

# On the Effects of Tax-deferred Saving Accounts\*

Anson T. Y. Ho <sup>†</sup>

## Abstract

This study investigates the quantitative impacts of tax-deferred saving accounts (TDAs) in a general equilibrium framework. These accounts implicitly provide consumption tax treatments on households' retirement savings. In the U.S., the use of TDAs exhibits substantial heterogeneity: 401(k) has a much higher contribution limit than IRA, but only 50% of workers are eligible for it. I developed an OLG model that captures the tax saving effects of TDAs and the heterogeneity in 401(k) eligibility in presence of nonlinear taxes. Results from this study show that providing universal 401(k) eligibility will increase the aggregate capital and the aggregate output are increased by 4.7% and 1.8% respectively. Also, the existence of TDAs substantially reduces the impacts of a flat tax reform. Compared to a model *without* TDAs, the increase in the aggregate output in the model *with* TDAs is lowered by 20.9%.

*JEL classification:* H24, E21, E64.

*Keywords:* Tax-deferred Accounts, Taxation, Flat Tax, Saving Incentives, Retirement.

---

\*I am indebted to Gustavo Ventura for his continuous advice and encouragement. I also benefited from the comments of R. Anton Braun, Shutao Cao, Steven P. Cassou, Norma Coe, Nadja Dwenger, Srihari Govindan, Kim P. Huynh, Remzi Kaygusuz, Shinichi Nishiyama, Brian Peterson, B. Ravikumar, James K. Self, Guillaume Vandenbroucke, Yuzhe Zhang, Jie Zhou, and seminar participants at the 2010 Canadian Economic Association Annual Conference, 2011 Allied Social Science Associations Annual Meeting, 2011 Netspar International Pension Workshop, the University of Northern Iowa, Kansas State University, and Bank of Canada. I am grateful to Alberto Segre for the provision of computing facilities at the Department of Computer Science of the University of Iowa. Part of this paper is completed when the author was affiliated with the University of Iowa. All errors are mine.

<sup>†</sup>Department of Economics, Kansas State University, 302D Waters Hall, Manhattan, KS 66506-4001, USA. Phone: +1(785)532-4582. Email: [atyho@ksu.edu](mailto:atyho@ksu.edu).

# 1 Introduction

Tax-deferred saving accounts (hereafter TDAs) are important instruments for retirement savings, and they are systematically used in many countries. For example, Canada, Germany, Italy, the Netherlands, and the United Kingdom all allow tax-deferred saving accounts with similar institutional settings.<sup>1</sup> In the U.S., TDAs include Individual Retirement Accounts (IRA), 401(k) for private-sector employees, 403(b) for nonprofit-sector employees, 457 plan for public-sector employees, and Keogh accounts. Generally, all types of TDAs have three common features. First, they provide favorable tax treatments on the capital income within these accounts. Assets within TDAs grow tax-free until withdrawal. Second, contributions to TDAs are tax deductible, and subsequent withdrawals are taxed as ordinary income. Third, there are contribution and withdrawal constraints on assets in TDAs. Contributions are restricted to certain legal limits and can only be made in working age. Early withdrawal before retirement age is subject to penalty payment in addition to the income taxes incurred from the assets withdrawn.

TDAs are of particular interest because of their impacts on the progressive income tax system. Since assets in TDAs and their returns are exempted from income tax until they are withdrawn for consumption, TDAs effectively provide consumption tax treatment on income set aside for retirement consumption. With about 7.7 trillion dollars of assets (about 52.7% of GDP) accumulated in TDAs in 2010, the TDA system has become a sizable part in the U.S. economy. The current U.S. income tax system is essentially a hybrid one made up of a progressive income tax component and a consumption tax component. As a result, the consumption tax treatment on TDA assets mitigates the distortions created by capital income tax and household income tax.

TDAs also offer a way for households to redistribute their taxable income over time. While contributions to TDAs decrease account holders' current period taxable income, subsequently withdrawals of TDA assets will increase their future taxable income. In a progressive income tax system, it represents a significant tax arbitrage opportunity, which is due to the difference in household income levels before and after retirement. Working-age households receive income from two sources - labor earnings and capital income, and thus, they face higher marginal income tax rates. After retirement, returns on capital become their sole source of income and place them in tax brackets with lower marginal tax rates. Households can contribute to TDA and take advantage of the differences in marginal tax rates. Thus, TDAs offer households another channel to respond to distortionary tax policies, in addition to the labor supply decisions in a conventional framework. Apart from the tax-preferred treatment on capital income in TDAs, these institutional settings of TDAs induce households to incorporate tax avoidance incentives

---

<sup>1</sup>See [Guiso, Jappelli, and Haliassos \(2001\)](#) for a detailed description.

into their decisions.

The purpose of this study is to investigate the quantitative impacts of TDAs in a general equilibrium framework. Since the purpose of TDAs is to encourage private savings, this paper first addresses the effects of TDAs on this aspect. Incremental savings is measured by the differences in the steady state capital levels between the benchmark economy with TDAs and an economy without TDAs. Then, this paper evaluates the impacts of granting 401(k) access to all working-age households, *i.e.* universal 401(k) eligibility. This experiment analyzes the effects of expanding the TDA system and removing heterogeneity in 401(k) eligibility. By expanding the consumption tax component, the provision of universal 401(k) eligibility is a way to move towards consumption taxation with minimal changes to the existing tax system. After that, this paper proceeds to address the implications of TDAs for a flat tax reform. Since the impacts of a tax reform depends on the conditions of the initial economy, these effects are potentially overstated when TDAs are not taken into account.

An overlapping generations (OLG) model is employed to study the use of TDA in the U.S. economy. Households are heterogeneous in their labor efficiencies, which also evolve stochastically over time. Each period they decide the amount of time they spend at work and allocate their income on consumption and savings. Households can accumulate assets in their TDAs and/or as their ordinary asset holdings (hereafter OAH). TDAs in this model are subject to contribution and withdrawal constraints and receive tax-free treatment on capital income. Households are also heterogeneous in their 401(k) eligibilities. While all households are eligible for IRA, only a fraction of households are eligible for 401(k). Households' 401(k) eligibilities depend on their labor efficiencies and the pattern is calibrated to match the data in the 2001 Survey of Income and Program Participation (SIPP). After TDA contributions are deducted, household's remaining income is taxed in a progressive tax system. Markets are incomplete and households face uninsurable risks on labor efficiency, mortality, and 401(k) eligibility. These model settings are able to capture key aspects of the TDA system.

The main findings of the study are summarized as follows: TDAs only increase the aggregate capital stock by 3.7%. Most of the assets in TDAs are reallocated from OAH. TDAs also have very little influences on the aggregate labor supply. At the household level, TDAs create tax incentives that significantly affect household's life-cycle decisions on labor supply and savings. Furthermore, 401(k) eligibilities households supply more labor hours and accumulate more assets than noneligible households. When 401(k) eligibility is extended to all households, the aggregate capital and the aggregate output further increase by 4.7% and 1.8% respectively. Note that expanding the TDA system through universal 401(k) eligibility only has small effects on the government revenue and requires no alternation to the income tax code. Hence, the government can increase the aggregate output at a low cost. This paper also shows that TDAs substantially

reduce the impacts of a flat tax reform. Under the same tax reform, the increase in aggregate output in an economy with TDAs is 20.9% lower than that in an economy without TDAs.

The contribution of this study is multifold. The model in this study captures multiple aspects of the current TDA system and provides a better understanding on the effects of TDAs on the aggregate economy. It also sheds light on how household life-cycle decisions are influenced by TDAs and their 401(k) eligibility. Moreover, this study illustrates that TDAs have important implications for a flat tax reform, suggesting that omitting TDAs, or tax avoidance technologies in general, in the assessment of tax policies can be misleading.

This paper is organized in the following way: Section 2 reviews the literature on TDA and tax reforms. Section 3 highlights the features and stylized facts about IRA and 401(k) in the U.S. Section 4 describes the model economy. Section 5 explains the model's calibration strategy. Section 6 shows the results of the benchmark model. Section 7 analyzes the effects of TDA on the aggregate economy and household life-cycle decisions. Section 8 investigates the impacts of universal 401(k) eligibility. Section 9 studies the implications of TDA for tax reforms. Finally, Section 10 concludes.

## 2 Related Literature

This study is related to three different streams of research in the tax literature, which can be broadly categorized into the study of tax avoidance, TDAs, and tax reform.

There are vast evidences that high income households respond to the incentives of minimizing tax liabilities. [Feldstein \(1995\)](#) estimates the sensitivity of taxable income to changes in tax rates by comparing the tax returns of taxpayers before and after the 1986 Tax Reform. He finds that the elasticity of taxable income with respect to marginal tax rate is at least 1. A complementary study by [Auerbach and Slemrod \(1997\)](#) focusing on the same tax reform conclude that there was a hierarchy of responses, with the most responsive decisions being activities that primarily affect reported income, while the least responsive ones being the real decisions of households and firms. These results provide support to the tax avoidance literature that the response of taxable income involves decisions more than just a change in labor supply. [Joulfaian and Richardson \(2001\)](#) study the characteristics of households participating in TDA plans, and they find that higher labor earnings and marginal tax rates increase the probability of TDA participation. It provides evidence that households do exploit the tax arbitrage opportunities offered by TDAs. And yet, the tax-based incentive of TDA has not been fully explored.

