An Equilibrium Model of Pension Provision and Wage Determination*

Fumihiko Suga†

February 29, 2016

Abstract

This paper proposes and estimates a structural model of screening by employer-provided pension. The model can explain why firms offer contributions to pension plans and why workers with pension plans earn more and stay longer with their employer than those without pension. In the model, there are firms with a fixed cost of hiring a worker, and they offer contributions to workers' pension plans in conjunction with low human capital rental rates in order to screen out "impatient" workers; low human capital rental rates discourage impatient workers from taking the offer, while pension benefits attract "patient" workers. Although this paper focuses on 401(K) plans, the basic idea can be applied to other forms of employer-provided pensions as well; this is one of the biggest advantages of the idea of screening by employer-provided pensions. The model is estimated by using U.S. data. The estimation results show how workers are sorted into jobs with pension and without pension according to their time preferences and human capital levels. The result of a counter-factual simulation in which the government eliminates tax benefits from 401(K) plans suggests that the pension balance of a worker with pension at age 61 declines by 50%, and the number of workers working for the pension providing firms decrease by 14%. So the tax benefit of the 401(K) plans is important especially for retirement asset accumulation. In addition, the result of a simulation with flat income tax suggests that the number of workers with employer-provided pension declines by 82%. It indicates that it is important for firms and workers to save tax payment by flattening income over the life-cycle by taking advantage of pension plans.

Keywords: Pension, Job-market screening
JEL codes: D31, D82, J21, J22, J32, G23

*I am grateful to my adviser John Kennan and Christopher Taber for helpful comments and suggestions. I appreciate that Chao Fu let me use her computational resources. I thank workshop participants at the University of Wisconsin-Madison, National Graduate Institute for Policy Study, Hitotsubashi University, the University of Tokyo (in July 2014 and February 2015), Aoyama Gakuin University, Kyoto University, and Osaka University. All errors that remain are mine.

†Economic and Social Research Institute, Cabinet Office, Government of Japan. f.suga37@gmail.com

1 Introduction

Employer-provided pensions are quite common in the United States. More than half of the employed workers in the United States have employer-provided pension. Despite its pervasiveness, the role of employer-provided pension is not quite clear. Why do firms offer pension plans? Why are there workers with pension and without pension? What is the difference between them? Further, how are employer-provided pensions related to wages? A simple regression analysis suggests that workers with pension plans tend to earn more and stay longer with an employer than those not having pension plans. Why is this so? This paper proposes and estimates a structural model to answer these questions. The results of counter-factual simulations show what happens in wage distribution, asset accumulation, and the rate of participation in pension plans when the government changes the tax codes. In the model presented in this paper, employers offer attractive pension plans to hire workers with low discount rates (low discounters or "patient" workers); here, employer-provided pension plays the role of a screening device.

The low discounters are those who place a high value on their future income and a low value on their present consumption or leisure. Therefore, firms hiring workers can screen out high discounters by offering a generous contribution to pension plans in conjunction with a relatively low wage rate; the pension benefits that become available only after retirement attract the low discounters, whereas the low wages offered today discourages the high discounters from accepting such offers. The idea of screening out workers through employer-provided pensions was first proposed by Ippolito (1992). Ippolito focused on a particular feature of a typical 401(K) plan by which employers and employees deposit money in employees' pension accounts; in many cases, employers make contributions that "match" their employees' contributions. In 2012, around 80% of the workers' with tax-deferred pension plans reported that their pension plans had this feature. Now, the question that arises is, why do employers link their contributions to their employees' contributions? Ippolito conjectured that firms offer such an apparently in-favor-of "savors" plans in order to attract the low discounters who are supposed to be more inclined to save money.

The question then is, why are firms more interested in low discounters? Ippolito collected some statistics and regression results to show the connection between several discount rate measures of workers and their working behavior. Ippolito found that those supposed to be low discounters

tend to work harder and stay longer with their employer. For example, in an imperfect monitoring environment, firms have an incentive to hire those who are not likely to shirk work. Since the low discounters place a lower value on their present leisure and a higher value on their expected future income, they are less likely to under-perform and more likely to work harder to keep off the possibility of being fired. For another example, if firms have to train their workers and the training cost is constant for all workers, they have an incentive to hire workers who work harder and stay longer. Several studies (e.g., Dorsey and Macpherson, 1997) show that pension-providing firms are more likely to provide training to their workers. Thus a training cost in an equilibrium pension-providing and wage-determining model is not a far-fetched assumption.

A big advantage of the idea of screening workers by pension is that different types of pension plans play the same role. Employer-provided pensions in the United States can be broadly categorized into two: the defined benefit (DB) plans and the defined-contribution (DC) plans. While the former comprise the traditional pension plans, the latter include currently the most common type of pension plans in the United States. The 401(K) plans are categorized under DC plans. These plans are so different that it is hard to regard them as playing the same role. A DB plan determines the pension benefits of workers according to their tenure, wage, and age, whereas a DC plan determines pension benefits according to how the employers and employees contribute to pension plans. Screening by pension is one of the few hypotheses that explain the role of DB and DC pension plans harmoniously such that they play the same roles[1]. This is important especially in the United States where there has been a dramatic shift from the DB to DC plans since the 1980s, although the rate of participation in pension plans has not changed so dramatically. Furthermore, there is an empirical finding that can be explained only by sorting/screening by pension: the job turnover rate of workers with DC plans is much lower than the job turnover rate of workers without pension even when all the related observed variables are controlled for. Since DC pensions do not influence workers to stay longer with their current employer, this should be a consequence of sorting/screening workers by pension; workers with pension plans stay longer with an employer

[1]Another approach to employer-provided pension in which DB pensions and DC pensions play the same role is as follows: If the income tax rate is progressive and the income after retirement is much lower than the income before retirement, employers and employees can be better off by carrying over a part of their pre-retirement income to periods after their retirement by contributing to pension plans. Since contributions to pension plans are tax-deferred (the benefit is taxed), employees can save tax payments by lowering their taxable income when they are young. In order to check to what extent this matters, we carry out a counter-factual simulation in which the government imposes a flat-rate income tax instead of a progressive income tax rate.

not because their pension plans give them an incentive to stay, but because "patient" workers are likely to be sorted into jobs with pension plans.

This paper proposes and estimates an equilibrium model in which 401(K) plans play the role of a screening device. The model works in the following manner: low discounters tend to work harder and accumulate more human capital stock through learning-by-doing; they place a lower value on their present leisure and a higher value on their future wages. In addition, a transitory shock exists on the disutility from working that can be reset by changing employers; therefore, those who receive a negative shock have an incentive to change employers. However, a job switch can cause human capital deterioration. So there is a trade-off between today's leisure and human capital stocks. Since a utility shock is transitory whereas human capital stock is critical for the rest of their lives, low discounters are less likely to change their employers than are high discounters. Thus, low discounters put in a larger amount of labor (human capital stock times hours of work times years of tenure) than high discounters do. On the employers' side of the model, some firms may have to pay a fixed cost of hiring a worker for production, such as the cost of training. These firms have an incentive to hire fewer workers with higher human capital levels, more working hours, and longer tenure. Therefore, they prefer low discounters, but cannot observe the workers' time preferences because of asymmetric information. Thus, it is beneficial for firms with a fixed cost of hiring to offer employer-matching contributions and human capital rental rates that are lower than the rates offered by the other type of firms so that the low discounters are attracted by pension benefits, whereas the low wages offered today discourage the high discounters from accepting their offers. The model has only 401(K) plans, but its basic mechanism can be applied to other types of deferred compensations as well, such as DB pensions, non-401(K) DC pensions, and hybrid plans[2].

The model is estimated by using the data from Survey of Income and Program Participation (SIPP). SIPP data comprise detailed individual information on types of pension plans, contributions to pension plans made by employers and employees, and balance of pension assets as well as hours of work, income, and non-pension assets. The estimation results show how people are sorted into jobs with pension and without pension according to their time preferences and human capital level; low discounters are likely to take up jobs with pension plans, but some low discounters take up

[2]A hybrid pension plan is a type of plan that has the feature of DB pensions and DC pensions. Some firms convert the DB pensions of their existing employees to hybrid plans and offer their newly hired employees DC pension.

jobs without pension plans because their human capital levels are too low and vice versa. The result of a counter-factual simulation in which the government eliminates tax benefits from 401(K) plans suggests that the pension balance of a worker with pension at age 61 declines by 50%, and the number of workers working for the pension providing firms decrease by 14%. So the tax benefit of 401(K) is important for retirement asset accumulation. In addition, the result of a simulation with flat income tax suggests that the number of workers with employer provided pensions declines by 82%. It indicates that it is important for firms and workers to save tax payment by flattening income over the life-cycle taking advantage of pension plans.

The rest of this paper is organized as follows. Section 2 summarizes the background with a brief description of previous literature. Section 3 presents the screening model, and Section 4 describes the data. The estimation method and identification strategy are covered in Section 5. Section 6 gives the estimation results, and Section 7 presents counter-factual simulation results. The last section summarizes our findings and concludes the paper.

2 Background and Related Literature

The literature has several hypotheses on the role of employer-provided pensions. When DB pensions had dominated the market, employer-provided pensions were mostly regarded as an "incentive contract." A typical DB pension has a "back-loading" structure in the sense that benefit accumulation is slow for the first ten to twenty years of employment, but suddenly accelerates right before the "normal" retirement age, the age at which full pension benefits become available[3]. The back-loading structure of benefit accumulation discourages workers from leaving their current employers and raises their fear of being fired before their normal retirement age. Thus, employees stay longer with their current employer and are less likely to under-perform. Furthermore, benefit accumulation stops and sometimes even reverses after the normal retirement age of a worker, leading to firms deciding the timing of their employees' retirement. The previous literature on the role of pension plans prior to the 1990s emphasizes on these DB pension incentive effects. The idea of considering employer-provided pension as an incentive contract, however, can be applied only to DB pensions.

[3]Note that the definition of normal retirement age is different from what it means for public pension or the social security system.

In the United States, DC pensions, especially 401(K) plans, have been replacing DB pensions since the 1980s; DC pensions do not have the back-loading structure or show decreasing benefits after the normal retirement age. Although DC pensions are supposed not to have the incentive effect, several previous empirical studies have found workers with DC pensions tending to work longer than those having no pension. Therefore, it is still not clear whether the longer tenure of workers with DB pensions is due to incentive effects.

It is more difficult to explain the role of DC pensions, including 401(K) plans, because they do not seem to directly affect the behavior of workers. If a worker starts working for an employer offering 401(K) plans, an individual account will be set up for the worker, and the employee, employer, or both of them would jointly contribute to the worker's pension plan. DC pension accounts are mobile and considered to belong to the worker; if a worker changes employer, he/she is allowed to roll over the money remaining in the pension account to a new pension account with the new employer if they offer pension, or to an individual retirement account (IRA) otherwise. Basically, unlike with DB pension, there is no penalty for a job switch in DC pensions, which do not have a direct influence on the timing of retirement. Thus, DC pensions are not considered to have an incentive effect. If this is so, why have many firms switched from DB pensions to DC pensions? Some previous studies have tried to answer this question by focusing on the tax advantage of DC pensions: income tax is not imposed on contributions to pension plans and the earnings accrued on pension accounts until the pension benefits become payable. Workers need to pay income tax only when they receive pension benefits. Since such earnings are exempt from income tax, pension assets grow rapidly; furthermore, the income tax that workers pay after retirement will be low because employees receive their pension benefits in the form of a monthly annuity. It is possible that the employers as well as employees are better off by taking advantage of this tax benefit of DC pensions, but the idea of treating employer-provided pensions as a tax savings device through deferred tax payment cannot be applied to DB pensions[4].

