An Equilibrium Model of Pension Provision and Wage Determination*

Fumihiko Suga†

June 24, 2014

Abstract

This paper proposes and estimates a structural model of screening by employer-provided pensions. The model illustrates how employed workers are sorted into jobs with pension plans and jobs without pension plans according to their time preferences and human capital levels. It also explains why firms provide pension plans and why workers with pensions earn more than workers without pensions. The basic structure of the workers' side of the model is a precautionary saving model with heterogeneity in discount rates, human capital accumulation, and pension asset accumulation. In the model, workers with low discount rates (low discounters) work harder and stay longer with an employer, and accumulate more human capital stocks. On the employers' side of the model, there are firms that have an incentive to hire these workers due to a fixed cost of hiring a worker. Firms cannot observe workers' time preferences due to asymmetric information, so the firms with a fixed cost of hiring make contributions to pension plans and offer relatively low human capital rental rates so that only workers who place higher values on future income are attracted by their offer and low wage today discourage high discounters to take their offer. Although this paper focuses on 401(K) plans, the basic idea can be applied to other forms of employer-provided pensions, which is one of the biggest advantage 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 plans and jobs without pension plans. The result of a counter-factual simulation shows that the tax benefit of 401(K) plans has a great impact on the way of workers to be sorted into jobs with pension plans and jobs without pension plans.

Keywords: Pension, Job-market screening

JEL codes: D31, D82, J21, J22, J32, G23

*I am grateful for helpful comments and suggestions from my adviser John Kennan and Christopher Taber. I thank workshop participants at University of Wisconsin-Madison, National Graduate Institute For Policy Study, and Hitotsubashi University. All errors are mine.

†Department of Economics, University of Wisconsin-Madison, suga@wisc.edu

1 Introduction

Employer-provided pensions are quite common in the United States. More than half of employed workers in the United States have employer-provided pensions. Despite its pervasiveness, the role of employer-provided pensions is not obvious. Why do some firms provide pension plans while the others do not? Why do there exist workers with pensions and without pensions? What is the difference between them? And how are employer-provided pensions related to wages? The result of a simple regression analysis suggests that workers with pension plans tend to earn more and work longer than workers without pension plans. How could this happen? This paper proposes and estimates a structural model that can answer these questions. The results of counter-factual simulations show what happens to the wage distribution, asset accumulation, and the rate of participation to pension plans if the government changes the regulation on pensions or the tax codes. In the model presented in this paper, employers offer pensions in order to hire workers with low discount rates (low discounters), and employer-provided pensions play the role of a screening device.

Low discounters are those who place higher values on their future income and lower values on today's consumption and leisure. So firms can screen high discounters out by offering a generous contribution to pension plans in conjunction with relatively low wage: low discounters are attracted by pension benefits which become available only after they retire, while low wage today discourages high discounters to accept their offers. The idea of screening by employer-provided pensions was first proposed by Ippolito (1992). Ippolito focused on a particular feature of typical 401(K) plans: in typical 401(K) plans, employers and employees put money in employees' pension account, and many employers make contributions to "match" employees' contributions. In 2012, around 60 to 65% of employed workers with pension plans had pension plans with this feature. The question is why employers link their contributions to their employees' contributions. Ippolito conjectured that such an apparently-in-favor-of "savors" plans might be intended to attract low discounters who are supposed to be more willing to save money.

The question is why firms are interested in low discounters. Ippolito collected some statistics and regression results that connect several measures of workers' discount rates and their working behavior. Ippolito found that those who are supposed to be low discounters tend to work harder and longer. In an imperfect monitoring environment, for example, firms have an incentive to hire those who are not likely to shirk. Since low discounters place lower values on the leisure today and higher values on their expected value of their future income, they are less likely to under-perform and work harder in order to keep the probability of getting fired. Another example is that if firms need to provide training and the cost of providing training is constant across workers, firms have an incentive to hire workers who work harder and stay longer. Several studies (e.g., Dorsey and Macpherson (1997)) suggests that pension-providing firms are more likely to be firms that provide training.