In addition to investigating the tax avoidance incentives of TDAs, a lot of effort has been spent on evaluating the effectiveness of TDAs in creating new savings. Since some of the assets in TDAs are shifted from other accounts, which would have been saved anyway, only

a fraction of TDA assets represent new savings. To measure the amount of new savings, an intuitive way is to look at the differences in asset levels between households with and without TDAs. However, household saving decisions are also influenced by other factors such as their demographic characteristics and income levels. Empirical estimates on the incremental savings have been done by using different sources of data and different methods for controlling household heterogeneity.<sup>2</sup> The results in these empirical studies are inconclusive and are sensitive to the method used. For example, [Gale and Scholz \(1994\)](#) control household heterogeneity through a structural approach and find that IRA has little effects on new savings. However, [Poterba, Venti, and Wise \(1995\)](#) control heterogeneity by grouping them in eligibility categories and find little evidence that 401(k) contributions substitute for other forms of personal saving. In the absence of a completely randomized control experiment, none of these empirical methods can perfectly isolate the effects of TDA from other factors. Furthermore, these empirical studies are not able to identify the saving incentives induced by tax arbitrage with the preferential tax treatment of capital returns.

In a general equilibrium framework, [Imrohoroglu, Imrohoroglu, and Joines \(1998\)](#) use an OLG model to evaluate the effectiveness of IRA on increasing aggregate capital, and they found that the incremental saving lies at the lower end of the range of estimates in the empirical literature. [Gomes, Michaelides, and Polkovnichenko \(2009\)](#) consider a more complex asset portfolio structure with direct and indirect stockholders and find that TDA only marginally increase net savings. [Kitao \(2010\)](#) extends the analysis by endogenizing household labor supply decisions and incorporating a progressive income tax system into her model to capture the effects of tax arbitrage opportunities offered by TDAs in a nonlinear tax system. She finds that TDAs have a strong impact on raising capital and output levels. [Nishiyama \(2011\)](#) further explores the transition dynamics and the short run effects of introducing TDAs under different government financing options. A common assumption among studies in theoretical framework is that households are homogeneous in their TDA eligibilities. For instance, [Imrohoroglu, Imrohoroglu, and Joines \(1998\)](#) and [Kitao \(2010\)](#) set the TDA contribution limit to be roughly the same as the IRA limit, excluding 401(k) from consideration. In contrast, the contribution limit in [Nishiyama \(2011\)](#) is similar to that of 401(k), implying that all households are 401(k) eligible. However, there is vast evidence that households are heterogeneous in terms of 401(k) eligibility.<sup>3</sup>

This paper focuses on the distributional impacts of TDAs and their implications on tax reforms. The model in this paper is closest to that in [Imrohoroglu, Imrohoroglu, and Joines \(1998\)](#) and [Kitao \(2010\)](#), but it differs in two major ways. It incorporates a nonlinear tax

---

<sup>2</sup>See [Poterba, Venti, and Wise \(1996\)](#), [Engen, Gale, and Scholz \(1996\)](#), and [Bernheim \(2002\)](#) for a review of the empirical methods used to estimate TDA induced incremental savings.

<sup>3</sup>See Section 2 for a detailed description about TDAs in the U.S.

system and hence the tax arbitrage opportunities associated with TDAs. Endogenous labor supply and the more complex labor efficiency process in this model generates heterogeneity in labor earnings similar to the U.S. data, that allows the model to explore the effects of TDAs on the aggregate economy. Furthermore, this model captures the heterogeneity of 401(k) eligibility among households. This is a critical factor in analyzing the impacts of TDA because it limits the extent to which TDAs are used by households. If 401(k) is excluded from the TDA system, the IRA contribution limit becomes an effective constraint on the amount of contributions that agents can make. As result, contributions are likely to be shifted from other sources of savings which would have been done anyway. Thus, the effects of TDA is limited and increasing the TDA contribution limit will have significant impacts on the economy. On the contrary, assuming that all households have access to 401(k) as in [Imrohoroglu, Imrohoroglu, and Joines \(1998\)](#) and [Kitao \(2010\)](#), most agents are unrestricted by the contribution limit. The impacts of introducing TDA to the economy will be more significant. Since a further increase in the contribution limit will only affect the small fraction of agents who contribute the maximum amount, the effects of relaxing an already generous TDA contribution limit is less significant. Also, the model in this paper is more carefully calibrated, taking certain nature of household assets into account.

There is vast research on tax reforms. Notable ones are [Auerbach \(1997\)](#), [Ventura \(1999\)](#), and [Altig, Auerbach, Kotlikoff, Smetters, and Walliser \(2001\)](#). These studies have taken the progressive tax system at its face value. However, studies on the relationship between tax avoidance and flat tax reform are scarce. As evidences in tax elasticity show that households can minimize their tax liabilities by adjusting their taxable income through various ways in addition to changing their labor supply, studying the impacts of tax policies by only considering the de jure tax rates can be misleading. Given that TDA is widely available and commonly used, studying the impacts of TDA on a flat tax reform shed some light on the importance.

### 3 Features of IRA and 401(k)

Tax-deferred saving accounts in the US has gone through a number of changes since they were introduced. IRA was created in 1974 with the aim of increasing people readiness for retirement. Eligible individuals can and are responsible for setting up their own IRAs with a variety of organizations. Initially, IRAs were limited to workers without a qualified employer retirement plan. The contribution limit was \$1,500 from 1975 to 1981. After the Economic Recovery Tax Act of 81 (ERTA '81) was passed, all individuals who are below 70.5 years old and receive compensation from work during the year can set up and contribute their pre-tax income to IRA.<sup>4</sup> The contribution limit was increased to \$2,000 from 1982-2001. The Internal Revenue

---

<sup>4</sup>Publication 590 (2009), Individual Retirement Arrangements (IRAs), <http://www.irs.gov/publications/p590/>

Service (IRS) further increased the contribution limit to \$3,000 in 2002, then \$4,000 in 2005, and \$5,000 in 2008.

As suggested by its name, 401(k) retirement plans were created according to section 401(k) of the Internal Revenue Code in 1978. 401(k) plans are offered by employers and allow a covered employee to have a portion of compensation contributed to her 401(k) plan as a pre-tax reduction in salary. To avoid 401(k) being an unfair instrument leaning towards to rich, the Tax Reform Act of 1984 (TRA '84) introduced employers to “nondiscriminating” rules to ensure that 401(k) does not discriminate in favor of highly compensated employees. The Tax Reform Act of 1986 (TRA '86) further tightened the nondiscrimination rules and substantially reduced the 401(k) contribution limit. In 2000, the 401(k) contribution limit per worker was \$10,500 or 25% of her annual salary, whichever is less. The contribution limit was increased to 100% of the worker’s annual salary in 2002. According to the Survey of Income and Program Participation (SIPP), only about 50% of total workers were offered 401(k) plans by their employers in 2001. Figure 1 shows the 401(k) eligibility of workers conditional on their labor earnings. Note that the fraction of workers eligible to 401(k) is positively correlated to labor earnings.

While the IRA contribution limit remained about 10% of average income, 401(k) contribution limit was about 50% average income. It imposes substantial heterogeneity in TDA eligibility and hence tax uncertainty. The much higher contribution limit of 401(k) implies that taking it into account is very important. According to the data from Survey of Consumer Finance (SCF), about 3.4 trillion of assets are saved in private pension funds and 4.3 trillion dollars are saved in IRAs.

## 4 The Environment

This study used an overlapping generation (OLG) model that incorporates several important household features. First, agents are heterogeneous in their labor efficiencies, which evolve stochastically over time. Second, agent’s labor supply decision is endogenous. Third, there is a risk-free asset that can be saved in two types of accounts: tax-deferred account (TDA) and ordinary asset holdings (OAH). TDA in this economy is composed of 2 elements: IRA and 401(k). The novel feature of this model is that agents are heterogeneous in their 401(k) eligibility. While all agents are eligible for IRA, their 401(k) eligibility depends on their labor efficiencies and are calibrated to match an endogenous distribution of 401(k) eligibility conditional on labor earnings. Agent’s income is subject to progressive income tax. The existence of TDA and endogenous labor supply allows agents to conduct tax arbitrage, and the stochastic TDA contribution limit controls the extent to which it can be done.

## 4.1 Agents

A large number of age one agents are born in each period, and the population grows at a constant rate of  $g$ . Agents live a maximum of  $T$  periods and they face mandatory retirement after living  $R$  periods. There is an exogenous survival probability,  $\gamma_j$ , that an agent of age  $j$  will survive to age  $j + 1$ . Let  $\mu_j$  be the fraction of population at age  $j$ , then  $\mu_{j+1} = \frac{\gamma_j}{(1+g)}\mu_j$  and  $\sum_{j=1}^T \mu_j = 1$ .

Agents are endowed with one unit of time every period. For an agent of working age  $j$  at time  $t$ , she allocates her time to work ( $l_{j,t}$ ) and leisure ( $1 - l_{j,t}$ ).<sup>5</sup> After her mandatory retirement in period  $R + 1$ , she allocates all of her time to leisure, *i.e.*  $l_{j,t} = 0$  for  $j = R + 1 \dots T$ . Thus, an agent's time constraint is expressed as

$$l_{j,t} \begin{cases} \in [0, 1] & \text{if } j \leq R \\ = 0 & \text{if } j \geq R + 1 \end{cases} . \quad (1)$$

She also decides the amount of consumption ( $c_{j,t}$ ). Her preferences is defined by

$$E_0 \left[ \sum_{j=1}^T \beta^{j-1} u(c_{j,t}, l_{j,t}) \right] \quad (2)$$

where  $u(c, l) = \frac{[c^\theta (1-l)^{1-\theta}]^{1-\sigma}}{1-\sigma}$ ,  $\sigma > 0$ , and  $\theta \in (0, 1)$ .

