if $N_M < N_P$, Traditional would be preferred.

It is important to note that in this model we give the asset management industry the benefit of the doubt, because we assume that every additional fund improves the utility of the average individual. It would have been also possible, consistent with the prevailing

Parameter	Description	Roth	Traditional
δ	30-year discount factor	0.545	0.545
γ	Preference for funds	3500	3500
$\tau = G/\omega$	Tax rate/Government Expenditure	0.18	0.18
τ_S	Subsidy on savings	0	0.18
φ	Labor fraction of one fund	0.000017%	0.000017%
φN_P	Optimal resources employed in asset management	0.87%	
φN_M	Actual resources employed in asset management	1.70%	1.88%
f	30-year level of fees	5.882%	5.341%

Table 6: Calibration of the general equilibrium model

empirical evidence, to write a model with captive demand (Gil-Bazo and Ruiz-Verdú, 2008, 2009; Bergstresser et al., 2009; Gennaioli et al., 2015) and shrouded fees (Gabaix and Laibson, 2006; Carlin, 2009; Henderson and Pearson, 2011). 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.

To simplify the calibration, assume that the tax rate is set to achieve a balanced budget in the absence of tax subsidies: $\tau_L = G/\omega$ where G is the exogenous expenditure and ω is the output, so the tax rate is simply the ratio of government expenditure / output. Thus, we can rewrite the implicit expression for N_M as:

$$\frac{1+\delta}{\gamma}N_M^2 + \frac{1+\delta}{\delta}N_M = \frac{1}{1-\tau_S}\cdot\frac{1-G/\omega}{\varphi}. \tag{15}$$

This new expression is very similar to the implicit expression for N_P, the planner solution:

$$4\frac{1+\delta}{\gamma}N_P^2 + N_P = \frac{1-G/\omega}{\varphi} \tag{16}$$

Under a Roth scheme, $\tau_S = 0$ and therefore the right hand side of the two expressions is the same. Because of the "4" coefficient on the quadratic term, $N_P \approx N_M/2$, and therefore a Roth should be preferred under most parameterizations. However, for very impatient investors ($\delta \approx 0.01$), the linear term prevails, obtaining the opposite result. The question, therefore, is whether under a reasonable calibration we obtain the result we anticipate.

Table 6 shows two possible calibrations, with Roth ($\tau_S = 0$) and with Traditional ($\tau_S = \tau$). The details of the calibration are explained in Appendix D. The table shows that the excess resources dedicated to the asset management industry in a Traditional scenario as compared to a Roth scenario are substantial: $\varphi\left(N_M^{Trad} - N_M^{Roth}\right) = 0.18\%$. Multiplying this figure by annual total output ($18 trillion), the amount of excess resources dedicated to the asset management industry is about 32.3 billion dollars per year.

6 Conclusion

In this paper we show that back-loading the taxation of retirement savings results in the government paying asset management fees on its large implicit portfolio. Relative to a front-loaded system, this results in a larger asset management industry. Under a rough calibration of a general equilibrium model, we show that back-loading taxation inefficiently increases the amount of resources spent on asset management, thus reducing welfare. Our model focuses only on this one aspect—fees—abstracting from other potential drivers of the choice between front-loaded and back-loaded taxation of retirement savings. A back-loaded taxation scheme has a few unique features that a front-loaded scheme cannot reproduce. First, in a progressive tax system, a back-loaded scheme creates a fairer tax burden for those who do not earn income every year of their life. Second, in a back-loaded scheme there is no gain from contributing underpriced assets as there is in a front-loaded scheme. Finally, from a behavioral perspective, under a progressive income tax system an EET system contains a powerful tax incentive to spend down the account gradually when the individual becomes eligible for withdrawals.[11] A front-loaded scheme also has unique benefits in addition to the elimination of wasteful fees. First, it is simple and it eliminates all uncertainty about future taxes and future government revenue. Second, front-loaded taxation seems to provide a more powerful behavioral incentive to save, as it "cheats" investors into saving greater amounts (Beshears et al., forthcoming).