A big advantage of the idea of screening by pensions is that different types of pension plans can play the same role. Employer-provided pensions in the United States can be categorized into two broad types: defined benefit (DB) plans, the traditional type of pension plans, and defined-contribution (DC) plans such as 401(K) plans. They are so different that it is hard to regard them as playing the same role: A DB plan is a type of pension plan which determines pension benefits according to tenure, wage, and age of workers, while a DC plan is a type of pension plan which determines how an employer and an employee make contributions to the pension plan. Ippolito's idea is one of the few hypotheses that can explain the role of DB pensions and DC pensions in a harmonious way in a sense that DC pensions can play the same role as DB pensions[1]. This is important especially in the United States, because there was a dramatic shift from DB plans to DC plans since the 1980s, even though the rate of participation to pension plans has not changed so dramatically. In addition, there are some empirical findings that are difficult to be explained by other reasons but sorting/screening by pensions. For example, the job turnover rate of workers with DC plans is much lower than the job turnover rate of workers without pensions, even if all the related observed variables are controlled. DB pensions have an effect to make workers to stay longer with the current employer, but DC pensions do not. So this can be a consequence of sorting/screening by pensions: workers with pension plans stay longer with an employer not because of the incentive effect of the pension plans, but because those who are likely to stay longer with an employer tend to be sorted into jobs with pension plans. Therefore it is meaningful to study an economic model which shed lights not only on the tax benefit, but also on the screening by pensions.

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 way; low discounters tend to work harder and accumulate more human capital stocks through learning-by-doing, because they place lower values on today's leisure and higher values on their future wages. In addition, there is a transitory shock on the disutility from working that can be reset by changing employers, so workers who receive a negative shock have an incentive to change employers. However, job switch cause human capital deterioration so that there is a trade-off between today's leisure and human capital accumulation. Since the utility shock is transitory while human capital matters over the rest of their lives, low discounters are less likely to change employers than high discounters do. As a result, low discounters provide larger amount of labor (human capital stocks times hours of work times years of tenure) than high discounters. On employers' side of the model, there are firms that have to pay a fixed cost of hiring a worker for production, such as the cost of providing training. These firms have an incentive to hire less workers with larger human capital stocks, more working hours, and longer tenure. Therefore they prefer low discounters, but they cannot observe workers' time preferences

[1] There is another way of viewing employer-providing pensions in which DB pensions and DC pensions can play the same role: if the income tax is progressive and income after retirement is much lower than the income before retirement, employer and employee can be better off by bringing some fraction of the wage to the retirement income by making contributions to pension plans. Since contributions to pension plans are tax-deferred (the benefit is taxed), employees can save tax payment by lowering taxable income when they are young. In order to check to what extent this matters, I carry out a counter-factual simulation in which government impose flat-rate income tax instead of progressive income tax.

due to the asymmetric information. Then it is beneficial for firms with fixed cost of hiring to offer employer matching contributions in conjunction with human capital rental rates lower than those offered by the other type of firms so that low discounters are attracted by pension benefits while low wage rates today discourage high discounters to accept their offers. The model has only 401(K) plans, but the basic mechanism of the model can be applied to other types of deferred compensations including DB pensions, non-401(K) DC pensions, hybrid plans[2].

The model is estimated by using the Survey of Income and Program Participation (SIPP) data. The SIPP data is individual data which provides a detailed information about types of pension plans, contribution 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 plans and without pension plans according to their time preferences and human capital levels; low discounters are likely to take jobs with pension plans, but there are low discounters who do not take jobs with pension plans because their human capital levels are too low, while there are high discounters taking jobs with pension plans due to high human capital levels. The result of a counter-factual simulation with the estimated model suggests that if the government eliminates the tax benefit from 401(K) plans, the pension balance at age 61 declines by 13%, while the rate of participation to pensions declines only by 28%.

The rest of this paper is organized as follows. Section 2 summarizes the background and previous literature. Section 3 presents the screening model. Section 4 describes the data. The estimation method and the identification strategy are covered in Section 5. Section 6 shows the results of estimation and counter-factual simulations. The last section summarizes the findings and gives concluding remarks.