## 4.2 Endowments

Labor efficiency of an agent consists of two components: an age-specific deterministic efficiency ( $\bar{z}_j$ ) and an uninsurable idiosyncratic labor efficiency shock ( $z_j$ ). The labor efficiency of an age  $j$  agent is specified as

$$e(z_j, j) = \exp(\bar{z}_j + z_j). \quad (3)$$

The labor efficiency shocks are idiosyncratic and they follow the AR(1) process

$$z_j = \rho z_{j-1} + \epsilon_j, \quad \epsilon_j \sim N(0, \sigma_z^2) \quad (4)$$

and

$$z_1 \sim N(0, \sigma_1^2) \quad (5)$$

where  $\epsilon_j$  is a random shock at time  $t$  drawn from a normal distribution with variance  $\sigma_z^2$ . The initial shock,  $z_1$ , is drawn from a normal distribution with variance  $\sigma_1^2$ . Thus, agents supply

---

<sup>5</sup>For the rest of the paper, the first subscript denotes the age of the agent and the second subscript denotes the time period. The time subscript will be dropped later as the paper focuses on the stationary equilibrium.

labor hours  $l_{j,t}$  at the effective wage rate  $w_t$  to earn  $w_t e(z_t, j) l_{j,t}$ .

### 4.3 Assets

Agents are born with no assets and they can hold a risk-free asset in two different accounts: OAH and TDA.<sup>6</sup> Assets in OAH and TDA are identical as a factor of production. Thus, they receive a common rate of return on capital  $r_t$  and are subject to a proportional capital income tax with rate  $\tau_k$ . This capital income tax is used to mimic the corporate profit tax paid by firms. Both OAH and TDA are also subject to zero borrowing constraints

$$s_{j,t}, a_{j,t} \geq 0 \quad \forall j, t \quad (6)$$

where  $s_{j,t}$  and  $a_{j,t}$  denotes the assets in OAH and TDA respectively. OAH is not subject to any contribution or withdrawal constraints.

In contrast, TDA is subject to contribution and withdrawal constraints. Contributions to TDA, denoted by  $q_{j,t}$ , can only be made in agents' working age ( $q_{j,t} \geq 0$  for  $j = 1 \dots R$ ) and assets in TDA can only be withdrawn after retirement ( $q_{j,t} \leq 0$  for  $j = R + 1 \dots T$ ). Since the only differences between IRA and 401(k) are the eligibility status and the contribution limits, the heterogeneity in 401(k) eligibility can be modeled as a stochastic TDA contribution limit. For agents who are not eligible for 401(k), their TDA contribution limit is  $\bar{q}_L$ , which is the statutory contribution limit of IRA. For those who have access to 401(k), they have a high TDA contribution limit  $\bar{q}_H$ , which is the sum of the statutory contribution limits of IRA and 401(k). To capture the fact that only a fraction of agents are eligible for 401(k), a stochastic TDA contribution limit,  $\bar{q}_t \in \{\bar{q}_L, \bar{q}_H\}$  with  $\bar{q}_L < \bar{q}_H$ , is imposed on working-age agents.

Since the model is calibrated to the US economy in 2000 (see section 5), the maximum 401(k) contribution is further limited to 25% of labor earnings. For an age  $j$  agent who is eligible for making 401(k) contribution, her TDA contribution limit is restricted to be 25% of her labor earnings plus the IRA limit or the sum of the IRA and 401(k) statutory limit, whichever is lower. An agent's TDA contribution constraint can be written as

$$q_{j,t} \begin{cases} \in [0, \min(0.25w_t e(z_t, j) l_{j,t} + \bar{q}_L, \bar{q}_H)] & \text{if } j \leq R \text{ and } \bar{q}_t = \bar{q}_H \\ \in [0, \bar{q}_L] & \text{if } j \leq R \text{ and } \bar{q}_t = \bar{q}_L \\ \leq 0 & \text{if } j \geq R + 1 \end{cases} \quad (7)$$

---

<sup>6</sup>This paper focuses on the tax arbitrage opportunities associated with the optimal asset *allocation* problem, and hence maintains a single asset environment for simplicity purpose. In reality, taxation on capital income is incomplete. For example, returns on municipal bonds are exempted from federal and state income tax. The optimal asset *location* problem has been studied extensively in the finance literature (e.g., Amromin, 2003; Dammon, Spatt, and Zhang, 2004; Shoven and Sialm, 2004; Huang, 2008).

The stochastic process of  $\bar{q}_t$  follows a Markov transition matrix conditional on the agent's labor efficiency with  $\Pr(\bar{q}_{t+1} = \bar{q}_H | \bar{q}_t = \bar{q}_H) = \eta_H(e(z_t, j))$  and  $\Pr(\bar{q}_{t+1} = \bar{q}_L | \bar{q}_t = \bar{q}_L) = \eta_L(e(z_t, j))$ .<sup>7</sup> For age one agents,  $\Pr(\bar{q}_t = \bar{q}_H) = \eta_1(e(z_1, 1))$  and  $\Pr(\bar{q}_t = \bar{q}_L) = 1 - \eta_1(e(z_1, 1))$ . The functional form and parameter values of  $\eta_L(\bullet)$ ,  $\eta_H(\bullet)$ , and  $\eta_1(\bullet)$  will further be explained in the calibration section. The total amount of assets in the TDA, denoted by  $a_{j,t}$ , grows as

$$a_{j+1,t+1} = [1 + (1 - \tau_k) r_t] a_{j,t} + q_{j,t}. \quad (8)$$

## 4.4 Taxes

After agents make their contributions to TDAs, the remaining of their income (labor earnings plus the returns on OAH minus TDA contribution) is taxed according to a piecewise linear progressive tax function  $T(\bullet)$ . The marginal tax rates are conditional on the total income of the agents after TDA contribution is deducted. Let  $I = \{I_1, I_2, I_3, I_4, I_5\}$  be the cutoff points of the tax brackets. For an agent with income  $w_t e(z_t, j) l_{j,t} + r_t s_{j,t} - q_{j,t} \in (I_4, I_5]$ , her tax liability is

$$\begin{aligned} T(w_t e(z_t, j) l_{j,t} + r_t s_{j,t} - q_{j,t}) &= \tau_1 (I_2 - I_1) + \tau_2 (I_3 - I_2) + \tau_3 (I_4 - I_3) \\ &+ \tau_4 ([w_t e(z_t, j) l_{j,t} + r_t s_{j,t} - q_{j,t}] - I_4) + \tau_k \cdot r_t s_{j,t}. \end{aligned} \quad (9)$$

The important difference between OAH and TDA is that the returns of assets in TDA are *not* included in agents' taxable income.

In addition to the income tax, agents also pay social security taxes on their labor income at rate  $\tau_{ss}$ . All retired agents receive an equal amount of social security benefits  $b_{j,t}$ , where  $b_{j,t} = 0$  for  $j \leq R$ . Accidental bequest ( $TR_t$ ) is distributed evenly across agents as lump sum transfers.

---

<sup>7</sup>The target of these stochastic 401(k) eligibility processes is to capture the positive correlation between labor earnings and 401(k) eligibility demonstrated in figure 1. This paper assumes that, from the macroeconomic point of view, the labor market is perfectly competitive and firms randomly offer 401(k) to their employees. Since labor efficiency is positively correlated to labor earnings, conditioning 401(k) eligibility on labor efficiency is a reasonable way to deliver the pattern of 401(k) eligibility in the data. An alternative way is to condition the probability on efficient labor supply  $e(z_t, j) l_{j,t}$ . In this sense, labor supply becomes an intertemporal decision because it also affects agent's probability of being 401(k) eligible next period. Given that 401(k) is randomly offered by employers, the former specification in which agents have no control on the 401(k) probability is a more sensible one. There is also a practical consideration in using the former specification. If the probability of 401(k) eligibility depends on labor earnings, then labor supply is no longer an intratemporal decision because it also affects agent's probability of being 401(k) eligible *next period*. It makes the computation of decision rules substantially more difficult.

The budget constraint for an agent of age  $j$  at time  $t$  is

$$c_{j,t} + s_{j+1,t+1} + q_{j,t} \begin{cases} = (1 - \tau_{ss}) w_t e(z_j, j) l_{j,t} + (1 + r_t) s_{j,t} & \text{if } j \leq R \\ -T(w_t e(z_j, j) l_{j,t}, r_t s_{j,t}, q_{j,t}) + TR_t & \\ = (1 + r_t) s_{j,t} - T(0, r_t s_{j,t}, q_{j,t}) + b_{j,t} + TR_t & \text{if } j > R \end{cases} \quad (10)$$

## 4.5 Individual Problem

The decision problem of a new born agent at time  $t$  can be written as

$$\max E_0 \left[ \sum_{j=1}^T \beta^{j-1} u(c_{j,t}, l_{j,t}) \right]$$

*s.t.* (1), (6), (7), (8), (10)