The previous literature on the role of employer-provided pensions focuses on the incentive effects (e.g., Lazear, 1979) and tax benefits of DC pensions, and pays less attention to the sorting/screening effect of pensions. Several papers show empirical results supporting the idea of sorting or screening of workers by pension, most of them focusing on the statistics or regression

[4]The tax advantage of flattening lifetime income can be applied to both types of pensions.

results showing smaller turnover rates for workers with DC pensions than for those with no pension (e.g., Ippolito, 1997, 2002; Even and Macpherson, 2005). Since a typical DC plan cannot keep the workers from leaving their current employers, the longer tenure of workers with DC pensions can be a consequence of workers being sorted/screened by pension; for example, low discounters, or patient workers, are sorted into jobs with pensions, and they work longer with an employer. In the model presented in this paper, workers with a pension plan work longer because low discounters are likely to be sorted into jobs with pension plans. Since the tenure gap between workers with pension and without pension is difficult to be explained by any other factor, this model can help identify the distribution of discount rates. In other words, we can identify the distribution of the discount factor because the tenure gap between workers with pension and without pension cannot be generated by a model with homogeneous discount rates.

The basic structure of the workers' side of the model consists of a life-cycle model with human capital and pension asset accumulation. Heiland and Li (2012) estimate a life-cycle model with pension and non-pension assets and explain the increase in labor supply of older workers since the mid-1990s by the shift from DB to DC pensions. Since Heiland and Li's (2012) model is purely a partial equilibrium model, it cannot explain why employers offer pensions and how workers are sorted into jobs with pension and without pension.

The model is estimated by using the method of simulated moments (McFadden, 1989). Gourinchas and Parker (2002) is a pioneering work of estimating a life-cycle model by using the method of simulated moments. Similar estimation methods are employed by French (2005), French et al. (2011), and Laibson et al. (2007) etc. These papers estimate a partial equilibrium model, while the wage rate is determined endogenously in the model presented in this paper. Lee and Wolpin (2006) estimate a life-cycle model with wages determined endogenously. They find an equilibrium wage function by trying numerous candidate wage functions for each candidate parameter set. We take a similar approach to find the equilibrium human capital rental rate function.

This paper is the first attempt in the literature to estimate a life-cycle model in which firms offer contributions to employees' pension plans for the purpose of screening workers. The model is unique in the sense that it illustrates how workers are sorted into different jobs according to a "deep" parameter such as the discount factor.

3 Model

The basic setup of the workers' side of the model is similar to Imai and Keane's (2004) model. The process of human capital accumulation in our model is simpler than that of Imai and Keane (2004) [5], but our model is more sophisticated in that it has pension and non-pension asset accumulation as well as human capital accumulation, and workers can choose their employers. In addition, the model describes how pension-providing firms determine the wages they offer in equilibrium. The workers' side of the model is characterized by heterogeneity in discount rates, borrowing constraints, human capital accumulation through learning-by-doing, human capital depreciation due to job switch, transitory shocks on disutility from working that can be avoided by job switch, and the accumulation of pension and non-pension assets. The employers' side of the model is characterized by two types of firms, those with and those without a fixed cost of hiring a worker. Asymmetric information exists with respect to workers' discount rates, human capital stocks, and hours of work.

This model works in the following manner: workers with low discount rates (low discounters) place a higher value on future income and a lower value on present leisure, and so they work harder and accumulate more human capital stock compared to high discounters when they are young so as to enjoy a high wage after they get older. In addition, low discounters are less likely to change their employers because of transitory shocks on the disutility from working and human capital deterioration caused by job switch. This is because low discounters give more importance to human capital accumulation over the rest of their lives than the transitory utility shocks that matter only for the current period. As a result, low discounters put in more labor (human capital stock times hours of work times years of tenure) and produce more with an employer. On the employers' side, firms with a fixed cost of hiring a worker have an incentive to hire hardworking workers with larger human capital stock, because the hiring cost is constant but the amount of labor supply differs across workers. Therefore, these firms prefer low discounters, but the discount rates of workers are not observable to firms owing to asymmetric information. Thus, firms with a fixed cost of hiring contribute to pension plans in conjunction with offering human capital rental

[5] This is because identification of the parameter in the human capital accumulation equation relies on the transition of not hours of work and wage rate but annual labor income only. In the model, we assume that hours of work and wage rates cannot be observed separately, and we match the moment of labor income at each period.

rates that are lower than those offered by the other type of firms in order to screen out the high discounters by taking advantage of the difference in values they place on present leisure and future income (pension benefits). The advantage of this model is that it can be applied to other forms of deferred compensation such as DB pensions and non-401(K) DC pensions, although the model presented in this paper covers only 401(K) plans.

This model is unique in the sense that its basic structure consists of a life-cycle model but it illustrates how the equilibrium human capital rental rates are determined when there is a fixed cost of hiring a worker. The fixed cost of hiring in the model is a training cost, and firms have to incur a training cost only once at the beginning of employment.

From this chapter onward, we refer to firms with a fixed cost of hiring as pension firms and the other type of firms as non-pension firms, even though pension-provision is an equilibrium outcome [6].

3.1 Workers' Problem

Workers live for T periods, but can work only during the first $T-1$ periods. In their last period, workers do not work but receive social security and pension benefits. The total number of time periods T is eight, and the length of one period is six years [7]. The last period lasts for twelve years and everyone dies after the last period.

Preference:

The current payoff function at period $t < T$ is

$$U(C_{it}, h_{it}) = \frac{(C_{it} - \underline{C}_t)^{1-\gamma}}{1-\gamma} - \nu_{it}\frac{h_{it}^{1+\eta}}{1+\eta} \quad \text{if} \;\; t < T \tag{1}$$

where C_{it} is the consumption at t, h_{it} is hours worked, and $\underline{C}_t$ is the minimum consumption level. The minimum consumption level depends on age t as follows:

$$\underline{C}_t \;\; = \lambda_1 + \lambda_2 t \qquad \text{if} \;\; t < T$$

[6] We estimate the model assuming that firms with a fixed cost of hiring a worker always offer pension. However, the participation constraint is checked after estimation.

[7] 24 to 29, 30 to 35, 36 to 41, 42 to 47, 48 to 53, 54 to 60, 61 to 66, and 67 and after

Now, let q_{it} denote the indicator of a worker having a new employer and $q_{i1} = 1 \ \ \forall i$. The coefficient of the disutility from working ν_{it} in Equation (1) depends on for whom worker i works at period t; if worker i starts working for a new employer ($q_{it} = 1$), the disutility from working is always low ($\nu_{it} = \underline{\nu}$). If the workers do not change their employer, ν_{it} is stochastic:

$$\begin{aligned} \nu_{it} &= \overline{\nu} \ \text{ with probability } \ \phi \\ \nu_{it} &= \underline{\nu} \ \text{ with probability } \ 1-\phi \end{aligned}$$

where $\overline{\nu} > \underline{\nu}$. Thus workers have an incentive to change their employers when the ν becomes high, although their human capital stock would deteriorate if they change their employers (as explained later).

Since the last period is different from the other periods in that it has a different length and could have bequest motives, the utility from consumption at T takes the following form which is different from the current payoff function of previous periods:

$$U(C_{iT}, h_{iT}) = \mu \frac{C_{iT}^{1-\gamma}}{1-\gamma}$$

Constraints:

The budget constraint of worker i at $t < T$ is given by

$$\begin{aligned} C_{it} + S_{it+1} &\leq (1 + (1-\tau_{it})r)\, S_{it} + (1-\tau_{it})\,(1-I_{it})\, Y_{it}, \quad S_{t+1} \geq 0 \\ C_{it} + S_{it+1} &\leq (1+\overline{r})\, S_{it} + (1-\tau_{it})\,(1-I_{it})\, Y_{it}, \quad S_{t+1} < 0 \\ S_{it+1} &\geq \underline{S} \end{aligned}$$

where S_{it} is savings or the non-pension assets at t ; τ_{it} is the income tax rates which is a function of income[8]; I_{it} is the rate of investment in the pension plans out of pre-tax income; r is the rate of market returns; and Y_{it} is labor income, which depends on human capital stock H_{it}, hours of work h_{it} and an indicator of working for a pension-providing firm p_{it}. It is assumed that workers are not allowed to borrow against their pension assets, and that $I_{it} \geq 0$. When worker i works for

[8]The income tax rate is progressive and depends on income.

a non-pension firm, I_{it} is the rate of investment in his/her IRA from labor income. The budget constraint of worker i at $t = T$ is

$$\left(1 + \left(1 - \tau_{iT}\right) r\right) S_{iT} + \left(1 - \tau_{iT}\right) \left(B_{iT} + P_{iT}\right) - C_{it} \quad \geq 0$$

where B_{iT} stands for social security benefit and P_{iT} pension asset. It is assumed that workers receive retirement income in the form of a monthly annuity and that social security and pension benefits are received over a period of 12 years. Note that workers are not allowed to die in debt.

Let $\tau\left(Y\right)$ be the income tax rate when the income is Y. Then the income tax rate for period $t < T$ is given by

$$\tau_{it} = \tau\left(rS_{it} + \left(1 - I_{it}\right) Y_{it}\right).$$

Note further that the last period is twice as long as the previous periods. For simplicity, it is assumed that the income tax rate in the last period is

$$\tau_{iT} = \tau\left(\frac{r * S_{iT} + B_{iT} + P_{iT}}{2}\right).$$

Pension Asset:

The accumulation of pension assets depends on the indicator of working for a pension-providing firm, p_{it}. Workers decide whether or not to work for a pension-providing firm at each period. The pension balance P_{it} is given by

$$P_{it} = \sum_{s=1}^{t-1} (1 + r)^{t-s} \left[p_{is} \psi\left(I_{is} Y_{is}\right) + \left(1 - p_{is}\right) I_{is} Y_{is}\right]$$

where $\psi\left(I_{is} Y_{is}\right)$ is the sum of worker i's contribution and his/her employer's contribution when the worker works for a pension-providing firm. The total contribution to worker i's pension account is

$$\psi\left(I_{is} Y_{is}\right) \quad = \quad \begin{cases} (1 + 0.898) I_{is} Y_{is} & \text{if } \ I_{it} < 0.05 \\ (1 + 0.898) 0.05 Y_{is} + (1 + 0.545)\left(I_{it} - 0.05\right) Y_{is} & \text{if } \ I_{it} \geq 0.05 \text{ and } I_{it} < 0.1 \\ (1 + 0.898) 0.05 Y_{is} + (1 + 0.545) 0.05 Y_{is} + \left(I_{it} - 0.1\right) Y_{is} & \text{otherwise} \end{cases}$$

The rate of matching contribution is calculated directly from the data [9]. The pension contribution function $\psi(.)$ is given, and so firms cannot manipulate the rate of their matching contribution to pension plans[10]. In order to see how sensitive the model is to this assumption, two simulations with different $\psi(.)$s are carried out with the estimated model for a robustness check. Note that earnings on pension assets are not taxed at period t. The US government has set an upper bound on contributions to pension plans; the upper bound of contributions to employer-provided pension plans is different from that of contributions to IRAs:

$$I_{it}Y_{it} < \overline{I}(p_{it}).$$

Income:

Pre-tax labor income is a function of human capital H_{it}, hours of work h_{it}, and the indicator of working for a pension-providing firm p_{it}:

$$Y_{it} = \begin{cases} \overline{w}H_{it}h_{it} & \text{if } \ p_{it} = 0 \\ \tilde{w}\left(H_{it}h_{it}\right)H_{it}h_{it} & \text{if } \ p_{it} = 1 \end{cases}$$

where $\overline{w}$ is the human capital rental rate offered by non-pension firms and $\tilde{w}\left(H_{it}h_{it}\right)$ is the human capital rental rate offered by pension firms. It is assumed that firms cannot observe H_{it} and h_{it} separately, and that they offer a human capital rental rate based on the amount of labor supply $H_{it}h_{it}$ [11]. This assumption can be interpreted as an assumption that hours of work in the model indicate the "intensity" of work, which is different from the observed hours of work, and that firms observe only the amount of labor input by each worker.