Our results have implications for public policies related to retirement saving. Policy options include mandating the use Roth accounts for new contributions, converting existing tax-deferred accounts to Roth, and explicitly segregating the virtual government account (with a mechanism similar to withholdings) so that the government is able to negotiate lower fees. In addition, since the government is paying investment management fees on its implicit portion of retirement accounts, our results suggest a rationale for the government to support policies that might reduce asset management fees, at least those fees that do not correspond to added value. One potential policy of this sort is the requirement that retirement advisers abide by a fiduciary standard, as established by the new U.S. Department of Labor fiduciary

[11] This was one of the arguments made in a recent debate in the United Kingdom over the optimal tax incentive scheme for retirement savings. For instance, an Economist editorial (Buttonwood, 2015) explains that "This tax charge is actually quite useful in that it stops people blowing their pension pot in a spending spree at 65."

rule set to go into effect by April 2017. Not only is the welfare of retirees at stake, but in addition to that, the government's revenue as well.

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 (1), 128–148.

Arif, S., Ben-Rephael, A., Lee, C. M., 2015. Do short-sellers profit from mutual funds? evidence from daily trades, stanford University Graduate School of Business Research Paper No. 14-35, available at http://ssrn.com/abstract=2496990.

Ayres, I., Curtis, Q., 2015. Beyond diversification: The pervasive problem of excessive fees and 'dominated funds' in 401(k) plans. Yale Law Journal 124 (5), 1346–1835.

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.

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 (10), 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 (6), 1269–1295.

Beshears, J., Choi, J. J., Laibson, D., Madrian, B. C., forthcoming. Does front-loading taxation increase savings? evidence from roth 401(k) introductions. Journal of Public Economics.

Bessembinder, H., Jacobsen, S., Maxwell, W., Venkataraman, K., 2016. Capital commitment and illiquidity in corporate bonds.

Bogle, J. C., 2014. The arithmetic of 'all-in' investment expenses. Financial Analysts Journal 70 (1), 13–21.

Bolton, P., Santos, T., Scheinkman, J. A., 2016. Cream-skimming in financial markets. Journal of Finance 71 (2), 709–736.

Brady, P., 2013. The tax benefits and revenue costs of tax deferral. Tech. rep., Investment Company Institute, Washington, DC.

Brightscope, Investment Company Institute, 2014. The BrightScope/ICI defined contribution plan profile: A close look at 401(k) plans. Tech. rep.

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 (3), 689–702.

Buttonwood, 5 Aug. 2015. EET your TEE, George. The Economist.

Carhart, M. M., 1997. On persistence in mutual fund performance. Journal of Finance 52 (1), 57–82.

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

Christoffersen, S. E., Musto, D. K., 2002. Demand curves and the pricing of money management. Review of Financial Studies 15 (5), 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., 2005. Are S&P 500 index mutual funds commodities? Investment Company Institute Perspective 11 (3).

Collins, S., Holden, S., Duvall, J., Chism, E. B., 2016. The economics of 401(k) plans: Service, fees and expenses, 2015. ICI Research Perspective 22 (4).

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

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

Diamond, P., Geanakoplos, J., 2003. Social security investment in equities. American Economic Review 93 (4), 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.

Edelen, R., Evans, R., Kadlec, G., 2013. Shedding light on "invisible" costs: Trading costs and mutual fund performance. Financial Analysts Journal 69 (1), 33–44.

Executive Office of the President of the United States, 2015. The effects of conflicted investment advice on retirement savings. Tech. rep.

Fama, E. F., 1965. The behavior of stock market prices. Journal of Business 38, 34–105.

Fama, E. F., French, K. R., 2010. Luck versus skill in the cross-section of mutual fund returns. The Journal of Finance 65 (5), 1915–1947.

Freeman, J. P., Brown, S. L., Pomerantz, S., 2008. Mutual fund advisory fees: new evidene and a fair fiduciary duty test. Oklahoma Law Review 61, 83–153.

French, K. R., 2008. Presidential address: The cost of investing. The Journal of Finance 53 (4), 1537–1573.

Gabaix, X., Laibson, D., 2006. Shrouded attributes, consumer myopia, and information suppression in competitive markets. Quarterly Journal of Economics 121 (2), 505–540.

Gao, X., Livingston, M., 2008. The components of mutual fund fees. Financial Markets, Institutions and Instruments 17 (3), 197–223.