## 4.6 Firms

Markets are competitive. There is a representative firm using capital ( $K$ ) and efficient labor ( $N$ ) as inputs to produce output ( $Y$ ) through a Cobb-Douglas production technology

$$Y_t = K_t^\alpha N_t^{1-\alpha} \quad (11)$$

The firm's profit maximization problem is

$$\max_{K_t, L_t} Y_t - (r_t + \delta) K_t - w_t N_t \quad (12)$$

where  $\delta$  is the depreciation rate, and  $r_t$  and  $w_t$  are the rate of return on capital and the efficient wage rate respectively. The first order conditions of the firm profit maximization problem implies

$$r_t = \alpha K_t^{\alpha-1} N_t^{1-\alpha} - \delta \quad (13)$$

and

$$w_t = (1 - \alpha) K_t^\alpha N_t^{-\alpha}. \quad (14)$$

## 4.7 Recursive Formulation

As this paper focuses on the economy with a stationary equilibrium, the time subscript is dropped whenever possible to maintain simplicity. In this model, agents are heterogeneous in their asset levels in OAH, TDA, TDA contribution limit, and the realization of their idiosyncratic

labor efficiency shocks. The state of an agent can be summarized by  $x = (s, a, \bar{q}, z)$ ,  $x \in X$ , where  $X = \mathbb{R}^+ \times \mathbb{R}^+ \times \{\bar{q}_H, \bar{q}_L\} \times \mathbb{R}$ . The agent's problem can be written recursively in the dynamic programming language

$$V(x, j) = \max_{c, l, s', q} u(c, l) + \beta E[V(x', j + 1)]$$

s.t.

$$c + s' + q = \begin{cases} (1 - \tau_{ss}) we(z, j)l + (1 + r)s - T(we(z, j)l, rs, q) + TR & \text{if } j \leq R \\ (1 + r)s - T(0, rs, q) + b + TR & \text{if } j \geq R + 1 \end{cases}$$

$$l \begin{cases} \in [0, 1] & \text{if } j \leq R \\ = 0 & \text{if } j \geq R + 1 \end{cases}$$

$$q \begin{cases} \in [0, \min(0.25we(z, j)l + \bar{q}_L, \bar{q}_H)] & \text{if } j \leq R \text{ and } \bar{q} = \bar{q}_H \\ \in [0, \bar{q}_L] & \text{if } j \leq R \text{ and } \bar{q} = \bar{q}_L \\ \leq 0 & \text{if } j \geq R + 1 \end{cases}$$

$$a', s' \geq 0$$

$$a' = [1 + (1 - \tau_k)r]a + q$$

In order to specify the model equilibrium, a probability measure  $\varphi_j$  defined on subsets of individual state space ( $X$ ) is used to describe the heterogeneity among agents of age  $j$ . Let the probability space be  $(X, B(X), \varphi)$ , where  $B(X)$  is the Borel  $\sigma$ -algebra on  $X$ . The probability measure must be consistent with the individual decision rules of OAH  $s(x, j)$  and TDA contribution  $q(x, j)$ , and the law of motion of the TDA contribution limit  $\bar{q}$  and the efficiency shock  $z$ . The distribution of individual states across age 1 agents is determined by the joint initial distribution of TDA contribution limit and labor efficiency shock. For agent of age  $j > 1$ , the probability measure is given by the recursion

$$\varphi_{j+1}(B) = \int_X \Pr(x, j, B) d\varphi_j. \quad (15)$$

where

$$\Pr(x, j, B) = \sum_{i \in \{H, L\}} \Pr(\bar{q}', \bar{q}_i) \int \Pr(z', z) dz \quad \text{if } (s(x, j), [1 + (1 - \tau_k)r]a + q(x, j), \bar{q}', z') \in B,$$

$$\Pr(x, j, B) = 0 \quad \text{otherwise.}$$

**Definition 1** A stationary equilibrium in this model is a set of decision rules  $c(x, j)$ ,  $s(x, j)$ ,  $q(x, j)$  and  $l(x, j)$ , factor prices  $r$  and  $w$ , taxes paid  $T(we(z, j)l(x, j), rs(x, j), q(x, j))$ , lump sum

transfer of accidental bequests  $TR$ , social security  $b$ , aggregate capital  $K$ , aggregate efficient labor  $L$ , government consumption  $G$ , a social security tax  $\tau_{ss}$ , a tax regime, and distributions of agents  $\{\varphi_j\}_{j=1}^T$  such that

1.  $c(x, j)$ ,  $s(x, j)$ ,  $q(x, j)$  and  $l(x, j)$  are the optimal decision rules on consumption, next period OAH, TDA contribution, and labor supply respectively.
2. Factor prices are determined competitively, *i.e.*  $w = (1 - \alpha) K^\alpha N^{-\alpha}$  and  $r = \alpha K^{\alpha-1} N^{1-\alpha} - \delta$ .
3. Markets clear

(a) Capital market clears

$$(1 + g) K = \sum_{j=1}^T \mu_j \int_X a(x, j) + s(x, j) d\varphi_j$$

(b) Labor market clears

$$L = \sum_{j=1}^T \mu_j \int_X l(x, j) e(z, j) d\varphi_j$$

(c) Goods market clears

$$C = \sum_{j=1}^T \mu_j \int_X c(x, j) d\varphi_j$$

$$Y = C + (g + \delta) K + G$$

4. Government maintains budget balance

$$G = \sum_{j=1}^T \mu_j \int_X T(w e(z, j) l(x, j), r s_j, q_j) d\varphi_j$$

5. Social security expenditure is equal to the social security tax receipt

$$\tau_{ss} w L = \sum_{j=R+1}^T \mu_j b$$

6. Accidental bequest is equal to transfers

$$TR = \sum_{j=1}^T (1 - \gamma_j) \mu_j \int_X (1 + r) [a(x, j) + s(x, j)] d\varphi_j$$

7. Distributions are consistent with individual behavior as stated in (15).

## 5 Calibration

The model is calibrated to the U.S. economy in 2000, before the tax cuts were implemented by President Bush and substantial changes in the TDA contribution limits. Each model period is equal to one year. Households enter the economy at age twenty-one, retire at age sixty-one ( $R = 40$ ), and live at most to one hundred years old. Thus, households live a maximum of eighty periods and retire after forty periods. The age conditional survival rate is based on the mortality data from the US Census Bureau in 2000, with the conditional survival rate from age one hundred to one hundred and one arbitrarily set to zero. The population growth rate is 0.01, matching the long term U.S. population growth from 1950 to 2000.

There are two components in the labor efficiency process to be calibrated. The deterministic path of household's age-earning profile is taken from the age conditional earnings estimates in Hansen (1993), with linear interpolation for age-efficiency points that are not available. The parameters of the labor efficiency shocks are directly taken from the estimates in Heathcote, Storesletten, and Violante (2008), and the variance of the persistent shock is computed as the average of those year-specific variances from 1991 to 2000 in the same study.

The risk aversion parameter ( $\sigma$ ) is 4. The coefficient of the consumption goods ( $\theta$ ) in the utility function is 0.324, such that the equilibrium average labor supply of workers to be 0.33. These parameters imply that the Frisch elasticity of labor supply of an agent supplying mean hours is about one.

The discount factor, capital share of output, and the depreciation rate are calibrated in the same fashion as in Cooley and Prescott (1995). Due to differences in the model structure, a different treatment is applied on the estimation of capital. First, owners occupied housing is excluded from the stock of capital. The model explicitly specifies the accounts through which agents can hold assets, but it does not specify the types of assets that can be held in each account. Under the single asset assumption in this model, the role of owners occupied housing is not explicitly considered and they are identical to other types of assets. Including owners occupied housing in the notion of capital implies that it can be partially saved in TDA. However, wealth in TDA is mainly composed of financial assets. Although imputed rents are untaxed in nature and the mortgage interests are tax deductible, principle payments are not tax deductible and capital gains from selling a house are subject to capital gain tax. As a result, assets accumulated as owners occupied housing is not tax deferred. This creates an inconsistency on the use of TDAs between the model and in reality. This mismatch is quantitatively important because empirically around 30% of household net worth is held in terms of owners occupied

housing. In particular, the fraction of assets in TDAs will be higher if owners occupied housing is included.<sup>8</sup> For consistency purposes, investment, depreciation, and imputed service flow from owners occupied housing are excluded from relevant calculations.

Second, as the government in this model only consumes output and does not invest in capital, government capital is also excluded from the calculation of capital. According to this calibration, the discount factor is 0.953, set to match a capital-output ratio of 1.87 in the model. The capital share of output ( $\alpha$ ) is 0.334. The depreciation rate ( $\delta$ ) is 0.096.

Tax brackets and the marginal tax rates are taken from the U.S. Internal Revenue Services. Tax brackets are expressed as fractions of average household income reported by the US Census Bureau in 2000. For personal exemptions, households are assumed to be married file jointly and take the standard deduction. Since a significant fraction of households itemizes their tax deductions, the lowest bracket is adjusted from 0.3 to 0.5 to take itemized deduction into account.

For the capital income tax rate, I estimated the capital income tax base by the capital share of output in each year. Then the capital income tax rate is calculated as average percentage of the corporate income tax receipts to capital share of GDP (net of depreciation) from 1987 to 2000. Social security tax is calculated as the average percentage of social security payment (OASDI) to the total compensation to employees and the labor share of proprietary income from 1987 to 2000.

## 5.1 TDA Contribution Limits

The TDA contribution limits are expressed as percentages of average household earnings in 2000. Since this model is not concerned about intra-household decisions and for simplicity purpose, I assume that there are two income earners in the household and both of them share the same 401(k) eligibility status. Thus, the IRA and 401(k) contribution limits are set to be twice of the per worker limits. The low TDA contribution limit is set to be that of IRA, and the high contribution limit is the sum of the IRA and 401(k) contribution limits. The tax brackets and the TDA contribution limits, expressed as fraction of average household income, are summarized in table 1.