There is a trade-off between human capital rental rates and employers' contributions to the pension plans when $\overline{w} > \tilde{w}\left(H_{it}h_{it}\right)$, and the human capital rental rate function $\tilde{w}\left(.\right)$ is determined

[9]The rate of matching contribution is calculated from the ratio of employers' contributions to employees' contributions for workers whose rate of contribution out of wages is less than 5% and 5% to 10% respectively.

[10]It is assumed that $\psi(.)$ is fixed because different combinations of $\psi(.)$ and human capital rental rate$\tilde{w}(.)$ can yield the same profit, which can cause an identification problem.

[11]It is computationally very burdensome to estimate the model if we assume that employers observe the human capital levels and hours of work separately. If H_it is observable, firms can infer workers type by referring to the human capital levels. Then $\tilde{w}(.)$ will be a function of $H_{it}h_{it}$ and H_it, and we have to try candidate $\tilde{w}(.)$ function for different levels of human capital.

in equilibrium.

Human Capital:

Workers can avoid the negative shock of disutility from working by changing jobs, but part of the human capital stock will be lost. The human capital depreciation due to a job switch can be interpreted as a loss of firm-specific human capital. Note that the fraction of human capital loss due to job switch is not tenure dependent. Since incorporating tenure as a state variable is computationally burdensome, it is assumed that the fraction of the firm-specific human capital is age-dependent and decreasing in age:

$$\begin{aligned} H_{it} &= D_t \tilde{H}_{it} \quad \text{if} \quad t > 1 \text{ and } q_{it} = 1 \\ H_{it} &= \tilde{H}_{it} \quad \text{otherwise} \\ D_t &= \delta_0 - \delta_1 t - \delta_2 t^2 \end{aligned}$$

where $\tilde{H}_{it}$ is the human capital stock carried over from previous periods, and H_{it} is the stock after the workers decide on whether or not to stay with the same employer. Note that labor income at period t depends on H_{it}, while the state variable which workers make decision based upon is $\tilde{H}_{it}$. Human capital accumulates in the learning-by-doing way. Human capital $\tilde{H}_{it}$ evolves according to the following transition equation:

$$\begin{aligned} \ln \tilde{H}_{it+1} &= \ln \left(H_{it} + \zeta \left(h_{it} - \underline{h} \right) \right) + \epsilon_{it+1} \quad \text{and} \quad \epsilon_{it+1} \sim \mathrm{N}(0, \sigma_\epsilon) \\ \ln H_1 &\sim N \left(\mu_{H_1}, \sigma_{H_1} \right) \end{aligned}$$

Shock ϵ_{it+1} is realized at the end of period t or at the beginning of period $t+1$, before the worker make decisions.

Social Security:

For simplicity, social security benefit is assumed to be proportional to the pre-tax income at $T-1$; that is,

$$B_{iT} = \gamma Y_{iT-1}.$$

The pre-tax income just before retirement, Y_{iT-1}, is considered as a crude measure of life time income or a proxy for the Average Index of Monthly Earnings (AIME).

Heterogeneity:

There are two types of workers, those with a high discount factor $\beta_i = \overline{\beta}$ (low discounters) and those with a low discount factor $\beta_i = \underline{\beta}$ (high discounters, $\overline{\beta} \geq \underline{\beta}$)[12]. The discount factors $\overline{\beta}$ and $\underline{\beta}$ are parameters to be estimated, and the probability of being a high discounter is fixed at 0.5[13].

Bellman Equation:

A worker's problem can be summarized by the Bellman equation. Let PP_{it} be an indicator of employer type at the last period. For a worker of a pension-providing firm during the previous period, $PP_{it} = 1$. The Bellman equation of worker i at $t < T$ is given by

$$
\begin{aligned}
V_1(S_{i1}, P_{i1}, \tilde{H}_{i1}) &= \max_{C_{i1}, h_{i1}, p_{i1}, I_{i1}} U(C_{i1}, h_{i1}) + \beta_i E_1 V_2 \left(S_{i2}, P_{i2}, \tilde{H}_{i2}, p_{i1}\right) \\
V_t(S_{it}, P_{it}, \tilde{H}_{it}, PP_{it}) &= \max_{\{C_{it}, h_{it}, p_{it}, q_{it}, I_{it}\}_{t=2}^{T}} U(C_{it}, h_{it}) + \beta_i E_t V_{it+1} \left(S_{it+1}, P_{it+1}, \tilde{H}_{it+1}, p_{it}\right) \\
&\text{s.t.} \\
C_{it} + S_{it+1} &\leq (1 + (1 - \tau_{it})r)\, S_{it} + (1 - \tau_{it})\,(1 - I_{it})\, Y_{it}, \text{ if } S_{it+1} > 0 \\
C_{it} + S_{it+1} &\leq (1 + \overline{r})\, S_{it} + (1 - \tau_{it})\,(1 - I_{it})\, Y_{it}, \text{ if } S_{it+1} \leq 0 \\
C_{iT} &\leq (1 + (1 - \tau_{iT})\, r)\, S_{iT} + (1 - \tau_{iT})\,(B_{iT} + P_{iT}) \\
P_{it} &= \sum_{s=1}^{t-1} (1 + r)^{t-s} \left[p_{it}(1 + \psi) I_{is} Y_{is} + (1 - p_{it})\, I_{is} Y_{is}\right] \\
H_{it} &= \begin{cases} \tilde{H}_{it} & \text{if } q_{it} = 0 \\ D_t \tilde{H}_{it} & \text{if } q_{it} = 1 \end{cases} \\
\ln \tilde{H}_{it+1} &= \ln\left(H_{it} + \zeta\,(h_{it} - \underline{h})\right) + \epsilon_{it+1}, \quad \epsilon_{it+1} \sim N(0, \sigma_\epsilon) \\
B &= \gamma Y_{it} \\
C_{it} &\geq 0, \quad h_{it} \geq 0, \quad I_{it} \geq 0 \quad \forall t
\end{aligned}
$$

[12] A higher β means less discounting of future value. So, workers with a lower β are high discounters.

[13] The probability of being a high discounter is fixed because its identification of it is not obvious.

The space of S, P, $\tilde{H}$ are discretized and the problem is solve by backward induction. The control variables h and I are discretized whereas C is continuous.

3.2 Firms' Problem

There are two types of firms, the "ordinary" firms (or non-pension firms) and the firms having a fixed cost of hiring a worker. Ordinary firms have a constant returns-to-scale technology in labor input and pay just as much as each worker produces minus tax payment in the equilibrium. The profit of an ordinary firms from worker i is given by

$$\overline{\pi}_i = \sum_{t=1}^{T-1} \left(\frac{1}{1+r}\right)^{t-1} (1-p_{it}) \left[\overline{\alpha} H_{it} h_{it} - \left(\frac{1+\tau_{fica}}{1-\tau_{fica}}\right) \overline{w} H_{it} h_{it}\right]$$

where τ_{fica} is the Federal Insurance Contributions Act (FICA) tax rate and $\overline{w}$ is the human capital rental rate offered to workers of "ordinary" firms[14]. It is assume that the market is so competitive that $\overline{w} = \overline{\alpha}\left(\frac{1-\tau_{fica}}{1+\tau_{fica}}\right)$ and ordinary firms make zero profit in equilibrium.

The other type of firms also have a constant returns-to-scale production technology in labor input [15], but the marginal revenue $\tilde{\alpha}$ is higher than $\overline{\alpha}$, and they incur a fixed cost of hiring a worker ξ[16]. The hiring cost ξ can be considered the cost of training the workers; the marginal productivity of workers increases on account of training. The effect of training is assumed to be firm-specific, and the marginal productivity of a worker goes down to $\overline{\alpha}$ if he/she switches to a non-pension firm. The profit of the pension-providing firm from worker i is

$$\tilde{\pi}_i = \sum_{t=1}^{T-1} \left(\frac{1}{1+r}\right)^{t-1} p_{it} \left[\tilde{\alpha} H_{it} h_{it} - \left(\frac{1+\tau_{fica}}{1-\tau_{fica}}\right) \tilde{w}\left(H_{it} h_{it}\right) H_{it} h_{it} - \tilde{\psi}\left(I_{it} \tilde{w}\left(H_{it} h_{it}\right) H_{it} h_{it}\right) - \xi \mathbf{1}\left[q_{it} = 1\right]\right] \quad (2)$$

where p_{it} is an indicator of worker i of a pension-providing firm at period t, ξ is the fixed cost of

[14] Note that $\overline{w} H_{it} h_{it}$ is pre-tax income (but FICA tax is already taxed) from which workers make contributions to pension plans and the income tax is imposed; $\frac{1}{1-\tau_{fica}} \overline{w} H_{it} h_{it}$ is pre-tax income (FICA tax not imposed yet), and $\frac{\tau_{fica}}{1-\tau_{fica}} \overline{w} H_{it} h_{it}$ is the FICA tax that employers directly pay to the government.

[15] To be more precise, revenue is constant returns to scale. Owing to the fixed cost of hiring a worker, profit is not constant returns to scale.

[16] Since non-pension firms do not incur ξ, pension firms should have higher marginal revenue so that they survive in equilibrium.

hiring a worker, and $\tilde{\psi}(.)$ is the employer's contributions to the worker's pension plan:

$$\tilde{\psi}\left(I_{it}Y_{it}\right) = \begin{cases} 0.898 \times I_{is}Y_{is} & \text{if } \ I_{it} < 0.05 \\ 0.898 \times 0.05Y_{is} + 0.545 \times \left(I_{it} - 0.05\right)Y_{is} & \text{if } \ I_{it} \geq 0.05 \text{ and } I_{it} < 0.1 \\ 0.898 \times 0.05Y_{is} + 0.545 \times 0.05Y_{is} & \text{otherwise} \end{cases}$$

The first term in the solid bracket in Equation (2) is what worker i produces at period t, the second term in the same bracket is the wage paid to this worker, and the third term is the employer's contribution to the worker's pension plan. The last term is the fixed cost of hiring a worker. The profit maximization problem of pension-providing firm j can be written as

$$\begin{aligned} \tilde{\Pi}_j \quad &= \quad \max_{\tilde{w}(Hh)} \sum_{i \in N_j} \sum_{t=1}^{T-1} \left(\frac{1}{1+r}\right)^{t-1} p_{it} \left[\tilde{\alpha} H_{it}h_{it} - \left(\frac{1+\tau_{fica}}{1-\tau_{fica}}\right) \tilde{w}\left(H_{it}h_{it}\right) H_{it}h_{it} - \tilde{\psi}\left(I_{it}\tilde{w}\left(H_{it}h_{it}\right) H_{it}h_{it}\right) - \xi \mathbf{1}\left[q_{it}=1\right]\right] \\ \text{s.t.} \quad &\tilde{V}_t(S_{it}, P_{it}, H_{it}|p_{it}=1) > \tilde{V}_t(S_{it}, P_{it}, H_{it}|p_{it}=0) \end{aligned}$$

where N_j is the set of workers working for firm j and $\tilde{V}_t(S_{it}, P_{it}, H_{it}|p_{it})$ is the expected life-time utility of worker i at period t when he/she chooses a pension/non-pension firm at period t. The human capital rental rate offered by the pension-providing firm $\tilde{w}\left(H_{it}h_{it}\right)$ is determined in equilibrium.