Gârleanu, N. B., Pedersen, L. H., 2015. Efficiently inefficient markets for assets and asset management, NBER Working Paper 21563, available at http://www.nber.org/papers/w21563.

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

Gennaioli, N., Shleifer, A., Vishny, R., 2015. Money doctors. Journal of Finance 70 (1), 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 (5), 2153–2183.

Greenwood, R., Scharfstein, D., 2013. The growth of finance. Journal of Economic Perspectives 27 (2), 3–28.

Grossman, S. J., Stiglitz, J. E., 1980. On the impossibility of informationally efficient markets. American Economic Review 70, 393–408.

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.

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

Hortacsu, A., Syverson, C., May 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, 2015. 2015 Investment Company Fact Book. Investment Company Institute.

Jensen, M. C., 1968. Problems in selection of security portfolios. The Journal of Finance 23 (2), 389–416.

Kahn, V. M., 14 July 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 (3), 331–339.

Livingston, M., Zhou, L., 2015. Brokerage commissions and mutual fund performance. Journal of Financial Research 38 (3), 283–303.

Malkiel, B. G., 1995. Returns from investing in equity mutual funds 1971 to 1991. Journal of Finance 50 (2), 549–572.

Malkiel, B. G., 2013. Asset management fees and the growth of finance. Journal of Economic Perspectives 27 (2), 97–108.

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.

Philippon, T., 2015. Has the U.S. finance industry become less efficient? on the theory and measurement of financial intermediation. American Economic Review 105 (4), 1408–1438.

Philippon, T., Reshef, A., 2012. Wages and human capital in the U.S. financial industry: 1909–2006. Quarterly Journal of Economics 127 (4), 1551–1609.

Pine, A., October 20 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., Forthcoming. It pays to set the menu: Mutual fund investment options in 401(k) plans. Journal of Finance.

Reuter, J., Zitzewitz, E., 2015. How much does size erode mutual fund performance? a regression discontinuity approach, working paper, available at http://ssrn.com/abstract=1661447.

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. Tech. rep., Deloitte Consulting LLP.

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

Salop, S. C., 1979. Monopolistic competition with outside goods. Bell Journal of Economics 10 (1), 141–156.

Shive, S., Yun, H., 2013. Are mutual funds sitting ducks? Journal of Financial Economics 107 (1), 220–237.

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 (2), 186–192.

Tirole, J., 1988. The Theory of Industrial Organization. MIT Press.

Wealthfront, 2016. Mutual fund fees. Online document. Visited on 11/6/2016.
URL `https://www.wealthfront.com/research/mutual-fund-fees`

Wermers, R., 2000. Mutual fund performance: An empirical decomposition into stock-picking talent, style, transaction costs, and expenses. Journal of Finance 55 (4), 1655–1703.

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

A Appendix: Measuring asset management fees on U.S. retirement accounts

In this section, to be added, we will estimate asset-weighted average asset management fees as a percent of assets under management. We will also attempt to estimate how costs vary as a function of assets and other relevant variables (e.g. number of accounts). We will be using the following data:

- Employer-sponsored plans:
  - A comprehensive database of Form 5500 filings, to get plan costs, number of participants, and assets under management for each fund in the plan.
  - A sample of hand-collected Form 5500 Schedule H Line 4i attachments, to obtain the amount of plan assets invested in each fund available within a plan
  - A sample of 404(a)(5) fee disclosure forms for 401(k) and 403(b) plans, to obtain plan menus, and fees for each fund on the menu
- IRAs
  - ICI data on aggregate balances in IRAs (Roth and Traditional)
  - Mutual fund fee information from Morningstar and other sources

B Extra information on fees

B.1 Account fees

Many investors hold their retirement savings in some kind of wealth management account (including brokerage accounts, managed accounts, or defined-contribution retirement plans). Typically, these accounts pay some kind of "wrap fee" or account management fee (typically a fixed percent of assets in the account). These fees cover basic account administration costs and, in many cases, premium services such as financial advisory. Advisory can also be provided as a separate service from the account, for a separate fee.