Estimating the probability of being eligible of making 401(k) contribution is more complex. In this model, 401(k) eligibility depends on agents' labor efficiency. However, the Survey of Income and Program Participation (SIPP) only reports the percentage of workers participated in 401(k) plans conditional on different categories of labor earnings, which also depends on agents endogenous labor supply decisions. Hence, the conditional probabilities on 401(k) eligibility are

---

<sup>8</sup>Technically, the empirical capital-output ratio increases when owners occupied housing is included in the notion of capital. To reproduce a higher capital-output ratio in the model, the discount factor ( $\beta$ ) has to increase to induce agents to accumulate more assets.

calibrated within the model. First, I assume the conditional probability functions  $\eta_L(\bullet)$  and  $\eta_H(\bullet)$  have logistic functional forms. Specifically,

$$\eta_i(e(z_j, j)) = \frac{1}{1 - \exp(-\pi^i)} \quad (16)$$

where  $\pi^i = \lambda_0^i + \lambda_1^i \ln[e(z_j, j)] + \lambda_2^i \ln[e(z_j, j)]^2$  for  $i \in \{L, H\}$ . With  $\eta_L(\bullet)$  and  $\eta_H(\bullet)$  specified, the probability of initial 401(k) eligibility,  $\eta_1(\bullet)$ , is taken from the stationary distribution conditional on labor efficiencies.

There are six parameters to be estimated within the model. Given the equilibrium wage rate and the optimal labor supply decisions, I simulated a hundred thousand agents and record their labor earnings and 401(k) eligibility, aiming at reproducing the 2001 participation data reported in the SIPP.<sup>9</sup> Since the SIPP categorizes workers' earnings into eight categories and reports the conditional 401(k) participation rate, two adjustments to the earning categories are done to make the results comparable. First, I transformed the earning categories in the SIPP from worker's earnings to household earnings by assuming that it is 1.7 times of the former one. Second, I expressed the household earnings categories as fractions of average household income. The parameters on 401(k) eligibility are estimated to minimize the sum of squared errors between the data and the simulated results.

## 5.2 Summary

Calibration values of the model parameters are summarized in table 3. To sum up, labor efficiency parameters ( $\rho$ ,  $\sigma_1^2$ , and  $\sigma_z^2$ ) are borrowed from the empirical estimates documented in Hansen (1993) and Heathcote, Storesletten, and Violante (2008). Demographic parameters on population growth and survival rates are taken directly from the U.S. data. Production parameters on capital share and depreciation rate are estimated from the NIPA data following the approach in Cooley and Prescott (1995). Income tax rates and tax brackets are taken from the IRS, while the capital tax rate and the social security tax rate are estimated using the data from the Office of Budget and Management. Preference parameters and the parameters for 401(k) eligibility are estimated within the model to deliver the target moments.

---

<sup>9</sup>As pointed out in Poterba, Venti, and Wise (1995), 401(k) eligibility is not the same as its participation. There are various background factors affecting agent's 401(k) participation decision that cannot be captured in the model (e.g., Bayer, Bernheim, and Scholz, 2009; Duflo and Saez, 2003; Madrian and Shea, 2001; Papke and Wooldridge, 1996). Calibrating the model to the 401(k) eligibility data in SIPP will deliver a higher participation rate than the data counterpart. Thus, the 401(k) eligibility in the model is calibrated to the participation rate to deliver a closer match with the data.

## 6 Benchmark Economy

Overall, the benchmark economy provides a good fit to the data. The distribution of 401(k) eligibility produced by the model is consistent with the data. Figure 2 shows that the model is able to capture the hump shape characteristics of the conditional 401(k) eligibility. The extent of TDA usage in the model is also consistent with that in the data. Table 5 shows the amount of capital held in each account. In the model, the fraction of aggregate capital held in TDAs is about 65.9%. Note that owners occupied housing are excluded from the notion of capital in this model. The percentage of capital in TDAs delivered by the model is consistent with the ratio of TDA-to-total financial assets (55.2%) calculated by Bergstresser and Poterba (2004) using the data from the Survey of Consumer Finances in 1998.<sup>10</sup> Descriptive statistics of the benchmark economy is reported in Table 4. The model also does reasonably well in delivering key aspects regarding the tax system. With TDAs and the U.S. tax system, the ratio of government expenditure to aggregate output in the model is 9.4%, which is close to the average of its empirical counterpart from 1987 to 2000 (9.9%).<sup>11</sup>

As shown in Table 6, eligibility to 401(k) has significant impacts on households' labor supply decisions. 401(k) eligible households supply more labor hours than non-eligible households in all age groups. Since 401(k) eligible households have higher contribution limits, they can utilize their TDAs to a greater extent for tax arbitrage. The differences in labor hours are most significant when households first enter the labor force (21-30 year olds) and are about to retire (51-60 years old).

On household savings, 401(k) eligible households accumulate more assets than non-eligible households in all age groups. The percentage difference in asset holdings is largest in the youngest age group, with eligible households saving 3.3 times more than non-eligible households. There are several factors contributing to these results. First, 401(k) eligible households have higher contribution limits that are less likely to be binding. Those eligible households who would *optimally* contribute more than the IRA limit can save more in their TDAs. Second, assets in TDAs grow much faster than that held as OAH because of the favorable tax treatment on capital income in TDAs. Due to the effect of compound growth, the effects of such favorable tax treatment are largest when households are young and lead to the large differences in asset levels in the youngest age group. Third, eligible households are potentially high wage earners

---

<sup>10</sup>These numbers should be interpreted with caution because the ratio reported in Bergstresser and Poterba (2004) only includes households with *both* TDA and non-TDA assets. Furthermore, the notion of capital in this model is not strictly limited to financial assets.

<sup>11</sup>In this model, government consumption is equal to its revenue, which is the sum of revenues from income tax and capital tax. Thus, the empirical government consumption is calculated as the sum of tax receipts from individual income taxes and corporate income taxes. Data is obtained from Table 2.1 in the *Budget of the U.S. Government* published by the Office of Management and Budget.

because they have higher labor efficiencies on average. As a result, they accumulate more assets than their non-eligible, low efficiencies counterparts.

## 7 Effects of TDA on Private Savings

Since TDAs provide tax incentives for households to save in those accounts, it is important to understand the quantitative impacts of TDA on the aggregate economy. The effects of TDAs on private savings is evaluated by steady-state comparison between the benchmark economy and the economy without TDA, *i.e.*  $\bar{q} = 0$ . Incremental savings is defined as the percentage increase in aggregate capital when TDAs are allowed in the economy. The tax brackets and the marginal tax rates are kept at the benchmark levels.<sup>12</sup> A new proportional income tax is introduced with a flat tax rate  $\tau_g$  to maintain government revenue neutrality. It only applies to income above  $I^*$ . Hence, the tax function of the economy without TDA is written as

$$\begin{aligned} T(w_t e(z_j, j) l_t, r_t s_t) &= \tau_1 (I_2 - I_1) + \tau_2 (I_3 - I_2) + \tau_3 (I_4 - I_3) \\ &\quad + \tau_4 ([w_t e(z_j, j) l_{j,t} + r_t s_{j,t}] - I_4) + \tau_k (r_t s_{j,t}) \\ &\quad + \tau_g ([w_t e(z_j, j) l_{j,t} + r_t s_{j,t}] - I^*) \end{aligned} \quad (17)$$

The exemption level of the new tax is chosen to be the same as that in the benchmark model ( $I^* = I_1$ ), implying that the government changes the marginal income tax rates of all income brackets by  $\tau_g$ .<sup>13</sup>

Results on the introduction of TDA are reported in the third column of table 7. Aggregate capital, labor supply, mean hours, and output of the benchmark model are normalized to one, and that of the economy without TDA are expressed as fractions relative to the benchmark model. In general, TDA increases the aggregate capital, but have very little effects on labor supply and mean hours. Consequently, the aggregate output level only increases by 1.2%.

The aggregate capital stock increases by 3.7%. This increase is due to the consumption tax treatment on income contributed to TDAs. First, households can contribute pre-tax income to their TDAs and exploit the tax arbitrage opportunities (difference between current and after-retirement marginal income tax rate). Second, capital income from assets in TDA is free of income tax. It increases the after-tax rate of return on capital and induces households to save more for retirement consumption. On the other hand, an increase in the aggregate capital lowers the rate of return on capital from 8.6% to 8.1%. This equilibrium price effect counteracts the

<sup>12</sup>Average household income in the benchmark model is used to calculate the tax brackets, so that any changes on average income in this counterfactual experiment does not affect the values of the tax brackets.

<sup>13</sup>Experiments on  $I^*$  equals to 0 and 25% average labor earnings are also conducted. Since those results are very similar (less than 0.5% difference) to that of  $I^* = I_1$ , they are not reported in this paper.

tax-incentives from TDA. The composition of capital in Table 5 shows that there is a strong reallocation of assets from OAH to TDA. The amount of assets in OAH decreases by 64.6% when TDA is introduced. Defining new savings from TDA as the increase in capital divided by assets in TDA, only 5.4% of assets in TDA are new savings in the economy.

At the aggregate level, labor supply and mean hours roughly stay the same. On one hand, increase in the capital stock also drives up the wage rate by about 1.2%. The price of leisure rises and the substitution effect induces agents to devote more time to work. Furthermore, TDA contributions enable some agents to move to other tax brackets with lower marginal tax rates. On the other hand, the income effect on leisure causes agents to devote more time to leisure. The income, the tax arbitrage incentives, and substitution effects roughly cancel out each other.