The firms' problem is solved under the assumption that firms with a fixed cost of hiring always offer pension plans. However, in order to check whether the participation constraint of the pension-providing firms (for providing pension) is satisfied or not, a simulation with the estimated model is performed. For this, it is assumed that a firm with a fixed cost of hiring choose the profit maximizing human capital rental rate function without contribution to workers' pension plans when all the other firms with a fixed cost of hiring offer pensions and the equilibrium human capital rental rate at the estimates. It is also assumed that workers can choose job offers with and without pension plans from employers having a fixed cost of hiring as well as job offers without pension plans from employers not having a fixed cost of hiring.

3.3 Equilibrium

It is assumed that the economy is in an overlapping generations environment with no population growth and aggregate shock. The equilibrium of this model is a sequence of quantities and human

capital rental rates

$$\{\{C_{it}^*, h_{it}^*, p_{it}^*, q_{it}^*, I_{it}^*\}, \overline{w}^*, \tilde{w}^*(.)\}$$

such that, given the human capital rental rates $\overline{w}$ and $\tilde{w}^*(.)$, consumption C_{it}^*, hours of work h_{it}^*, investment in pension assets I_{it}^*, whether to work for a pension-proving firm p_{it}^*, and job switch q_{it}^* maximize the life-time utility of workers. Allowing for workers' responses, firms choose the human capital rental rates they offer. The labor market is assumed to be so competitive that the firms cannot but offer the most generous human capital rental rate with positive (but almost zero) profit in equilibrium. As a result, $\overline{w} = \overline{\alpha}\left(\frac{1-\tau_{fica}}{1+\tau_{fica}}\right)$, $\overline{\pi} = 0$, and $\tilde{\Pi} = 0$ hold in equilibrium.

Since no closed form solution exists for pension providing firms' problem, it is solved numerically in the following manner: Let $l_{it} \equiv H_{it}h_{it}$ be the labor supply of worker i at time t. In the space of l, G_l grid points $\{l_1, l_2, \ldots, l_{G_l}\}$are picked up; in the space of $\tilde{w}$, G_w grid points $\{\tilde{w}_1, \tilde{w}_2, \ldots, \tilde{w}_{G_w}\}$ are picked up; and for each grid point of l_j, $j \in \{1, 2, \ldots, G_l\}$, one grid point of human capital rental rate $\tilde{w}^{(j)} \in \{\tilde{w}_1, \tilde{w}_2, \ldots, \tilde{w}_{G_w}\}$ is picked up. By connecting these point $\left\{\left(l_j, \tilde{w}^{(j)}\right)\right\}_{j=1}^{G_l}$ with a monotone cubic spline (Fritsch and Carlson, 1980), a human capital rental rate function $\tilde{w}(l)$ is constructed. Since firms have an incentive to offer the workers with larger labor supply a better deal, $\tilde{w}(l)$ will be an increasing function in l. When the number of grid points of l is four and that of $\tilde{w}$ is ten, there are 205 candidate human capital rental rate functions. In a competitive market, firms offer the most generous human capital rental rate with positive expected profit. Therefore, among the 205 candidate human capital rental rate functions, the one with the highest rate of pension participation with positive profit is regarded as the equilibrium human capital rental rate. Figure 3 illustrates the construction of human capital rental rate function $\tilde{w}(l)$. X axis is the labor supply $l_{it} \equiv H_{it}h_{it}$ and Y axis is the human capital rental rate.

3.4 Assumptions on the Human Capital Rental Rate

In the model, firms with a fixed cost of hiring have an incentive to hire workers who work longer. Thus, if the probability of leaving the current employer is constant across workers of different ages, firms have an incentive to offer higher wages to younger workers. However, the probability of leaving the current employer is decreasing in the model. Therefore, it depends on the parameter value if the firm prefer younger workers or older workers. For simplicity, it is assumed that firms cannot make

an age-dependent wage offer. The Age Discrimination in Employment Act of 1967 prohibits firms from discriminating against older workers. In the model, firms and workers observe the amount of labor input provided by each worker, and it is assumed that firms would be accused of violation of the Age Discrimination in Employment Act of 1967 if they offer different wages to workers with the same amount of labor supply but different ages[17]. In addition, we assume that firms cannot make a tenure-dependent wage offer. If firms were allowed to make a tenure-dependent wage offer, they would screen out high discounters by offering a deferred wage contract. They would thus screen out high discounters by offering a lifetime wage profile such that the starting wage rates are lower than the workers' productivity level, but the wage rates grow at a higher rate than the growth rate of workers' productivity. Those who place more emphasis on future income than present consumption would choose such an offer, and high discounters can be screened out without the use of pensions. Therefore, we assume that firms are not allowed to do so[18]. This assumption might appear to be very strong, but the possibility of screening by such a deferred wage contract should be limited in the U.S. This is because, in order for such an offer to be accepted, firms have to commit to workers' future wage and they have to be trusted their firms. Younger workers may not accept a low starting wage rate if they do not believe that the firms pay as much as they promise in the future, or if they suspect that firms may fire them before the wage rates catch up with their productivity. In the United States, the job turnover rate is relatively high and firing a worker is easier than in other developed countries, for example, Japan. Therefore, younger workers find it riskier to accept a lower wage rate in the United States. Another justification of this assumption is that making commitment to the future wage rate may be too risky for firms when their economic prospects could become clouded. Therefore, the assumption of no tenure-dependent wages can be interpreted as an assumption of lack of trust by (younger) workers or firms' uncertainty about their economic prospects.

[17] In the real world, however, the rule of non-discrimination against older workers is applied only to workers above the age of 40.

[18] We can think about a firm offer tenure-dependent matching contributions to pensions. However, this is likely to violate the "non-discrimination rule" of defined contribution pensions: The Internal Revenue Service do not allow firms to discriminate against workers with lower wages, and the ratio of contributions to pension plans to wages have to be constant across workers with different wage levels.

4 Data

The data set used in this paper is taken from the Survey of Income and Program Participation (SIPP) data. The SIPP data provides detailed information about income, assets, pension participation, and contributions to pension plans by surveying a nationally representative sample of individuals aged 15 and above, interviewed every four month over a period of 32 to 52 months[19]. Although each panel consists of 8 to 13 waves, the detailed information about pensions is provided only in the topical modules appearing at most on three waves in each panel. In addition, some important variables relating to pension plans before the 1996 panel are not available. Therefore, we construct a repeated cross-section data by gathering one wave from each of the 1996, 2002, 2004, and 2008 panels.[20]. Since the earnings on pension assets are tax-deferred and all pension-providing firms contribute to pension plans in the model, in this paper pension workers are defined as those who have tax-deferred pension accounts and whose employers also contribute to match the employees' contributions. Therefore, workers with different types of pension plans such as DB plans, non-401(K) defined contribution plans, or hybrid plans are excluded. Although workers with a pension plan in the model can choose not to contribute themselves, only those who contribute are regarded as working for pension-providing firms. This is because workers in the sample not contributing do not provide enough information about their pension plans. Since the model does not have self-employed workers, the self-employed workers in the sample are eliminated. Public workers are also excluded because the objectives of governments are supposed to be different from those of private firms. This study focuses on male workers, because female workers face various problems such as childbirth that the model does not deal with. Although SIPP data cover all individuals above the age of 15, the number of workers above 50 with 401(K) plans is quite small. This is because 401(K) plans became common only during the mid-1980s. Therefore, the data of individuals aged between 24 and 48 are collected, and then the moments of annual labor income, the rate of job switch, and the share of pension workers calculated from the actual and generated data are matched for the first to fourth (ages 24 to 29, 30 to 35, 36 to 41, and 42 to 47). Similarly, the moments of pension assets and non-pension assets at the beginning of the first to fifth period (age around 24, 30, 36, 42, and 47) calculated from the actual and generated data are matched.

[19]Thus, each panel of SIPP data has 8 to 13 four-month-long waves.

[20]All the variables are regressed on cohort and year dummies and the effect of cohort and year are eliminated.

There are in all 2791 workers with pension plans and 3780 workers without pension [21]. Since the workers of different cohorts are mixed, we adjust the price levels by using the consumer price index (CPI) and eliminate the cohort and year effects by regressing all the related variables on year and cohort dummies so that the sample can be considered homogeneous.

4.1 Summary Statistics

Figure 1 shows the annual income from wages and salary of male full-time workers aged between 23 and 52 in the SIPP data. The figure shows that workers with a pension plan earn more than workers with no pension. Figure 2 indicates that the difference between the annual income of workers with pension and without pension is due to the difference in their hourly wages (annual income divided by annual hours of work) and hours of work.

4.2 Simple Regression Analysis

Table 8 shows the results of simple ordinary least squares (OLS) regressions. The dependent variables are annual income, hourly wage, and tenure. Each of these variables is regressed on age; age squared; the dummy variables of education level, occupation, and industry; the number of employees working for an employer at the location where the individuals are working; the total number of employees working for the employer; and the year dummies [22] In addition, in order to control for the difference in employers' preference on how the pecuniary/non-pecuniary compensation is provided, the dummy variables of fully financed and partly financed health insurance are included. The regression results indicate that workers with employer-provided pension earn $5,200 more per year and $2.5 more per hour, and stay 22 months longer than workers without pension plans do.

4.3 Employers Matching Provision

Figure 4 shows the fraction of workers whose employers link their contributions to their employees' contributions for workers with tax-deferred pension accounts. Around 80% of workers with tax-deferred pension accounts have employers whose contributions depend entirely or partly on their

[21] Since the workers with different types of pension plans, such as DB pensions, are eliminated, the number of pension workers we estimate is less than the actual number of pension workers.

[22] Wage rates are also included as an independent variable for the tenure regression model.

own contributions. Figure 5 shows the distribution of the ratio of employers' contributions to employees' contributions. We find that a majority of workers contribute to the level up to which a 100% match is provided.

5 Estimation

5.1 Estimation Method

The model is estimated by using the method of simulated moments (MSM). The estimates of the structural parameters minimized the "distance"between the moments calculated from the SIPP data and those calculated from the generated (simulated) data. The moments to be matched are annual labor income, non-pension asset holdings, pension asset holdings, rates of changing employers (job switch) of the workers with and without pension respectively over four to five periods. In addition, the share of pension workers and the standard deviations of annual labor income are matched. Since 401(K) plans have become so common since the mid-1980s, many of the older workers with 401(K) plans are those who started their career with DB pension. Therefore, we can match only the moments of the first four periods of the annual income, the rate of job switch, and the rate of participation and the moments of the first five periods of the non-pension asset and pension asset holdings.

Let θ be the vector of the structural parameters to be estimated, and let $m(\theta, \tilde{w}(l))$ denote the vector of moments calculated from the data generated by the model and $\overline{m}$ denote the vector of the corresponding moments obtained from the SIPP data. The vector of parameters to be estimated, θ, include parameters of workers' preferences, parameters in the transition equation of human capitals, the marginal product of labor of firms with a fixed cost of hiring $\tilde{\alpha}$, and the fixed cost of hiring ξ.