Part of the advisory service is covered by mutual fund distribution fees (i.e., 12b-1 fees) that are already included in the expense ratio. However, investors buying shares from mutual fund brokers may incur "load" fees, e.g., upfront fees or fees upon redemption of their shares. These load fees directly pay for the advisory services the broker performs at the time when the mutual fund is selected.[12] Investors paying explicit account fees—such as managed accounts fees, or plan management costs paid by the sponsors of defined-contribution employer plans—typically have access to no-load funds inside the account. Thus, load fees and account management fees do not *usually* appear together. Overall, advisory and distribution fees (excluding 12b-1 fees, which are already included in the net expense ratio) average about 50 bps (Bogle, 2014).

B.2 Mutual fund fees

The net expense ratio includes three types of costs. First, paperwork costs: custodial fees, legal fees, record-keeping fees, etc. These fees typically cover the cost of inevitable services provided by third parties unaffiliated with the mutual fund. The second type of costs are distribution and service fees (so-called 12b-1 fees). 12b-1 fees cover two types of expense: distribution costs, i.e., commissions to the sales force (capped at 75 bps), and shareholder servicing costs, e.g., cost of providing internet access to fund filings, etc.. (capped at 25 bps).[13] 12b-1 fees are included in the fund's expense ratio and they are taken from the fund's NAV. Third, the net expense ratio includes asset management advisory fees, i.e., the actual revenue of the money management company that sponsors the fund in the first place.[14]

[12]Both the distribution costs component of 12b-1 fees and sales loads constitute compensation for the broker, rather than the fund manager. However, the two fees are not overlapping or mutually exclusive. 12b-1 fees are taken year after year out of fund assets, loads are directly paid by the investor to the broker. For instance, with a 5% load, an investor giving $100 to the broker is only investing $95. If the fund has 12b-1 fees in addition to loads these fees will be levied upon the $95. The same fund may have multiple classes of shares. According to Morningstar's Glossary, "In a typical multi-class situation, the class A fund has a front-end load and either a 0.25% distribution fee or a 0.25% service fee. Class B shares usually have a contingent deferred sales charge and a corresponding 0.75% 12b-1 fee, plus a maximum 0.25% service fee. [...] Class C shares customarily charge a level load with the same fee structure found in a class B share."

[13]12b-1 fees are so called after SEC Rule 12b-1 under the Investment Company Act of 1940. FINRA regulations from 1993 establish the caps on these fees. See SEC > Mutual Funds Fees and Expenses (https://www.sec.gov/answers/mffees.htm).

[14]Typically, advisory fees are not set at arm's length because the fund is a captive customer of the management company. It is generally believed that market forces curb excessive advisory fees, because of the threat of investors withdrawing their money and taking it to a different fund (e.g., Coates and Hubbard, 2007). Others contend that market forces are not sufficient to keep fees in check because no fund's fees are set at arm's length; even if a fund's fees appear "reasonable" with respect to the competition, they need not be reasonable overall (Freeman et al., 2008). The Supreme Court (*Jones et al. v. Harris Associates L.P.*, 2010) rejects the "market" argument, in part because it is conscious of the lack of arm's length prices, arguing instead in favor of the "workable standard" set in the *Gartenberg* case, i.e., that in order for high fees to be evidence of breach of fiduciary duty, they must be so disproportionately high that they bear no resemblance to the services provided and could not be the result of arm's length bargaining. Evidence of breach of fiduciary duty must otherwise be found in the process by which the mutual fund board has reviewed the advisor's fees.

C Appendix: Full discussion of the model

The model features a two-period economy. When individuals are young (time t), they work, produce, consume, and save for retirement by investing via mutual funds in a mix of government bonds and financial assets. When they are old (time t'), they receive passive retirement income, consume, and die without bequest. Mutual funds require a fixed amount of labor to operate, and therefore must pay a competitive salary to individuals who otherwise would work to produce goods.

C.1 Individuals

Individual $i \in \mathcal{I}$ lives two periods and leaves no bequest. In period t the individual works, saves, and allocates the savings. The individual starts with a net worth of 0 and is endowed with one unit of labor to spend either producing consumption goods, or managing a mutual fund. The production technology is linear: if the individual allocates L units of labor to the production of goods, the output is $F(L) = \omega L$. In order to attract labor, therefore, mutual funds must pay a competitive wage ω. The individual's total supply of labor is inelastic. Overall, the individual's pretax income is ω regardless of how labor is allocated. Income from work is taxed at rate τ_L. In period t', the individual retires and depletes all the savings.