In spite of the small effects of TDA at the aggregate level, Table 6 shows that TDA has important implications on household decisions in different age groups. In the model without TDA, households in first half of their working lives (age 21-40) supply more labor hours than those in the benchmark model, and the reverse is observed in the later half of their working lives (age 41-60). The direction of these differences is robust to households 401(k) eligibilities. In the benchmark economy, households in the two elder age groups supply more labor hours because of the tax deductibility of TDA contribution. Since households in those two age groups on average are more efficient than younger households, they potentially have higher income and face higher marginal income tax rates. By making TDA contributions, they can reduce their taxable income without decreasing their labor hours. Furthermore, returns on capital also affect households marginal tax rates at different age. As households accumulate more assets for retirement, they receive more capital income. That effectively put them into tax brackets with higher marginal tax rates. In the benchmark economy, however, capital income from assets in TDA represents is tax free and is a large part of households' capital income. In all age groups younger than 50, households in the benchmark economy have lower levels of savings than those in the economy without TDA. It is because of the compound growth of assets in TDA and the income effect induced by the higher after-tax capital rate of return.

The new income tax in the economy without TDA is negative, meaning that the government *reduces* the marginal income tax rates by 3.2% to maintain revenue neutrality when TDAs are eliminated. It is because the government collects extra tax revenue without the tax-free treatment on capital return and the tax arbitrage opportunities offered by TDAs. The decrease in marginal income tax rates also have implications on the aggregate effects of TDAs. On aggregate capital, an decrease in income tax rates means that returns on capital in the economy without TDA are subject to lower tax rates, and thus it increases the amount of savings. Also, lower tax rates also encourage agents to increase their labor hours. As a result, the effects of eliminating TDAs are reduced.

## 8 Universal 401(k)

This section is devoted to study the quantitative impacts of heterogeneity of 401(k) eligibility. In particular, I explore the effects of giving all households eligibility in making 401(k) contributions. This experiment expands the TDA system in which all agents have a common TDA contribution limit  $\bar{q}_H$ .<sup>14</sup> Another way to view this reform is that the government imposes heterogeneous IRA limits on households conditional on their 401(k) eligibility, such that all households have a common TDA contribution limit. For agents who do not have access to 401(k), their IRA contribution limits are the sum of the current IRA and 401(k) limit, while the IRA contribution limit for 401(k) eligible agents remains the same. As in the previous section, a new proportional income tax of rate  $\tau_g$  and exemption level  $I_1$  is introduced to maintain government revenue neutrality.

This TDA experiment serves multiple purposes. First, it sheds light on the impacts of raising the TDA contribution limit. With the current hybrid system of income and consumption taxation, one way to move towards consumption tax is by expanding the TDA system. By increasing the contribution limit, the portion of consumption tax also increases. This policy reform can easily be carried out with minimal changes to the tax system. Second, it also removes the heterogeneity in 401(k) eligibility, which imposes uncertainty on households' TDA contribution limits and hence their future income tax rates. Third, this reform reflects the importance of capturing the characteristics of 401(k) eligibility in the data. It shows how much the impacts of TDA are overestimated when a broad definition of TDA that includes IRA and 401(k) is used without considering the pattern of 401(k) eligibility.

Results are reported in the second column of table 7. Overall, removing heterogeneity of 401(k) eligibility has impacts of similar degree to the introduction of TDA. Aggregate capital rises by 4.7%. Note that the magnitude of this increase is comparable to the effects of introducing the TDA system in the economy (in Section 6.1). More assets are reallocated from OAH to TDA (Table 5). Compared with the benchmark model, the amount of assets in OAH decreases by 17.8% and that in TDA increases by 16.5%. A large fraction of incremental TDA assets are new savings in the economy, and most of which are attributed to households with high earning potentials. It is because they are the ones who can be benefited the most from making TDA contributions beyond the IRA limit.

Aggregate labor supply and mean hours only increase by 0.4 and 0.5% respectively. The capital-output ratio increases with universal eligibility, and consequently the interest rate decreases and the effective wage rate increases. Aggregate output also increases by 1.8%. Although

---

<sup>14</sup>Same as the tax brackets, the contribution limit is set with respect to the average household income in the benchmark model, such that the contribution limit is not affected by the macroeconomic impacts of the TDA reform.

providing universal 401(k) eligibility and introducing the TDA system have similar aggregate effects, universal 401(k) eligibility does not have much effects on government revenue. The new proportion tax is only 0.12%, meaning that the government only has to increase the income tax rate marginally to balance its budget.

The importance of capturing 401(k) heterogeneity can be illustrated by comparing the economy with universal 401(k) to that without TDA. The impacts of introducing TDA (with universal 401(k) eligibility) almost doubled. For instance, aggregate capital increases by 8.6% instead of 3.7% in the benchmark analysis. It shows that when a broad notion of TDA, *i.e.* IRA and 401(k), is considered, it is quantitatively important to capture the pattern of 401(k) eligibility in order to have a realistic estimates on the effects of TDA.

The effects of TDA on the household decisions in different age groups (shown in Table 6) intensifies as 401(k) becomes universally eligible. The labor supply of households in the 21-30 and 31-40 age groups decreases and the labor supply of households in the 41-50 and 51-60 age groups rises. Since TDA contributions are tax deductible and households are certain that they can use both IRA and 401(k) to accumulate retirement savings, they save more in later periods of their working lives, at which point they have higher efficiencies on average.

## 9 Tax Reform

The effects of a tax reform is highly sensitive to the conditions of the initial economy. However, studies on tax reforms has taken the progressive tax system at its face value. Results of tax reforms are based on comparisons to baseline economies in which no tax arbitrage opportunity is available. Thus, the effects of a tax reform is potentially overstated when the de facto consumption tax component from TDA is ignored. The purpose of this section is to explore the implications of TDA for a flat tax reform. Here I consider a flat tax reform promoted in [Hall and Rabushka \(1995\)](#) and quantitatively studied in [Ventura \(1999\)](#). By comparing the outcomes of a flat tax reform in the benchmark model to that in a model without TDA (hereafter referred to as the "alternative" model), I examine the extent to which the outcomes of a flat tax reform are affected by the existence of TDA.

With a flat tax reform, the progressive income tax system is replaced by a proportional tax with rate  $\tau_{flat}$  chosen to maintain government revenue neutrality. Only labor earnings above exemption level  $I^*$  is subject to taxation. Investments are tax deductible (or subsidized if net tax payment is negative) and the same treatment applies to both investments in OAH and TDA. Since OAH receives the same tax treatment as TDA but without any withdrawal constraint, OAH is preferred to TDA. Thus, TDA is redundant and can be dropped from the model. The

tax function in (9) is replaced by

$$T = \begin{cases} \tau_{flat} [we(z_j, j)l - \mathbf{I}^* + rs - (s' - s)] & \text{if } we(z_j, j)l > \mathbf{I}^* \\ \tau_{flat} [rs - (s' - s)] & \text{otherwise} \end{cases}.$$

Depending on the exemption level, the new tax scheme can be viewed as a progressive consumption tax ( $\mathbf{I}^* > 0$ ) or a proportional tax ( $\mathbf{I}^* = 0$ ). In the case of a progressive consumption tax,  $\mathbf{I}^*$  is set to be 20% of average household income in the benchmark model. All consumption is subject to taxation under the proportional consumption tax scheme.

Further explanation on model parameter values is needed. If the benchmark model's discount factor ( $\beta$ ) and consumption share ( $\theta$ ) are used in the alternative model, the latter one would not be able to match the target capital-output ratio and mean hours specified in section 3. As the goal of this exercise is to highlight that TDA is an important component ignored in a standard model, I re-calibrated the alternative model so that it matches the same target capital-output ratio and mean hours as the benchmark model. The calibrated parameter values of the benchmark and the alternative model are reported in Table 8. Since the alternative model excludes the favorable tax treatment on TDA, the discount factor has to be higher than that in the benchmark model to achieve the same capital-output ratio. The absence of tax arbitrage opportunity from TDA also leads to higher effective marginal income tax rates for productive agents, so the consumption share in the utility function also has to be higher in the alternative model to offset the work-leisure substitution effect.

Results of the flat tax reform are reported in Table 9. Regardless of the exemption level ( $\mathbf{I}^*$ ) and the existence of TDA, the flat tax reform increases aggregate capital, aggregate labor supply, and output in both the benchmark and the alternative model. These changes are bigger in the case of a proportional consumption tax. Intuitively, since the tax base is larger when there is no tax exemption, the tax rate required to achieve government revenue neutrality is about 2% lower.

More importantly, the quantitative impacts of the flat tax reform in the benchmark model is significantly smaller than that in the alternative model. In the case that  $\mathbf{I}^* = 0.2$ , aggregate capital increases by 27.5 and 37.1% in the benchmark and the alternative model respectively. That means the effects of a flat tax reform on aggregate capital is lowered by 25.9% when the effects of TDA are properly considered in the model. Similar differences are observed in the increase in aggregate labor (16.0%), mean hours (23.0%), and output (20.9%).

## 10 Conclusion

This study investigates the macroeconomic impacts of TDAs and their implications for a flat tax reform. The model in this study includes two key features. First, with endogenous labor supply and a nonlinear tax scheme, this model incorporates the additional tax-saving effect from reallocating taxable income over time through contributions to TDAs. Second, this model also considers the heterogeneity in 401(k) eligibility. Since 401(k) has a much higher contribution limit and is only accessible by a fraction of households, capturing the pattern of 401(k) eligibility is critical in determining the extent of TDA usage in the economy.