5.2 Estimation Procedure

The estimation procedure is as follows: for a candidate parameter vector $\tilde{\theta} \in \Theta$, we solve the workers' problem for each candidate human capital rental rate function $\tilde{w}(l)$ by backward induction and obtain the workers' policy function. We then use this policy function to run simulations and

obtain the vector of simulated moments $m\left(\tilde{\theta}, \tilde{w}(l)\right)$ and the expected value of profit of pension-providing firms for each candidate $\tilde{w}(l)$. Then the equilibrium human capital rental rate function $\tilde{w}^*\left(l, \tilde{\theta}\right)$ is the one which maximizes the number of workers who work for pension-providing firms with a positive expected profit [23]. Then we match the simulated moments $m\left(\tilde{\theta}, \tilde{w}^*\left(l, \tilde{\theta}\right)\right)$ with the data moments $\overline{m}$. The objective function is given by

$$\left[\overline{m} - m\left(\tilde{\theta}, \tilde{w}^*\left(l, \tilde{\theta}\right)\right)\right]' W \left[\overline{m} - m\left(\tilde{\theta}, \tilde{w}^*\left(l, \tilde{\theta}\right)\right)\right]$$

where W is a weighting matrix whose diagonal elements are the inverse of the variance of each elements of the data moments.

5.3 Identification

In general, identification of the discount factors and the relative risk aversion coefficient is ambiguous when they are estimated by matching the life-cycle path of asset holdings or consumption. However, the model estimated in this paper has two assets, and the portfolio choice depends on the time preferences and the relative risk aversion coefficients: workers hold non-pension assets for precautionary motives which heavily depends on the relative risk aversion coefficient, and pension asset accumulation depend on workers preference for future consumption, which is closely related to discount factors. Therefore these parameters can be separately identified by matching the life-cycle path of pension and non-pension assets holdings respectively. However, it is difficult to identify the coefficient on the utility at period T μ separately from the discount factor β, because the moments are matched only for relatively younger workers. So μ is fixed at 1.

The identification of parameters in the human capital accumulation function is also ambiguous. It is assumed that human capital stocks and hours of work are not separately observed and the observed hours of work do not correspond to the hours of work in the model, so the moments of wage rates and hours of work should not be used. Thus, only the life-cycle path of annual income is matched, but the parameters in the human capital accumulation equation cannot be identified

[23] The equilibrium human capital rental rate function is the most generous one with positive (but almost zero) profit. Since the space of the human capital rental rate function is discretized, the most popular one with positive profit is regarded as the most generous one. The zero profit condition is imposed by adjusting the fixed cost of hiring ξ after finding the equilibrium human capital rental rate function.

solely by matching annual income. However, the life-cycle path of hours of work in the model is determined by other parameters such as discount factor β or relative risk aversion coefficient γ. Therefore the parameters in the human capital accumulation equation can be identified when the preference paramtparameters such as β and γ are identified and the life-cycle path of hours of work is pinned down by matching the moments other than the life-cycle path of annual income. The hiring cost ξ and the marginal product of labor $\tilde{\alpha}$ jointly determine workers' choice of employer (pension-providing employer or not), and the marginal productivity of labor is also related to wages. Thus, they are identified from the life-cycle path of pension participation and annual labor income. The rate of changing employers is closely related to the disutility from working $\underline{\nu}$ and $\overline{\nu}$ and the depreciation rate of human capital (δ_0, δ_1, and δ_2). In addition, the disutility from working is also related to hours of work, and the depreciation rate is related to wages. Thus $\underline{\nu}$, $\overline{\nu}$, δ_0, and δ_1 can be identified from the life-cycle path of annual labor income and the rate of changing employers.

The discount factors $\overline{\beta}$ and $\underline{\beta}$ can be identified from the life-cycle profile of hours of work, non-pension assets, and the pension assets of pension workers and non-pension workers when the low discounters and high discounters are perfectly sorted into pension jobs and non-pension jobs. However, workers are sorted not only in accordance with their time preferences, but also human capital levels. Therefore, a model with homogeneous β can generates a life-cycle path of hours of work and asset accumulation which is close to the actual life-cycle path of hours of work and asset accumulation, because higher human capital can result in higher hours of work (due to the substitution effect) and larger asset holdings. This can cause a serious identification problem with respect to the heterogeneity of discount factor β. However, the gap between the rates of job switch can only be explained by the heterogeneity in the time preferences. Therefore discount factors $\overline{\beta}$ and $\underline{\beta}$ can be separately identified by matching the rate of job switch and the other moments of workers with and without pension.

Note that the fixed cost of hiring a worker ξ cannot be pinned down at one value without the zero-profit condition if the space of human capital rental rate function $\tilde{w}(l)$ is discretized: $\tilde{w}^*(l)$ is the equilibrium human capital rental rate function as long as the rate of participation in pension plans is higher than that at the other candidate human capital rental rate functions. Therefore, $\tilde{w}^*(l)$ can be an equilibrium human capital rental rate function even when the profit is positive when ξ stays within a boundary $\left[\underline{\xi}^*, \overline{\xi}^*\right]$ where $\tilde{w}^*(l)$ is an equilibrium human capital rental rate:

if $\xi < \underline{\xi}^*$another human capital rental rate function with a larger rate of participation in pension plans will be an equilibrium human capital rental rate function. If $\xi > \overline{\xi}^*$, firms with a fixed cost of hiring make negative profits and $\tilde{w}^*(l)$ will not be the equilibrium human capital rental rate anymore. The estimates and the simulated moments of the model will not change as long as ξ stays within this boundary, but the counter factual simulation results can change. Therefore, the cost of hiring a worker ξ is pinned down at $\xi = \overline{\xi}^*$ so that the pension-providing firm's profit at equilibrium is zero.

6 Estimation Results

6.1 Model Fit

For the estimated parameters, see Table 3 and Table 4. The estimates of discount factor $\overline{\beta}$ and $\underline{\beta}$ are 0.958 and 0.926, respectively, in annual terms. Figure 6 shows the moments of hours of work, savings, pension assets, and the rate of a job switch calculated from SIPP data and the data generated from the estimated model. The estimated model is successful in replicating the life-cycle path of asset accumulation, but overestimates the life-cycle path of annual income and rates of job switch of workers with pension.

6.2 Low Discounters and High Discounters

The simulated moments of low discounters ($\beta = \overline{\beta}$) and high discounters ($\beta = \underline{\beta}$) are given in Table 7. Although the difference in their estimated discount factors is only 3% in annual terms, the low discounters and high discounters reveal surprisingly different behavior. Workers with a higher β, that is, low discounters, work harder, accumulate more human capital, stay longer with their current employer, and are more likely to take up jobs with pension plans, which is consistent with our intuition.

Figure 9 illustrates the low and high discounters' rate of taking up jobs with pension plans by human capital level. Note that the level of human capital stock of most of the workers ranges between 15 and 30. This shows that low discounters and highly productive workers (who have larger human capital stock) are more likely to take up jobs with pension plans.

6.3 Human Capital Rental Rate Function

Figure 8 illustrates the human capital rental rate function $\tilde{w}^*(l)$. The figure shows that the human capital rental rates offered by pension firms is lower than the marginal productivity of workers when the labor supply $H_{it}h_{it}$ for over six years is less than 350,000. This means that there is a trade-off between lower wage rates and pension benefits for such workers that work for 2000 hours per year for an employer without pension who pays less than \$30 per hour (before income tax is imposed). In the model, more than 90% of workers face such a trade-off.

6.4 Robustness Check

In our model, the rate of matching contribution is given and fixed. In order to see how sensitive the model is to this assumption, we run simulations with the following two plans:

Plan 1: Firms offer 100% match for up to 5% of wage

$$\psi_1\left(s_{is}Y_{is}\right) = \begin{cases} 2 \times I_{is}Y_{is} & \textbf{if} \quad s_{it} < 0.05 \\ 2 \times 0.05 \times Y_{is} + \left(I_{it} - 0.05\right) Y_{is} & \textbf{otherwise} \end{cases}$$

Plan 2: Firms offer 50% match for up to 10% of wage

$$\psi_2\left(s_{is}Y_{is}\right) = \begin{cases} 1.5 \times I_{is}Y_{is} & \textbf{if} \quad s_{it} < 0.1 \\ 1.5 \times 0.1 \times Y_{is} + \left(I_{it} - 0.1\right) Y_{is} & \textbf{otherwise} \end{cases}$$

Firms are allowed to change their human capital rental rates according to the rate of matching contributions. The simulation results are given in Figure 10. The figure shows the model very sensitive to the parametric assumption of $\psi(.)$. Thus, the model is not robust to the parametric assumption with respect to $\psi(.)$.

7 Counter-factual Simulations

7.1 Roth 401(K)

In order to study the importance of tax deferrals on pension asset earnings, we perform a counter-factual simulation in which the government imposes income tax on worker and firm contributions to pension plans and the earnings accrued in pension and individual retirement accounts, but do not impose income tax on the pension benefits received after the workers retire. This resembles the pension plans called "Roth 401(K)."

Figure 11 summarizes the simulation results. The pension assets balance of a worker at age 61 declines by 50% from the deprivation of tax benefit, rendering pension not as attractive as before. In addition, it is more difficult for firms to finance contributions to worker's pension plans, firms become picky: firms offer lower human capital rental rates for workers with smaller labor supply than the case with ordinary 401(K) (See Figure 12). As a result, share of workers with pension declines by 14%.

7.2 Flat Income Tax

Employers and employees can be better of by flattening the life-time income of workers by using pension when the income tax is progressive. Private pension can be regarded as an inter-temporal income transfer from younger self to older self. Since income declines when a worker retires, income transfer from working younger self to retired older self enables workers to avoid a high marginal tax rate while they are working. This can be the role of pension: firms offer pension for the tax saving purpose.

In order to check the importance of tax saving by flattening the life-time income (by using pension), we carried out a counter-factual simulation where income tax is flat. The tax rate is set at the level where the tax revenue is the same as with the progressive income tax. The simulation result is summarized in Figure 13. The share of workers with pension declines by more than 80%, which means tax saving by flattening lifetime income is crucial. Since it is hard for firms and workers to be better of by tax saving, the human capital rental rate offered by pension-providing firms is very low (See Figure 14). As a result, only few workers are attracted by the jobs with

pension, and the share of workers with pension declined dramatically.

8 Conclusion

This paper proposes and estimates a structural model of screening by employer-provided pensions. The model illustrates how workers are sorted into jobs with pension plans and without pension according to their time preferences and human capital level. This paper is the first attempt in the literature to estimate a structural model where workers decide their consumption/savings and participation in/contributions to pension plans. The study explicitly incorporates firms' incentives to offer contributions to pension plans. This study is unique in that although its basic structure is a life-cycle model, it illustrates how workers are sorted into different jobs according to a "deep" parameter such as the discount factor.

The proposed model is estimated by using SIPP data. Our estimation results show how workers with different discount rates behave: low discounters work harder, accumulate human capital, stay longer with their current employer, and are more likely to take up jobs with pension plans, consistent with our intuition. The model also shows how workers are sorted into jobs with and without pension: workers with a low discount rate and high human capital level are more likely to take up jobs with pension plans, and vice versa; however, workers cannot be sorted solely by discount rates: there are some low discounters who choose jobs without pension because of low human capital levels, and vice versa.

Finally, The result of a counter-factual simulation in which the government eliminates tax benefits from 401(K) plans suggests that the pension balance of a worker with pension at age 61 declines by 50%, and the number of workers working for the pension providing firms decrease by 14%. So the tax benefit of 401(K) is important especially for the retirement asset accumulation of workers with pension. In addition, the result of a simulation with flat income tax suggests that the number of workers with employer provided pensions declines by 82%. It indicates that it is important for firms and workers to save tax payment by flattening income over the life-cycle by taking advantage of pension plans.

References

[1] Dey, M. & C. Flinn (2005). An Equilibrium Model Of Health Insurance Provision and Wage Determination, *Econometrica*, 73(2), 571-627.