The individual draws utility from current consumption (C) and from discounted future consumption ($\delta \cdot C'$). In addition, the individual prefers a fund of type ι_i, and derives disutility that increases with the distance between the chosen fund type and one's own preference ($d_{i,j} \equiv |\iota_j - \iota_i|$). The individual saves an amount S. A fraction a of the individual's savings is invested in financial assets paying a return ρ, and the remainder in government bonds paying a return r. The government grants the individual a deduction for savings at a rate τ_S. All investment is carried out via mutual fund j charging a proportional fee $f_j S$. Thus, the utility of individual i is:

$$U_i = \max_{C,C',S,a,j} \ln C + \delta \ln C' - \gamma d_{i,j} \tag{17}$$

subject to the budget constraints:

$$C = \omega \left(1 - \tau_L\right) - S \left(1 - \tau_S\right), \tag{18}$$

$$C' = S \left(1 - f_j\right) \left[1 + a\rho + \left(1 - a\right) r\right] \left(1 - \tau_R\right). \tag{19}$$

C.1.1 The individual's savings and asset allocation decisions

The individual decides how much to save, S, and the fraction to allocate to stocks, a. It is simpler to solve for a first.

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

$$\frac{\partial}{\partial a}\ln C+\delta\frac{\partial}{\partial a}\ln C'=0,$$

i.e.

$$r=\rho. \tag{20}$$

This condition permits us to rewrite the time-t' budget constraint (19) as:

$$C'=S\left(1-f_j\right)\left(1+\rho\right)\left(1-\tau_R\right) \tag{21}$$

Next, the individual's first-order condition with respect to savings S is

$$\frac{\partial}{\partial S}u\ln C+\delta\frac{\partial}{\partial S}\ln C'=0.$$

Rewrite using the budget constraints (18) and (21) and (20):

$$\frac{\partial}{\partial S}\ln\left(\omega\left(1-\tau_L\right)-S\left(1-\tau_S\right)\right)+\delta\frac{\partial}{\partial S}\ln\left(S\left(1-f_j\right)\left(1+\rho\right)\left(1-\tau_R\right)\right)=0$$

to obtain the Euler equation:

$$\left(1+\rho\right)\delta\frac{C}{C'}=\frac{1}{1-f_j}\frac{1-\tau_S}{1-\tau_R}. \tag{22}$$

C.1.2 The individual's choice of a fund

Individuals choose fund j to satisfy the following criterion:

$$j^*=\arg\max_j\ln C+\delta\ln C'-\gamma d_{i,j}$$

Since C does not depend on f_j, simplify

$$j^*=\arg\max_{j\in\{1,2,\ldots N\}}\delta\ln C'\left(f_j\right)-\gamma d_{i,j} \tag{23}$$

where $C'\left(f_j\right)$ is written in that way to underscore that future consumption depends on the fees paid on one's chosen fund.

C.2 Mutual funds

Mutual funds can differentiate themselves over one qualitative characteristic, ι, defined on the $[0, 1)$ circle. Individuals are uniformly distributed over this circle, and their utility is decreasing in the distance from their chosen fund. Every fund $j \in \mathcal{J}$ needs a fixed amount of labor φ just to be able to operate. Fund profits are equal to revenue minus cost:

$$\pi_j = f_j Q_j - \varphi\omega,$$

where Q_j is the fund's assets under management, so that $f_j Q_j$ is the fund's total revenue. The fund's problem is to choose $f_j$and ι_j to maximize π_j, taking into account that Q_j depends on the fund's choices of pricing (f_j) and location (ι_j), and taking competitors' choices as given.

Mutual fund profits accrue to their managers and are taxed as other income. However, because of free entry, profits will be zero, so we do not keep track of who the profit accrues to.