Results in this paper show that TDAs moderately increase the aggregate capital by 3.7%. About 65.9% of the aggregate capital is held in TDAs, and most of these assets are shifted from households' ordinary accounts, which would have been saved anyway. TDAs have virtually no effect on the aggregate labor supply. As a result, aggregate output only increases by 1.2%. When access to 401(k) is given to all households, the aggregate capital further increases by 4.7% and the aggregate labor supply only rises by 0.4%, resulting in a 1.8% increase in aggregate output. Note that extending 401(k) eligibility is an expansion of the TDA system, and hence its consumption tax treatment, with little changes in the current tax code. These results show that increasing the IRA contribution limit, or promoting employers to provide 401(k), is a relatively simple way to increase the aggregate output in the economy.

Although TDAs only have moderate impacts on the aggregate economy, their existence significantly changes households' life-cycle labor supply and saving decisions. With households' upward sloping life-cycle efficiency profiles, the tax-saving effects of TDAs induce households decrease their labor supply when young and increase their labor supply when old. These changes in life-cycle labor supply, together with the effect of tax-free compound growth of assets in TDAs, households accumulate less capital when young but they also possess more capital for retirement.

This paper demonstrates that neglecting the consumption tax treatment offered by TDAs exaggerates the progressiveness of the U.S. income tax code, and overestimates the efficiency loss arise from the distortions in labor supply and capital accumulation decisions. Under a flat tax reform, the increase in aggregate capital and aggregate output in an economy with TDA is 25.9% and 20.9% lower than that in an economy without TDAs. These results highlight the importance of considering the effects of tax-favored instruments in tax policy analysis.

Future work on this aspect will include a welfare analysis on the introduction of TDAs. Although TDAs only have moderate impacts on the aggregate economy, their influences on household's life cycle decisions imply that TDAs potentially have important effects on welfare. Furthermore, household responses to changes in the income tax rates should further be investigated.

In this study, only conventional IRA and traditional 401(k) are considered. Since Roth IRA was introduced in 1998 and Roth 401(k) was created in 2006, these Roth accounts have become more popular among households. Instead of allowing households to defer their income tax liabilities, contributions to Roth accounts are made from after-tax earnings and subsequent asset withdrawals from Roth TDAs are tax-free. Future research will explore the implications of the coexistence of Roth and traditional TDAs, and study household choices on their types of TDA. This decision problem is of particular interest when the future tax rates are uncertain.

## References

- ALTIG, D., A. J. AUERBACH, L. J. KOTLIKOFF, K. A. SMETTERS, AND J. WALLISER (2001): “Simulating Fundamental Tax Reform in the United States,” *American Economic Review*, 91(3), 574–595.
- AMROMIN, G. (2003): “Household Portfolio Choices in Taxable and Tax-Deferred Accounts: Another Puzzle?,” *European Finance Review*, 7(3), 547–582.
- AUERBACH, A. J. (1997): “The Future of Fundamental Tax Reform,” *American Economic Review*, 87(2), 143–46.
- AUERBACH, A. J., AND J. SLEMROD (1997): “The Economic Effects of the Tax Reform Act of 1986,” *Journal of Economic Literature*, 35(2), 589–632.
- BAYER, P. J., B. D. BERNHEIM, AND J. K. SCHOLZ (2009): “The Effects of Financial Education in the Workplace: Evidence from a Survey of Employers,” *Economic Inquiry*, 47(4), 605–624.
- BERGSTRESSER, D., AND J. POTERBA (2004): “Asset allocation and asset location: Household evidence from the Survey of Consumer Finances,” *Journal of Public Economics*, 88(9-10), 1893–1915.
- BERNHEIM, B. D. (2002): “Taxation and saving,” in *Handbook of Public Economics*, ed. by A. J. Auerbach, and M. Feldstein, vol. 3 of *Handbook of Public Economics*, chap. 18, pp. 1173–1249. Elsevier.
- COOLEY, T. F., AND E. C. PRESCOTT (1995): *Frontiers of Business Cycle Research* chap. Economic growth and business cycles, pp. 1–38. Princeton University Press, Princeton, NJ.
- DAMMON, R., C. SPATT, AND H. ZHANG (2004): “Optimal Asset Location and Allocation with Taxable and Tax-Deferred Investing,” *Journal of Finance*, 59(3), 999–1037.
- DUFLO, E., AND E. SAEZ (2003): “The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence From a Randomized Experiment,” *The Quarterly Journal of Economics*, 118(3), 815–842.
- ENGEN, E., W. GALE, AND J. SCHOLZ (1996): “The illusory effects of saving incentives on saving,” *The Journal of Economic Perspectives*, 10(4), 113–138.
- FELDSTEIN, M. (1995): “The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act,” *Journal of Political Economy*, 103(3), 551–72.
- GALE, W. G., AND J. K. SCHOLZ (1994): “IRAs and Household Saving,” *American Economic Review*, 84(5), 1233–60.

- GOMES, F., A. MICHAELIDES, AND V. POLKOVNICHENKO (2009): “Optimal Savings with Taxable and Tax-Deferred Accounts,” *Review of Economic Dynamics*, 12(4), 718–735.
- GUISSO, L., T. JAPPELLI, AND M. HALIASSOS (eds.) (2001): *Household Portfolios*. The MIT Press.
- HALL, R. E., AND A. RABUSHKA (1995): *The Flat Tax*. Hoover Institution Press, Stanford, CA, 2nd ed. edn.
- HANSEN, G. D. (1993): “The Cyclical and Secular Behaviour of the Labour Input: Comparing Efficiency Units and Hours Worked,” *Journal of Applied Econometrics*, 8(1), 71–80.
- HEATHCOTE, J., K. STORESLETTEN, AND G. L. VIOLANTE (2008): “The Macroeconomic Implications of Rising Wage Inequality in the United States,” NBER Working Papers 14052, National Bureau of Economic Research, Inc.
- HEER, B., AND A. MAUSSNER (2005): *Dynamic General Equilibrium Models: Computation and Applications*. Overlapping Generations Models, pp. 357–412. Heidelberg: Springer.
- HUANG, J. (2008): “Taxable and Tax-Deferred Investing: A Tax-Arbitrage Approach,” *Review of Financial Studies*, 21(5), 2173–2207.
- HUGGETT, M. (1996): “Wealth distribution in life-cycle economies,” *Journal of Monetary Economics*, 38(3), 469–494.
- IMROHOROGLU, A., S. IMROHOROGLU, AND D. H. JOINES (1998): “The Effect of Tax-Favored Retirement Accounts on Capital Accumulation,” *American Economic Review*, 88(4), 749–68.
- JOULFAIAN, D., AND D. P. RICHARDSON (2001): “Who Takes Advantage of Tax-Deferred Saving Programs? Evidence from Federal Income Tax Data,” *National Tax Journal*, 54(3), 669–688.
- KITAO, S. (2010): “Individual Retirement Accounts, saving and labor supply,” *Economics Letters*, 108(2), 197–200.
- MADRIAN, B. C., AND D. F. SHEA (2001): “The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior,” *The Quarterly Journal of Economics*, 116(4), 1149–1187.
- NISHIYAMA, S. (2011): “The Budgetary and Welfare Effects of Tax-Deferred Retirement Saving Accounts,” *Journal of Public Economics*, forthcoming.
- PAPKE, L. E., AND J. M. WOOLDRIDGE (1996): “Econometric methods for fractional response variables with an application to 401(k) plan participation rates,” *Journal of Applied Econometrics*, 11, 619–632.
- POTERBA, J. M., S. F. VENTI, AND D. A. WISE (1995): “Do 401(k) contributions crowd out other personal saving?,” *Journal of Public Economics*, 58(1), 1–32.

- POTERBA, J. M., S. F. VENTI, AND D. A. WISE (1996): "How Retirement Saving Programs Increase Saving," *Journal of Economic Perspectives*, 10(4), 91–112.
- SHOVEN, J. B., AND C. SIALM (2004): "Asset location in tax-deferred and conventional savings accounts," *Journal of Public Economics*, 88(1-2), 23–38.
- TAUCHEN, G. (1986): "Finite state Markov-chain approximations to univariate and vector autoregressions," *Economics letters*, 20(2), 177–181.
- VENTURA, G. (1999): "Flat tax reform: A quantitative exploration," *Journal of Economic Dynamics and Control*, 23(9-10), 1425–1458.

## Appendix A. Calibration and Equilibrium Computation

The following is an outline of the strategy used to calibrate the benchmark model and the algorithm employed to compute the equilibria in this study :

1. Guess the values of  $\beta$  and  $\theta$  so that the benchmark model will deliver the targeted capital-output ratio and labor hours.
2. Guess the values of  $\{\lambda_i^H\}_{i=0}^2$  and  $\{\lambda_j^L\}_{j=0}^2$  in the probability function specified in equation (16) for an agent's next period 401(k) eligibility.
3. Guess of the steady state values of the aggregate capital ( $K$ ), aggregate efficient labor supply ( $N$ ), and accidental bequest ( $TR$ ).
4. Compute the values of effective wage rate ( $w$ ), rate of return on capital ( $r$ ), social security benefits ( $b$ ), and the calibrated tax brackets cutoff points  $\{\mathbb{I}_i\}_{i=1}^5$  in the benchmark model.
5. Discretize the asset spaces into grid points and approximate the labor productivity shocks in (4) by a Markov transition matrix by the method described in Tauchen (1986).
6. Calculate the optimal decision rules on consumption  $c(x, j)$ , labor supply  $l(x, j)$ , savings as ordinary asset holdings  $s(x, j)$ , and contributions to TDAs  $q(x, j)$  by backward induction.
7. Compute the model-implied aggregate capital stock ( $K$ ), aggregate efficient labor supply ( $N$ ), and accidental bequest ( $TR$ ) in the model.
8. If the model implied values of  $K$ ,  $N$ , and  $TR$  in step 7 equals to the guessed values in step 3, then the algorithm has found a steady state equilibrium given the parameter values in step 1 and 2. Calibration of the benchmark model proceeds to step 9. Otherwise, guess new values and repeat step 4 to 7.
9. Update  $\{\lambda_i^H\}_{i=0}^2$  and  $\{\lambda_j^L\}_{j=0}^2$  to search for the values that minimize the sum of squared error between the model-implied distribution of 401(k) eligibility and its empirical counterpart. Repeat step 3 to 8 for each new guess of  $\{\lambda_i^H\}_{i=0}^2$  and  $\{\lambda_j^L\}_{j=0}^2$ .
10. Update  $\beta$  and  $\theta$  until the benchmark model matches the calibration targets. Repeat step 2 to 9 for each new guess.