2 Background and Related Literature

There are several hypotheses on the role of employer-provided pensions. Unfortunately, regarding employer-provided pensions as a screening device has not been so common. When the DB pensions had dominant share in the market, employer-provided pensions were mostly regarded as playing a role of an "incentive contract". Typical DB pensions have a "back-loading" structure in a 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 a full benefit becomes available[3]. This structure of benefit accumulation discourages workers from leaving current employers and augments the fear of getting fired before the normal retirement age. As a result, employees stay longer with the current employer and are less likely to under-perform. In addition, benefit accumulation stops and sometimes even reverses after the normal retirement age, which enables firms to affect the timing of

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

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

employees' retirement. Previous literature on the role of pension plans before the 1990s put emphasis on these incentive effects of DB pensions. The idea of regarding employer-provided pensions as an incentive contract, however, can only be applied to DB pensions. In the United States, DB pensions are being replaced by DC pensions, especially 401(K) plans, since the 1980s, and DC pensions do not have the "back-loading" structure nor the decreasing benefits after the normal retirement age. Even though DC pensions are supposed not to have the "incentive effect", several previous empirical studies found that workers with DC pensions tend to work longer than workers without pensions.

It is more difficult to figure out the role of DC pensions, including 401(K) plans, because DC pensions do not seem to have a direct effect on workers' behavior. If a worker starts working for an employer who offers 401(K) plans, an individual account will be set up for the worker, and the employee, the employer, or both make contributions to pensions. DC pension accounts are basically mobile and regarded as belonging to the worker: if the worker change employers, the worker is allowed to roll over his money in the pension account to a new pension account if the new employer offers pensions, or individual retirement accounts (IRAs) otherwise. Basically, unlike DB pensions, there is no penalty on job switch. The most important thing about DC plans is that the contributions to pension plans made by employees, employers, or both and the earnings accrued on pension accounts are not taxed until the pension benefits become available. There are several studies investigating why pension-providing firms shifted from DB plans to DC plans, and many of them found that the tax advantage of DC pensions is one of the most important factors. It is possible that both employers and employees would be better off by taking advantage of this tax benefit of DC pensions, but the idea of regarding employer-provided pensions as a tax saving device through tax-deferral cannot be applied to DB pensions [4].

Previous literature focusing on the role of employer-provided pensions puts emphasis on the incentive effects (e.g., Lazear (1979)) and the tax benefit of DC pensions, and less attention was paid to the sorting/screening effect of pensions. There are several papers which support the idea of sorting or screening by pensions, and their reasons for supporting the sorting/screening effect are mostly based on two empirical findings: the difference in tenure between workers with and without pensions and low elasticity of participation to pension plans to the rate of employers' matching contributions. The turnover rates of workers with the defined-contribution plans were smaller than the turnover rates of workers without pension plans (e.g., Ippolito (1997, 2002), Even and Macpherson (2005)) . Several papers pointed out that the job turnover rate of workers with DC pensions are lower and the tenure is longer than those of workers without pension plans. Since typical defined-contribution plans do not have an effect to keep workers from leaving the current employers, this can be a result of sorting out high discounters who are likely to be quit sooner. By using NLSY79 and SIPP data, I run a simple OLS in order to see the relationship between workers' tenure and participation to pension plans. The regression results suggest that those who have DC pensions are likely to stay

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

1.5 to 2 years longer with an employer than those who do not [5]. The estimate of the coefficient on the indicator of pension participation is so large that a worker without pension plans will stay with an employer as long as workers with pension of the same age, occupation, industry and education levels when their hourly wage is more than 10 dollars higher. The model presented in this paper has the property that workers with pension plans work longer because low discounters are likely to be sorted into jobs with pension plans. Since the gap of tenure is hard to be explained by any other factors, it can help identification of the distribution of the discount rates.