C.2.1 Funds' location and pricing decisions

Fund j sets fees f_j and location ι_j to maximize profits:

$$\max_{f_j,\iota_j} f_j q_j S - \varphi\omega. \tag{24}$$

Because all funds choose simultaneously whether to enter the market, the problem is symmetric for every fund j, which implies that funds will distribute at a constant distance over the circle (Tirole, 1988)

$$d_{j,j+1} \equiv |\iota_j - \iota_{j+1}| = {}^1\!/\!{}_N \qquad \forall j \in \mathcal{J}, \tag{25}$$

and all funds will charge the same fees

$$f_j = f \qquad \forall j \in \mathcal{J}. \tag{26}$$

Finally, because entry is free, the number of funds that choose to enter, N, will be such that

$$\pi_j = 0 \qquad \forall j \in \mathcal{J}. \tag{27}$$

C.3 Government

The government spends an exogenously given amount G. This expenditure is inevitable and it does not affect the utility of agents. At time t, the government taxes income at a rate τ_L, and grants individuals a deduction for savings at a rate τ_S. The government can also borrow an amount B at the market interest rate R to satisfy the government budget constraint:

$$G = \tau_L \omega - S\tau_S + B$$

To pay off the bonds at time t', the government can tax retirement income at a rate τ_R so as to satisfy

$$B\left(1+r\right) = \tau_R S\left(1-f\right)\left[1 + a\rho + \left(1-a\right)r\right]$$

where f is the equilibrium level of fees.

C.4 Market clearing

C.4.1 Consumption good market

Total output of goods is equal to consumption plus savings plus government expenditure:

$$C + S + G = \omega L \tag{28}$$

C.4.2 Labor market

Total labor employed in the production of goods (L) and by the N mutual funds (φN) must be equal to the working population, i.e., equal to one.

$$L + \varphi N = 1. \tag{29}$$

C.4.3 Asset management market

Total assets under management by funds are equal to total savings by individuals:

$$\sum_{j\in\mathcal{J}} Q_j = \sum_{i\in\mathcal{I}} S_i = S.$$

Define market share $q_j = \frac{Q_j}{S}$, then the market clearing condition becomes

$$\sum_{j\in\mathcal{J}} q_j = 1 \tag{30}$$

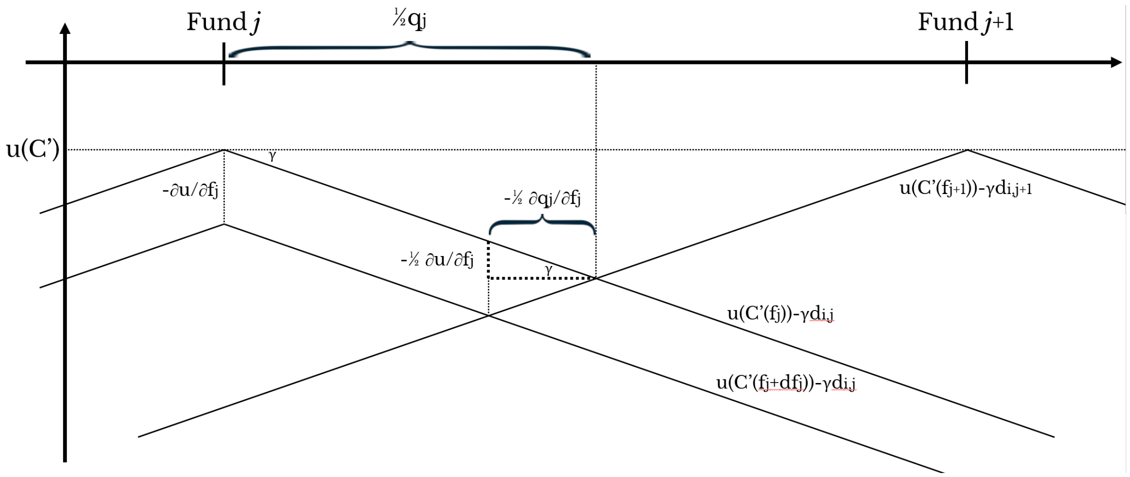

Figure 3: Geometric intuition: competitive fee equilibrium for mutual funds.