[2] Dorsey, S., & Macpherson, D. A. (1997). Pensions and training. *Industrial Relations: A Journal of Economy and Society*, 36(1), 81-96.

[3] Even, William E., & David A. Macpherson (2005). The effects of employer matching in 401 (k) plans. *Industrial Relations: A Journal of Economy and Society*, 44(3), 525-549.

[4] French, E. (2005). The effects of health, wealth, and wages on labour supply and retirement behaviour. *The Review of Economic Studies*, 72(2), 395-427.

[5] French, E., & Jones, J. B. (2011). The Effects of Health Insurance and Self- Insurance on Retirement Behavior. *Econometrica*, 79(3), 693-732.

[6] Fritsch, F. N., & Carlson, R. E. (1980). Monotone piecewise cubic interpolation. *SIAM Journal on Numerical Analysis*, 17(2), 238-246.

[7] Gourinchas, P. O., & Parker, J. A. (2002). Consumption over the life cycle. *Econometrica*, 70(1), 47-89.

[8] Gustman, A. L., Mitchell, O. S., & Steinmeier, T. L. (1994). The role of pensions in the labor market: A survey of the literature. *Industrial and Labor Relations Review*, 417-438.

[9] Heiland, F., & Li, Z. (2012). Changes in Labor Force Participation of Older Americans and Their Pension Structures: A Policy Perspective. *Boston College Center for Retirement Research Working Paper*, (2012-18).

[10] Imai, S., and Keane, M. P. (2004). Intertemporal labor supply and human capital accumulation*. *International Economic Review*, 45(2), 601-641.

[11] Ippolito, R. A. (1997). *Pension plans and employee performance: Evidence, analysis, and policy*. University of Chicago Press.

[12] Ippolito, R. A. (2002). Stayers as" Workers" and" Savers": Toward Reconciling the Pension-Quit Literature. *Journal of Human Resources*, 275-308.

[13] Kusko, A. L., Poterba, J. M., & Wilcox, D. W. (1994). *Employee decisions with respect to 401 (k) plans: Evidence from individual-level data* (No. w4635). National Bureau of Economic Research.

[14] Laibson, D., Repetto, A., & Tobacman, J. (2007). *Estimating discount functions with consumption choices over the lifecycle* (No. w13314). National Bureau of Economic Research.

[15] Lee, D., & Wolpin, K. I. (2006). Intersectoral labor mobility and the growth of the service sector. *Econometrica*, 74(1), 1-46.

[16] McFadden, D. (1989). A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica: Journal of the Econometric Society*, 995-1026.

[17] Mitchell, O. S., Utkus, S. P., & Yang, T. (2006). *Dimensions of 401 (k) plan design. Restructuring retirement risks*. Oxford University Press, UK, 186-203.

Figure 1: Pension and Annual Income

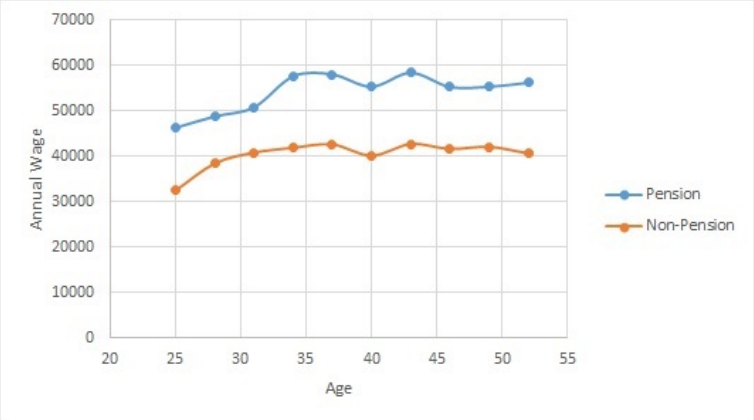

Figure 2: Pension and Hourly Wage (left), Pension and Hours Worked (right)

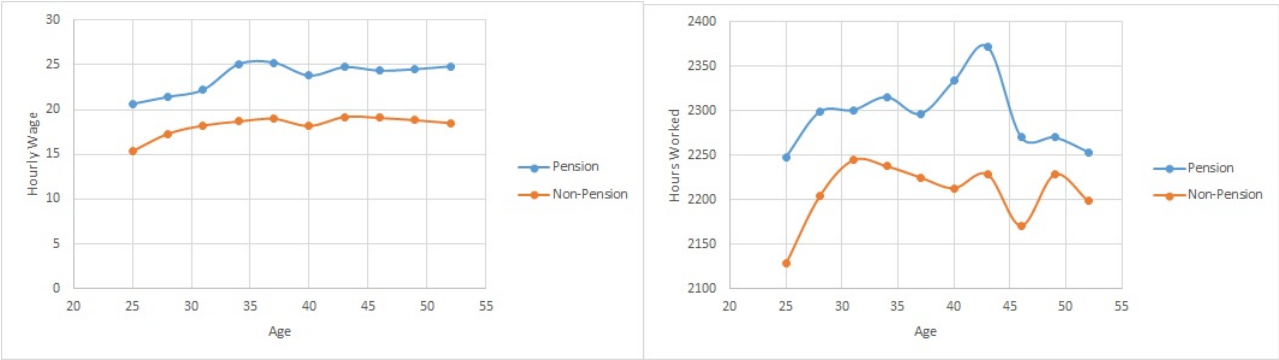

Figure 3: How to Find $\tilde{w}(.)$

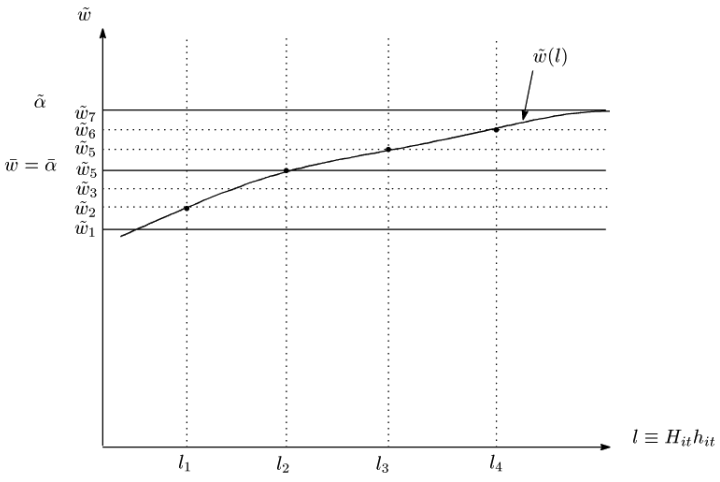

Table 1: Annual, Hourly Wage and Pension

	Dependent variables		
	Annual Income	Wage Rate	Tenure (month)
Some College	5189	2.49	-19.05
	(1033)	(0.41)	(4.63)
Bachelor's Degree	11760	5.17	-42.08
	(1364)	(0.55)	(6.24)
Master and Ph.D	21788	9.30	-46.89
	(2324)	(0.93)	(10.48)
Health Insurance (full)	6624	2.40	17.79
	(1436)	(0.58)	(6.59)
Health Insurance (part)	7955	2.90	11.90
	(1186)	(0.48)	(5.57)
Age	2042	0.82	7.62
	(328)	(0.13)	(1.48)
Age Squared	-22.77	-0.0089	-0.04
	(3.85)	(.0015)	(0.017)
Pension	**6814**	**2.73**	**22.12**
	(972)	(0.93)	(4.38)

Dummy variables of occupations and industries, the number of employees working for the employer at the location where individuals are working, the total number of employees working for the employer, and year dummies are also included.

Figure 4: The Link between the Contributions of Employers and Employees

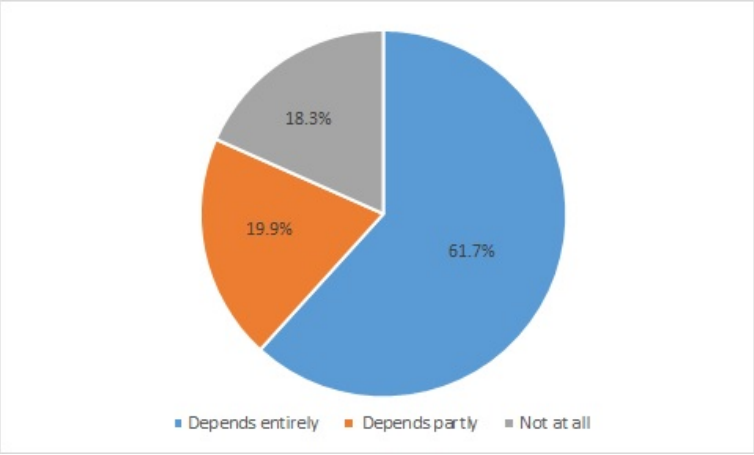

Table 2: Summary of the SIPP data

	Pension Workers	workers without pension plans
Annual Hours of Work (Age 26-40)	2,292.34	2,172.145
	(423.60)	(473.72)
Annual Hours of Work (Age 41 - 55)	2,299.44	2,179.93
	(395.20)	(452.47)
Hourly Wage (Age 24 - 28)	19.14	14.410
	(8.31)	(6.74)
Hourly Wage (Age 38 - 42)	25.09	18.132
	(9.70)	(8.85)
Hourly Wage (Age 53 - 57)	24.26	17.941
	(9.68)	(9.04)
Non-Pension Assets (Age 24 - 28)	15,749.29	10,232.39
	(26,987.02)	(21,675.30)
Non-Pension Assets (Age 38 - 42)	23,960.63	12,403.43
	(36.819.92)	(31,462.27)
Non-Pension Assets (Age 53 - 57)	30766.29	19254.87
	(41,564.94)	(37,645.84)
Pension Assets (Age 24 - 28)	15,376.94	5,351.956
	(24,449.15)	(11,658.03)
Pension Assets (Age 38 - 42)	43,673.17	10,106.92
	(56,156.45)	(27,556.13)
Pension Assets (Age 53 - 57)	79,747.87	26441.11
	(95,5949.49)	(63,540.81)
Participation Rate (Age 26 - 40)	0.371	
Participation Rate (Age 41 - 55)	0.455	

Numbers in brackets are standard errors. The numbers of non-pension assets and pension balance are in thousands.

Figure 5: The Ratio of Employers' Contribution to Employees' Contribution

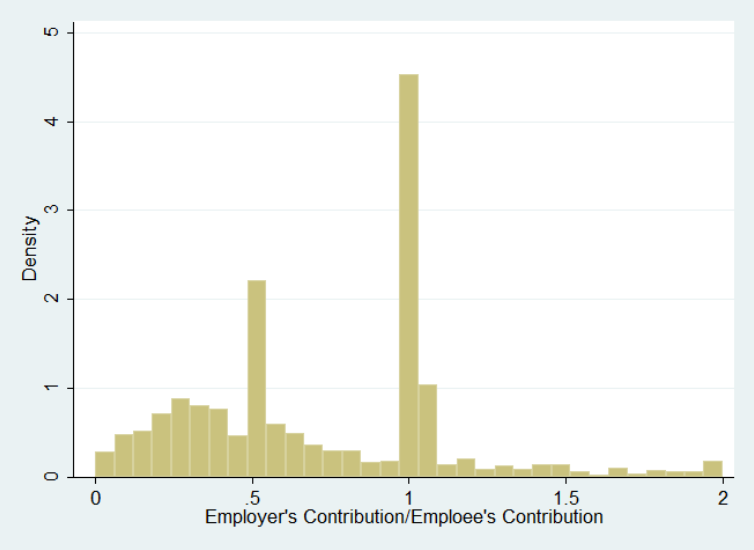

Table 3: The Ratio of Employers' Contribution to Employees' Contribution

Parameters		Estimates
Utility Function		
Discount Factor (high)	$\overline{\beta}$	0.772 (0.958 in annual terms)
		(—)
Discount Factor (low)	$\underline{\beta}$	0.631 (0.926 in annual terms)
		(—)
Marginal disutility from working (low)	$\underline{\nu}$	0.991
		(—)
Weight of disutility from working (high)	$\overline{\nu}$	1.119
		(—)
Probability of getting $\nu = \overline{\nu}$	ϕ	0.309
		(—)
Minimum Consumption at $t = 1$	λ_1	134.348
		(—)
Coefficient of Increase in $\underline{C}$	λ_2	24.119
		(—)
CRRA coefficient	γ	0.531
		(—)

Notes: The standard errors will be calculated right before submission to an academic journal, because it costs a lot of computational resources.