And there are several papers finding that the estimates of the effect of the rate of employer matching contributions to the pension participation rate are smaller than expected (Kusko, Poterba and Wilcox (1998) and Even and Macpherson (2005)). Even and Macpherson (2005) found that the estimates became larger if they use instrumental variables to deal with the endogeneity of matching contributions (caused by the sorting effect), which can be interpreted that sorting by pensions is operative. If workers are sorted by pensions and human capital rental rates are adjusted to the rate of employer-matching contributions, the rate of participation to pension plans might appear to be inelastic to the rate of employer-matching contributions. In the model presented in this paper, the rate of employer matching is fixed and the human capital rental rate is determined in the equilibrium. For instance, if the rate of employer matching get higher, the equilibrium human capital rental rate will be lower and the fraction of pension workers may not change. The insensitivity to the rate of employer matching is not inconsistent with the model structure, even though the model does not have heterogeneity in the rate of employer matching.

The basic structure of the workers' side of the model is a life-cycle model with human capital, and pension asset accumulation. Heiland and Lee (2012) estimates a life-cycle model with pension and non-pension assets in order to show that the increasing labor supply of older workers since the mid-1990s can be explained by the shift from DB pensions to DC pensions. Since Heiland and Lee (2012)'s model is purely partial equilibrium model, it cannot explain why employers offer pensions and how workers are sorted into jobs with pension plans and jobs without pension plans. This paper is the first attempt to estimate a structural model where workers make decisions about consumption/saving and participation/contribution to pension plans, and firms' incentive to offer contributions to pension plans is explicitly incorporated. It is unique in a sense that the basic structure is a life-cycle model, but it illustrates how workers are sorted into different jobs according to such a "deep" parameter as the discount factor. It is often the case with estimation of a life-cycle model that discount factors are not separately identified from the relative risk aversion coefficients. Since my model has two different assets and the portfolio choice depends on the discount factors and relative risk aversion coefficients, they are separately identified from the life-cycle path of saving and pension assets accumulation, which is a great advantage of my model over the other models with life-cycle models or precautionary saving models.

[5] I control for age, education, occupation, industry, size of employers. In order to eliminate the effect of temporary workers, I excluded workers whose tenure is shorter than one year.

3 Model

The basic setup of the workers' side of the model is similar to Imai and Keane (2004). The human capital accumulation process in my model is simplified[6], but my model has pension asset accumulation and allows workers to choose employers. In addition, the model describes how the wages offered by pension-providing firms are determined 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 which can be avoided by job switch, and pension and non-pension asset accumulation. The employers' side of the model is characterized by two types of firms: Although the production technology of both types of firm is constant returns in labor inputs, there are firms with and without a fixed cost of hiring a worker[7]. There is asymmetric information with respect to workers' discount rates, human capital stocks, and hours of work: I assume that workers' discount rates are completely unobservable to firms, and firms can observe only the the labor supply (human capital stocks times hours of work) but cannot observe human capital levels and hours of work separately[8].

This model works in the following way; workers with low discount rates (low discounters) place higher values on future income and lower values on today's leisure, so they work harder and accumulate more human capital stocks than high discounters do when they are young in order to enjoy high wage after they get older. In addition, low discounters are less likely to change employers because of the transitory shocks on the disutility from working and the human capital deterioration due to job switch, because they put more emphasis on the human capital accumulation over the rest of their lives than transitory utility shocks which matters only today. As a result, low discounters provide more labor (human capital stocks times hours of work time years of tenure) and produce more. On the employers' side, firms with a fixed cost of hiring a worker have an incentive to hire hard-working workers with larger human capital stocks because the hiring cost is constant across workers but the amount of labor supply differs across workers. Therefore these firms prefer low discounters, but discount rates of workers are not observable to firms due to asymmetric information. Thus firms with a fixed cost of hiring make contributions to pension plans in conjunction with human capital rental rates that might be lower than the human capital rental rates offered by the other type of firms in order to screen out high discounters by taking advantage of difference in the value placed on today's leisure and the future income (pension benefits). The advantage of this model is that it

[6]This is because the identification of the parameter in human capital accumulation equation relies on the transition of hours of work and wage rate, but I assume that they are not separately observable and match the moment of labor income at each period.

[7]It may not precise to call constant returns to scale technology when there is a fixed cost of hiring. To be more precise, the amount of goods produced is proportional to the aggregate labor supply, but the hiring cost is proportional to the number of workers hired.