C.4.4 Government bond market

After fees, the capital available to buy government bonds is $(1-a)\,S\,(1-f)$. The government gives a subsidy τ_S to saving, and borrows to make up the difference:

$$(1-a)\,S\,(1-f) = B. \tag{31}$$

C.5 Market solution

C.5.1 Asset management market equilibrium

In a hypothetical equilibrium there is a marginal investor who is indifferent between funds j and $j+1$. For this individual,

$$\delta \ln C'\left(f_j\right) - \gamma d_{i,j} = \delta \ln C'\left(f_{j+1}\right) - \gamma d_{i,j+1}. \tag{32}$$

Equation (26) says that, in equilibrium, all funds charge the same fees f because of symmetry considerations. Moreover, when fund j chooses f_j, it takes all other funds' fees as given. Thus, we can write f_{j+1} as simply the aggregate average level of fees, f. Because of the definition of marginal investor, and because of the symmetry of the problem, the distance between fund j and the marginal investor is equal to one-half q_j, the market share of fund j:

$$d_{i,j} = \frac{1}{2} q_j.$$

Finally, equation (25) says that the distance between fund j and fund $j+1$ is known to be $1/N$, so

$$d_{i,j+1} = d_{j,j+1} - d_{i,j} = \frac{1}{N} - \frac{1}{2}q_j.$$

Replacing distances in terms of demand within (32) we obtain the following demand function that fund j faces *in equilibrium*:

$$q_j = \frac{1}{N} + \frac{\delta}{\gamma}\left(\ln C'\left(f_j\right) - \ln C'\left(f\right)\right) \tag{33}$$

The geometric intuition behind this result is represented in Figure 3.

Given the fund's objective function (24), we can rewrite the first-order condition for maximization as

$$\frac{1}{N} + \frac{\delta}{\gamma}\left(\ln C'\left(f_j\right) - \ln C'\left(f\right)\right) - f_j\frac{\delta}{\gamma}\frac{1}{1-f_j} = 0. \tag{34}$$

The optimal level of fees f_j^* solves this condition.

Now, consider the market equilibrium. Because every fund faces the same problem, $f_j^* = f$, so that the second term of (34) cancels out. The equilibrium level of fees of fund j, equal to all other funds, is then

$$f_j^* = f = \frac{1}{1 + \frac{\delta}{\gamma}N}. \tag{35}$$

From (27) we know that because of free entry, profits are zero, i.e., revenues are equal to costs. Thus,

$$fQ = \frac{1}{1 + \frac{\delta}{\gamma}N} \cdot \frac{S}{N} = \varphi\omega \tag{36}$$

C.5.2 Goods market equilibrium

The individual's budget constraint (18) pins down S as a function of C:

$$C = \omega\left(1 - \tau_L\right) - S\left(1 - \tau_S\right)$$

but we also know, from the Euler equation, that

$$C = \left(1 - \tau_S\right)\frac{S}{\delta}$$

and thus we have a fixed, closed-form expression for savings that does not depend on the number of funds N:

$$S = \omega\frac{1-\tau_L}{1-\tau_S}\frac{\delta}{1+\delta} \tag{37}$$

Because S does not depend on N, total profits of the fund industry are a strictly decreasing function of N:

$$\sum_{j=1}^{N} \pi_j = \left(f_j^* Q_j - \varphi\omega\right) N = \frac{1}{1+\frac{\delta}{\gamma}N} S - \varphi\omega N$$

(the second equality follows from the definition of the equilibrium level of fees f^* and because fund market share $Q_j = S/N$). More funds means more competition with lower fees (first term), and more total fixed costs (second term). Since demand is unvaried, lower revenues and higher costs translates into lower profits. This is not obvious; if industry profits are ever to reach zero for some N^*, they have to be a decreasing function of N *past a certain point*, but not monotonically.

C.5.3 Number of funds

The updated zero-profit condition (36) and the expression for savings (37) help us pin down the parameter of interest: N, the equilibrium number of funds under the market solution.

$$\frac{\delta}{\gamma} N^2 + N = \frac{S}{\varphi\omega} = \frac{1}{\varphi} \cdot \frac{\delta}{1+\delta} \cdot \frac{1-\tau_L}{1-\tau_S}. \tag{38}$$

Recall that under the market solution, savings are fixed:

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

where τ_L is the labor income tax rate and τ_S is the subsidy to saving. In the case of a traditional $\tau_S = \tau_L$, and for a Roth $\tau_S = 0$. Therefore, without solving explicitly for N, it is already evident that S and N, are higher with a Traditional than they are with a Roth scheme.