Table 4: Estimates of Parameters (Human Capital and etc.)

Parameters		Estimates
Human Capital Accumulation		
Marginal Effect of Learning	ζ	0.971
		(—)
Threshold of Accumulating/Decumulating	$\underline{h}$	11.620 (1973 hours in annual terms)
		(—)
Standard Deviation of Shock	σ_h	0.249
		(—)
Human Capital Depreciation due to Job Switch		
Initial Distribution of Human Capital		
The Depreciation at $t = 1$	δ_0	0.967
		(—)
The Rate of Depreciation Rate Depreciate	δ_1	0.0005
		(—)
The Rate of Depreciation Rate Depreciate	δ_1	0.003
		(—)
Mean of $\ln H_{i1}$	μ_H	3.048
		(—)
Std. Dev. of $\ln H_{i1}$	σ_{μ_H}	0.262
		(—)
Pension Firms' Variable		
Marginal Product of Labor Input ($H_{it}h_{it}$)	$\tilde{\alpha}$	1.047
		(—)
Hiring Cost for Pension Firms	ξ	26.644
		(—)

Notes: The standard errors will be calculated right before submission to an academic journal, because it costs a lot of computational resources.

Figure 6: The Data Moments and Simulated Moments at the Estimates

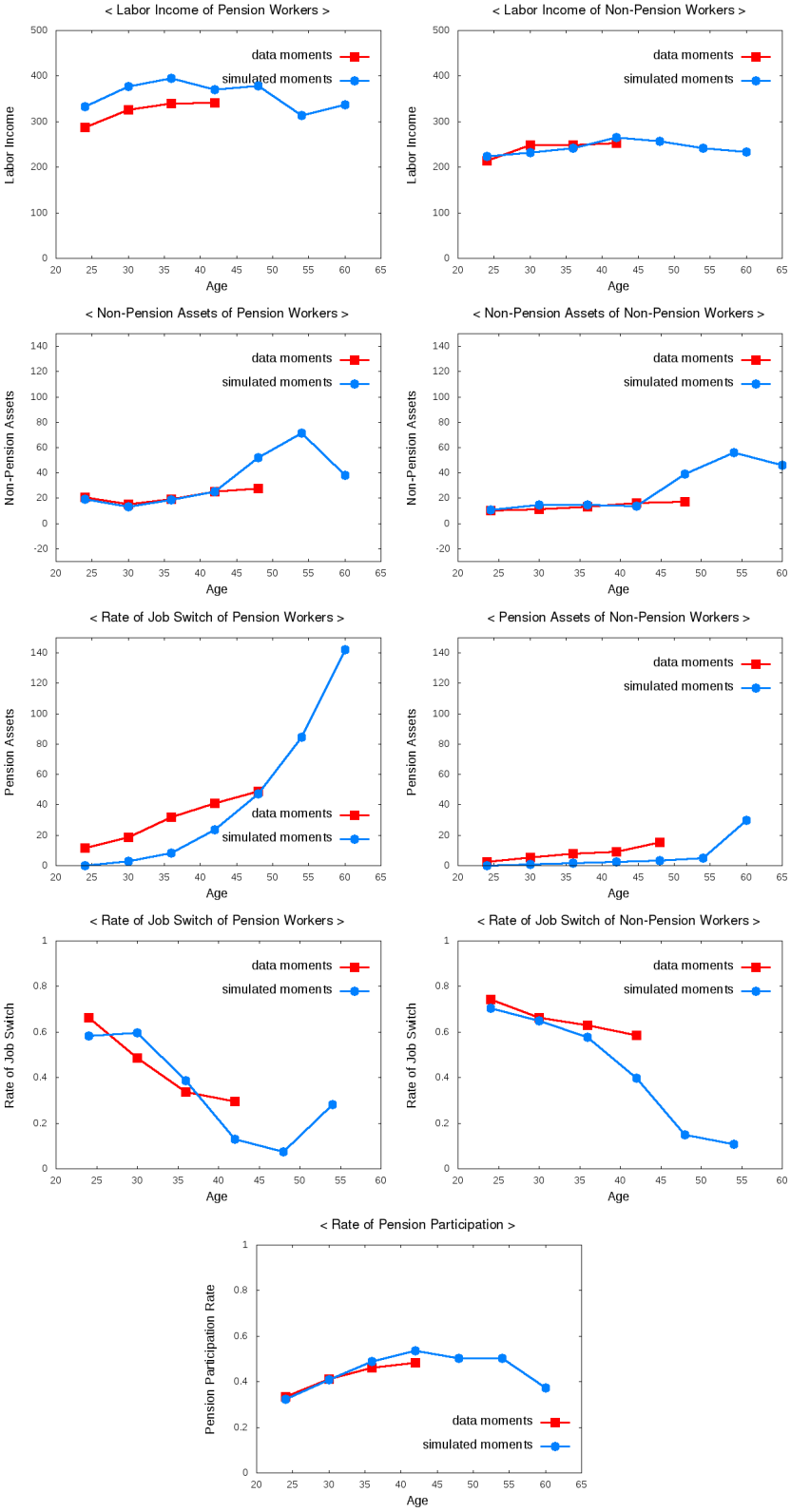

Figure 7: Simulated Moments of High/Low Discounters

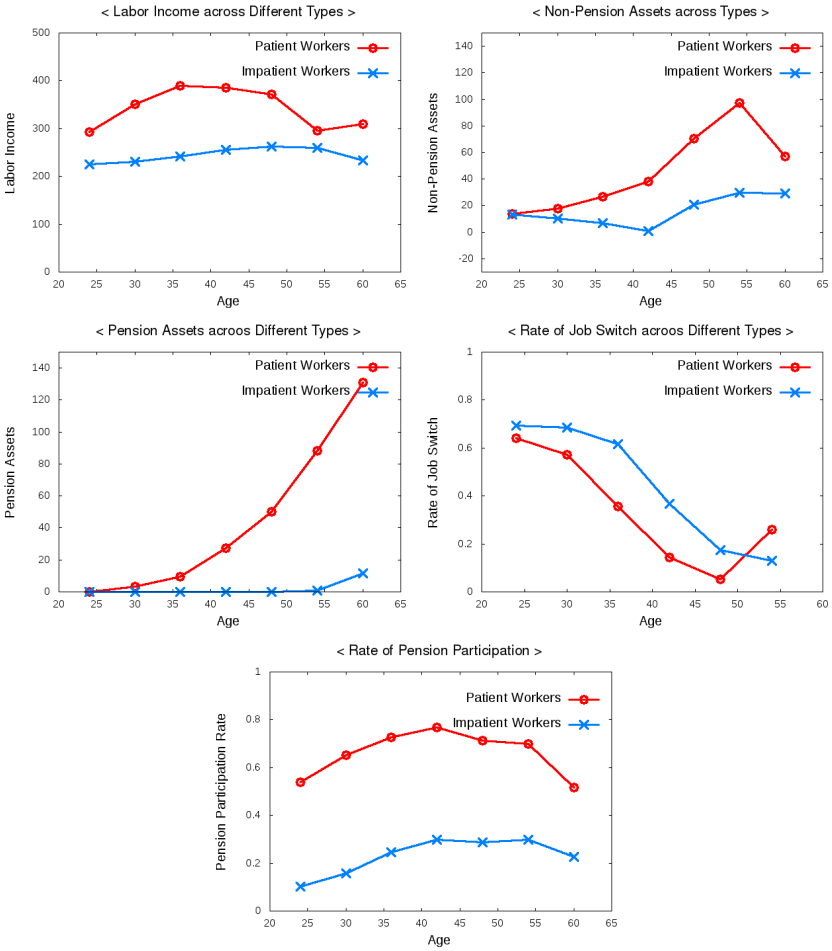

Figure 8: Labor Supply And Human Capital Rental Rate

Figure 9: Human Capital and Pension Participation

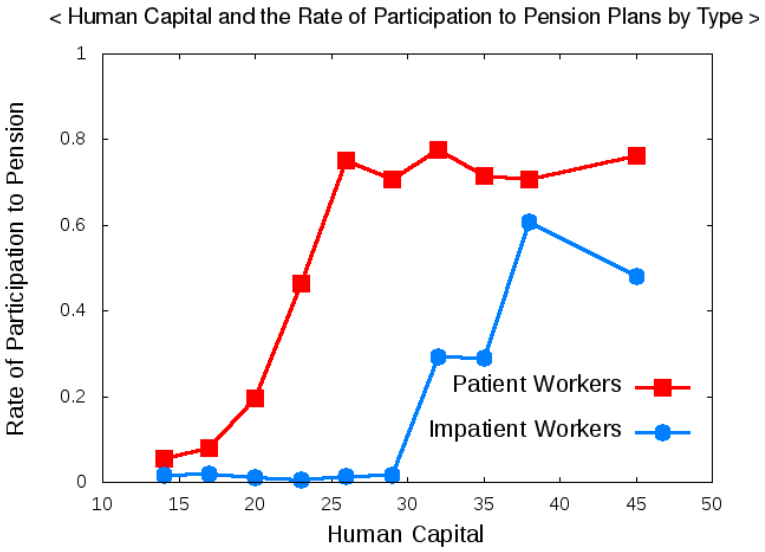

Figure 10: Robustness Check: Different Matching Contribution

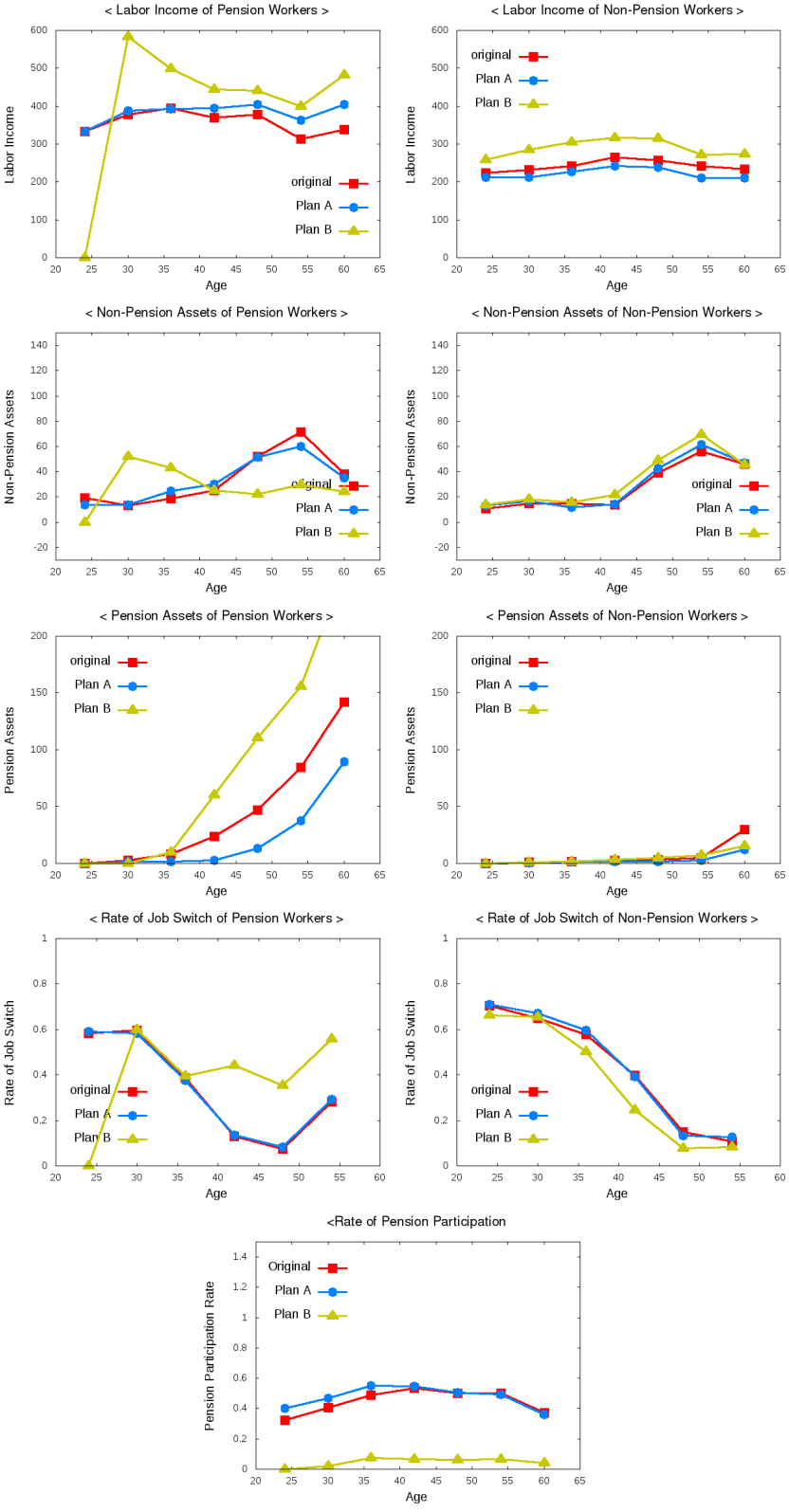

Figure 11: Counter-factual Simulation 1: Roth 401(K)

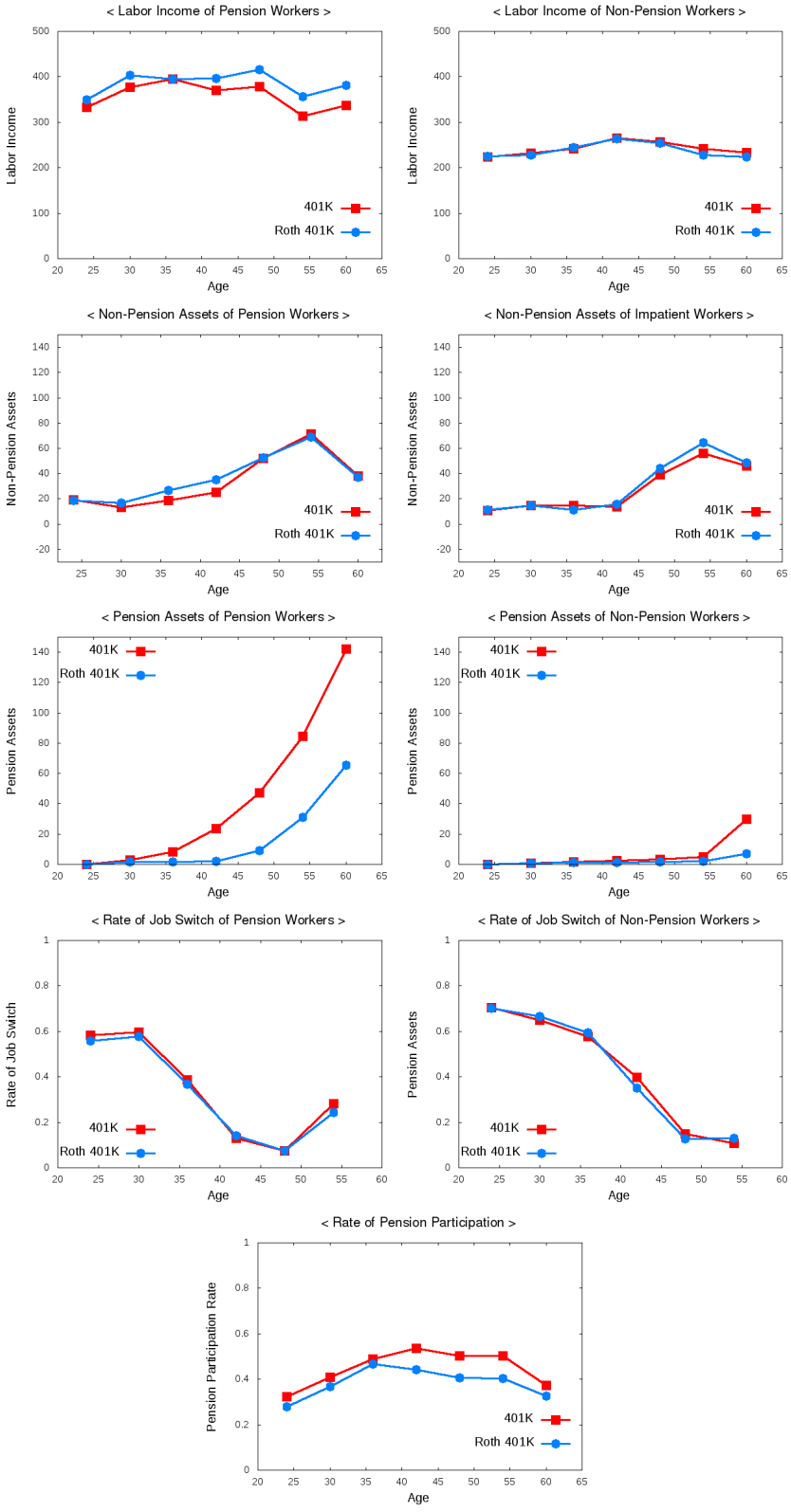

Figure 12: Counter-factual Simulation 1: Roth 401(K)

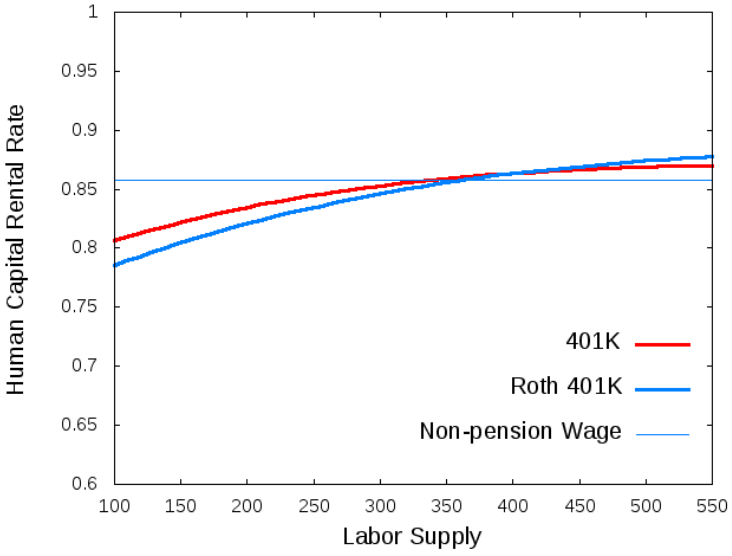

Figure 13: Counter-factual Simulation 2: Flat Income Tax

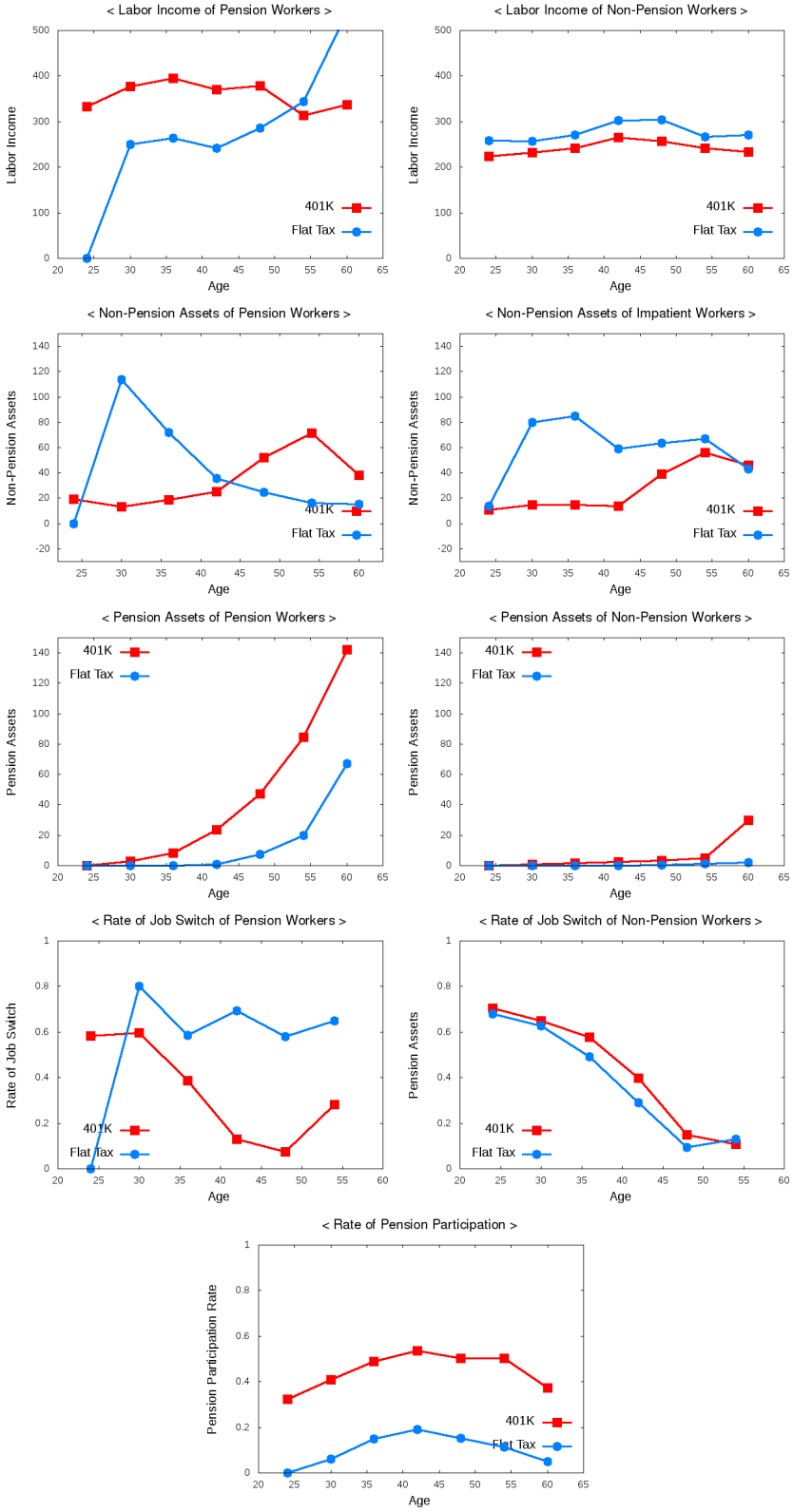

Figure 14: Counter-factual Simulation 2: Flat Income Tax