Step 1 and 2 are only performed to calibrate the benchmark model. The method used to compute the equilibria, *i.e.* step 3 to 8, is similar to that used in Huggett (1996), Ventura (1999), and Heer and Maussner (2005).

Table 1: Tax brackets, marginal income tax rates, and TDA contribution limits

Labor earnings	Avg. income ( $\chi$ )	Marginal tax rate
$(0, I_1]$	$0.0\chi - 0.5\chi$	.
$(I_1, I_2]$	$0.5\chi - 1.268\chi$	$\tau_1 = 0.15$
$(I_2, I_3]$	$1.268\chi - 2.354\chi$	$\tau_2 = 0.28$
$(I_3, I_4]$	$2.345\chi - 3.326\chi$	$\tau_3 = 0.31$
$(I_4, I_5]$	$3.326\chi - 5.547\chi$	$\tau_4 = 0.36$
$> I_5$	$> 5.547\chi$	$\tau_5 = 0.396$
$\bar{q}_L$	$0.07\chi$	
$\bar{q}_H$	$0.438\chi$	

Table 2: 401(k) participation data in 2001 from SIPP

Earnings Category	401(k) Participation
$< 0.274\chi$	3.9%
$0.274\chi - 0.685\chi$	17.4%
$0.685\chi - 1.369\chi$	44.0%
$1.369\chi - 2.054\chi$	60.7%
$2.054\chi - 2.738\chi$	67.4%
$2.738\chi - 3.560\chi$	77.3%
$3.560\chi - 4.107\chi$	69.2%
$> 4.107\chi$	66.3%

Note: Average household income ( $\chi$ ) is \$58,208.4 in 2001.

Table 3: Parameter values of the calibrated model

Parameters	Name	Value	Target / Data Source
$T$	max period	80	max age = 100
$R$	last period of work	40	retire at age 61
$\beta$	discount rate	0.9553	capital-output ratio = 1.873
$\sigma$	risk aversion parameter	4	
$\theta$	consumption coefficient	0.324	match labor supply = 0.333
$g$	population growth rate	0.01	Avg. growth 1950-2000
$\{\gamma_j\}$	survival probability		2000 mortality rate from SSA
$\{\bar{z}_j\}$	deterministic ability		Hansen (1993)
$\rho$	persistence of ability shocks	0.9733	Heathcote, Storesletten, and Violante (2008)
$\sigma_z^2$	variance of the random term	0.0176	Heathcote, Storesletten, and Violante (2008)
$\sigma_1^2$	var of initial distribution	0.1242	Heathcote, Storesletten, and Violante (2008)
$\alpha$	capital share of output	0.3342	Avg. capital share 1959-2000
$\delta$	depreciation rate	0.0960	Avg. depreciation 1959-2000
$\tau_{ss}$	social security tax	0.1064	Estimated from OMB data
$\tau_k$	tax on capital income	0.0984	Estimated from OMB data

Table 4: Descriptive statistics of the benchmark economy

K/Y	N	Mean hours	Y	r	w	G/Y
1.873	0.370	0.33	0.503	0.082	0.915	0.094

Table 5: Composition of Capital

	Benchmark	Without TDA	Universal 401(k)
K	1.000	0.964	1.047
$K_{TDA}$	1.000	0.000	1.165
$K_{OAH}$	1.000	2.831	0.822
$K_{TDA}/K$	0.659	0.000	0.733

Note: K refers to the amount of aggregate capital, which is the sum of assets in TDA and OAH.  $K_{TDA}$  is the amount of capital in TDAs.  $K_{OAH}$  is the amount of capital held as OAH. K,  $K_{TDA}$ , and  $K_{OAH}$  in the benchmark model are normalized to one. Values in the experiments are expressed as fractions relative to their counterparts in the benchmark model.

Table 6: Decisions of households in different age groups

	21-30 yr	31-40 yr	41-50 yr	51-60 yr
Labor hours per worker				
Benchmark	0.357	0.358	0.323	0.274
401(k) eligible	0.365	0.359	0.324	0.283
401(k) non-eligible	0.353	0.356	0.322	0.267
Without TDA	0.371	0.361	0.315	0.254
Universal 401(k)	0.353	0.354	0.329	0.278
Assets per worker				
Benchmark	0.092	0.568	1.207	1.806
401(k) eligible	0.174	0.722	1.478	2.283
401(k) non-eligible	0.052	0.421	0.932	1.365
Without TDA	0.130	0.708	1.338	1.658
Universal 401(k)	0.079	0.561	1.263	1.970

Note: Benchmark, without 401(k), and universal 401(k) show the average values per worker in the three different scenarios. For the benchmark case, households are further divided into those who are eligible and non-eligible for 401(k). Average values conditional on 401(k) eligibility are reported separately in rows beneath the benchmark case.

Table 7: Model Results with Different TDA Contribution Limits

	Benchmark	Without TDA	Universal 401(k)
K	1.000	0.964	1.047
N	1.000	1.000	1.004
Mean hours	1.000	0.997	1.005
Y	1.000	0.988	1.018
r	0.081	0.086	0.076
w	0.915	0.904	0.928
$\tau_g$		-3.20%	0.12%

Note: Aggregate capital (K), aggregate efficient labor (N), mean hours, and output (Y) in the benchmark model are normalized to one. Values in the experiments are expressed as fractions relative to their counterparts in the benchmark model.  $K_{TDA}/K$  refers to the fraction of aggregate capital held in TDA.

Table 8: Parameter values of the benchmark and alternative models

	Benchmark (with TDA)	Alternative (without TDA)
$\beta$	0.955	0.961
$\theta$	0.324	0.330

Table 9: Flat tax reform in the benchmark and the alternative model

	Benchmark model				Alternative model			
	Progressive	$\mathbf{I}^* = 0.4\chi$	$\mathbf{I}^* = 0.2\chi$	$\mathbf{I}^* = 0$	Progressive	$\mathbf{I}^* = 0.4\chi$	$\mathbf{I}^* = 0.2\chi$	$\mathbf{I}^* = 0$
K	1.000	1.197	1.221	1.242	1.000	1.336	1.366	1.390
N	1.000	1.081	1.108	1.130	1.000	1.096	1.128	1.150
Mean hours	1.000	1.042	1.074	1.100	1.000	1.061	1.100	1.130
Y	1.000	1.119	1.145	1.167	1.000	1.171	1.202	1.225
$\tau_{flat}$		15.15%	12.29%	10.39%		17.44%	14.17%	12.00%

Note: Columns with header "Progressive" denote the economies with progressive income tax  $T(\bullet)$  and proportional capital tax  $\tau_k$ . The average household income in these economies are denoted as  $\chi$ . Aggregate capital (K), aggregate efficient labor (N), mean hours, and output (Y) in those economies are normalized to one. Values in the flat tax economies are expressed as fractions relative to their counterparts in the progressive income tax economy.

Figure 1: 401(k) participation conditional on labor earnings in 2001

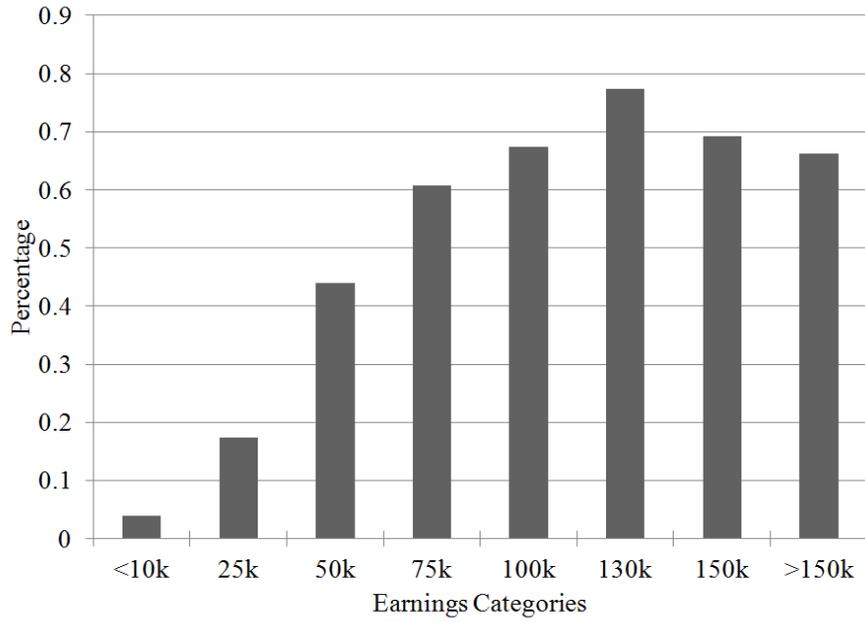


Figure 2: 401(k) participation in the model and the 2001 SIPP data