C.6 Planner solution

The planner also takes G as exogenously given, and chooses savings S, and number of funds N directly to maximize social utility:

$$U = \max_{C,C',S,N} \ln(C) + \delta \ln(C') - \gamma \bar{d} \tag{39}$$

with $\bar{d}$ the average distance between an investor and their fund. The planner's budget constraints are simply

$$C = \omega(1-\varphi N) - S - G, \tag{40}$$

$$C' = S(1+\rho) = (\omega(1-\varphi N) - C - G)(1+\rho). \tag{41}$$

The N funds are equally spaced along the circle.[15] Individuals' distance from the nearest fund is uniformly distributed over $[0,\ 1/(2N)]$, hence the density $f(x) = 2N$ for all x, and the average distance is (obviously):

$$\bar{d} \equiv \int_{i\in\mathcal{I}} d_{ij} = \int_0^{1/(2N)} xf(x)\,dx = 2N\left[\frac{1}{2}x^2\right]_0^{1/(2N)} = \frac{1}{4N}. \tag{42}$$

Using (40), (41), and (42), the utility function can be rewritten as

$$U = \max_{S,N} \ln(\omega(1-\varphi N) - G - S) + \delta \ln(S(1+\rho)) - \frac{\gamma}{4N} \tag{43}$$

The first-order condition with respect to N is

$$C = \frac{4\varphi\omega}{\gamma}N^2 \tag{44}$$

The first-order condition with respect to S is

$$S = C\delta \tag{45}$$

Using (40), rewrite the time-t budget constraint (40) as

$$C = \omega(1-\varphi N) - C\delta - G$$

$$C = \frac{\omega(1-\varphi N) - G}{1+\delta} \tag{46}$$

Finally, using (44) and (46), obtain an implicit expression for the socially optimal number of funds:

$$4\frac{1+\delta}{\gamma}N^2 + N = \frac{1-G/\omega}{\varphi} \tag{47}$$

D Appendix: Calibration details

The parameters we need to calibrate are

- δ, the discount rate between working life consumption and retirement consumption.
- G/ω, government spending as a fraction of total output

[15] In principle, the planner gets to choose the optimal placement of funds too, but it is easy to show that equal spacing is indeed optimal.

- φ, the percentage of the labor force employed by the average mutual fund (or, more intuitively, φN, the percentage of labor force employed by the mutual fund industry). Total labor force is normalized to one in the model, so φ is expressed as a fraction of total labor force in the economy.

- γ, i.e. how much people care about convenience in the choice of a mutual fund, vs. consumption.

D.1 Discount rates

δ is the one-period discount rate. How long is one period? The first period is working life (~35 years) and the second period is retirement (~25 years). The average number of years between savings and consumption is about 30 years. In order for δ to be low enough that the equilibrium number of funds is too low, individuals must be unrealistically impatient (e.g. with a 15% per annum discount rate, the 30-year δ is about 0.01). For our calibration, we adopt a discount rate of 2% per annum, i.e. $\delta = (1.02)^{-30} = .545$ over a 30-year period.

D.2 Government expenditure

G/ω is government expenditure as a percent of total output. Note that in this model government expenditure matters only to the extent that it subtracts real resources from the economy. To exclude transfers, we use the "Real Government Consumption Expenditures and Gross Investments" from the Bureau of Economic Analysis. This expenditure is about 18% of gross domestic product.

D.3 Mutual fund industry

φ can be directly estimated by combining information about the number of funds in the U.S. economy, the total number of workers, and the size of the labor force. According to ICI (2015), in 2013 the asset management industry had employed 166,000 people, and featured 7,713 funds. The Bureau of Labor Statistics reports a total employment of about 130 million people, implying $\varphi = 0.000017\%$.

D.4 Investors' preferences for convenience

γ has no real-world equivalent and thus it can only be calibrated indirectly. However, the level of fees is observable in the real world. In the model, f is the level of fees (per period). We have an equation for fees as a function of δ, γ, and N.

$$f = \frac{1}{1 + \frac{\delta}{\gamma} N}$$

Once again, we need to adjust f to take into account the fact that one period is about 30 years. We choose 0.20% as the minimum cost of money management, based on the expense ratios of large index funds, and we choose a 30-year horizon, implying $f = 1 - (1 - 0.20\%)^{30} = 5.83\%$. We then change γ in the model until f is about 6%. This yields $\gamma = 3500$.