[8]It is too computationally burdensome to estimate the model if I assume that employers can observe human capital levels and hours of work separately. This assumption can be translated into an assumption that the intensity of work differs across workers, and hours of work is too noisy to measure the amount of labor that each worker provides.

can be applied to other forms of deferred compensation, including DB pensions and non-401(K) DC pensions, even though the model presented in this paper covers only the 401(K) plans.

This model is unique in a sense that the basic structure of the model is a life-cycle model but it illustrates how the equilibrium human capital rental rates offered by pension-providing firms are determined when there is a fixed cost of hiring a worker. The fixed cost of hiring in the model can be interpreted as a training cost. There is an empirical evidence presented by previous literature (e.g., Dorsey and Macpherson (1997)) that employer-provided pensions are likely to be offered by firms which provide training. Thus a training cost in an equilibrium model of pension provision and wage determination is not a far-fetched assumption. I assume that firms have to pay the cost only once at the beginning of employment.

For simplicity, I assume that firms cannot make age-dependent wage offer. The Age Discrimination in Employment Act of 1967 prohibits firms to discriminate against older workers. In the model, firms and workers can observe the amount of labor supplied by each worker, and my assumption can be interpreted that firms will be accused of violation of the Age Discrimination in Employment Act of 1967 if the firm offer different wages to workers with the same amount of labor supply but of different ages[9]. In addition, I assume that firms cannot make a tenure-dependent wage offer. If the firms were allowed to make a tenure-dependent wage offer, firms can screen out high discounters by making a deferred wage contract: Firms can screen out high discounters without using pensions if they can offer a lifetime wage profile such that the starting wage rates are lower than workers' productivity, but the wage rates grow at a higher rate than the growth rate of workers' productivity. Therefore I assume that firms are not allowed to do so. This assumption might appear to be very strong, but such a deferred wage contract can be accepted by workers only if firms are trusted by workers; younger workers do not accept low starting wage rates if they doubt that the firms pay as much as they promised as they get older or suspect that firms fire workers before the wage rates catch up with workers' productivity. In the U.S., the job turnover rate is higher and firing a worker is easier than other developed countries, such as Japan, so accepting lower wage rates is riskier for younger workers in the U.S. Or committing to future wage rates may be too risky for the firms when economic prospects are clouded. Therefore the assumption of no tenure-dependent wage can be translated into an assumption of the lack of trust by workers or uncertainty about economic prospects.

From this chapter on, I call the firms with a fixed cost of hiring a worker as pension firms and the other type of firms as non-pension firms, even though pension-provision is an equilibrium outcome[10].

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

[10] I allow for the possibility that firms with a fixed cost of hiring a worker do not offer pension. However, I focus on the equilibrium where firms with a fixed cost of hiring offer pension, because the fit of the model will be worse when there is no pension providing firms.

Workers' Problem

Workers live T periods. They work only for the first $T-1$ periods. In the last period, they do not work and receive social security and pension benefits. The number of time period T is 8, and the length of one period is 6 years [11]. I assume that the last period last 12 years and everyone dies after the last period.

Preference:

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

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

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

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

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

$$\begin{aligned}
\nu_{it} &= \overline{\nu} \text{ with probability } \phi \text{ if } q_{it} = 0 \\
\nu_{it} &= \underline{\nu} \text{ with probability } 1-\phi \text{ if } q_{it} = 0 \\
\nu_{it} &= \underline{\nu} \text{ if } q_{it} = 1
\end{aligned}$$

where $\overline{\nu} > \underline{\nu}$. Therefore workers have an incentive to change employers if the ν gets high, but the human capital deteriorates if they change employers (explained later).
Since the last period can be different from the other periods in that it has different length and there might be bequest motives, the utility from consumption at T takes a form which is different from previous periods:

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

[11] 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

Constraints:

The budget constraint of a 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 saving or non-pension assets at t , τ_{it} is the income tax rates which is the function of income[12], I_{it} is the rate of investment to the pension plans out of the pre-tax income, r is the rate of market returns, and Y_{it} is labor income which depends on the human capital stock H_{it}, hours of work h_{it} and an indicator of working for a pension-providing firm p_{it}. I assume that workers are not allowed to borrow against their pension assets, and $I_{it} \geq 0$. When worker i works for a non-pension firm, I_{it} is the rate of investment to individual retirement accounts (IRAs) out of labor income. I call workers with employer-provided pensions as pension workers here. The budget constraint of worker i at $t = T$ is given by

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

where B_{iT} is social security benefit and P_{iT} is pension asset. I assume that workers receive retirement income in a life-annuity form, and the workers receive social security and pension benefits over 12 years. Note that workers are not allowed to die in debt.
Let $\tau(Y)$ be the income tax rate when the income is Y. The income tax rate for period $t < T$ is given by $\tau_{it} = \tau\left(rS_{it} + (1 - I_{it})\,Y_{it}\right)$. Note that the last period is twice as long as the previous periods. For simplicity, I assume that the income tax rate in the last period is given by the following equation:

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

Pension Asset:

Let p_{it} be an indicator of working for a pension-providing firm. Workers decide whether or not to work for a pension providing firm at each period. The pension balance P_{it} is given by

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

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

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

pension account is given by

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

[13]I assume that $\psi(.)$ is given and firms cannot manipulate[14]. In order to see how the model is sensitive to this assumption, I do simulations with two different $\psi(.)$s for a robustness check. Note that earnings on pension assets are not taxed at period t. The U.S. government sets an upper bound in contributions to pension plans, and the upper bound of contributions to employer-provided pensions is different from the upper bound of contributions to IRAs;

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

Income:

The pre-tax labor income of a worker without pension plans is given by

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

where H_{it} is the human capital stock, $\overline{w}$ 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. I assume that firms cannot observe H_{it} and h_{it} separately, and they offer human capital rental rate based on the amount of labor supply $H_{it}h_{it}$. Therefore there is a trade-off between wage 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 in equilibrium.

Human Capital:

I assume that firm-specific human capital stocks decrease if a worker changes the employers. The human capital depreciation due to job switch can be interpreted as a loss of firm-specific human capital, but the amount of human capital depreciation is not tenure dependent. Since incorporating tenure as a state variable is computationally burdensome, I assume that the fraction of firm-specific

[13]I obtained the rate of matching contribution by calculating the ratio of employer's contribution to employee's contribution for employees whose rate of contribution out of wage is less than 5% and 10% respectively.

[14]I assume $\psi(.)$ is fixed because different combination of $\psi(.)$ and human capital rental rate$\tilde{w}(.)$ can yield same profit, which can cause an identification problem.

human capital is age-dependent and decreasing in age;

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

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

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

The shock ϵ_{it+1} realizes 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,

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

I take Y_{iT-1} as a crude measure of life time income or a proxy for the Average Index of Monthly Earnings (AIME).

Heterogeneity:

I assume that there are workers with high discount factor $\beta_i = \overline{\beta}$ (low discounters) and low discount factor $\beta_i = \underline{\beta}$ (high discounters, $\overline{\beta} \geq \underline{\beta}$)[15]. I estimate $\overline{\beta}$ and $\underline{\beta}$, and the probability of being a high discounter is fixed at 0.5[16].

Bellman Equation:

The worker's problem can be summarize by the Bellman equation. Let PP_{it} be an indicator of the employer of the last period being a pension-providing firm. The Bellman equation of worker i at

[15] Higher β means discounting future value less. So workers with lower β are high discounters.

[16] I assumed the probability of being a high discounter to be fixed because the identification of it is not obvious.

$t < T$ is given by

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

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

Firm's Problem

I assume that there are two types of firms. One is a group of "ordinary" firms (or non-pension firms) that have a constant-return-to-scale technology in labor inputs and pays just as much as each worker produces in equilibrium. The profit of the ordinary firms from a 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 $\overline{w}$ is the human capital rental rate[17]. I assume 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 the equilibrium, and τ_{fica} is the FICA tax rate.

The other type of firms also have constant-returns-to-scale production technology in labor input [18], but the marginal revenue $\tilde{\alpha}$ is higher than $\overline{\alpha}$, and there is a fixed cost of hiring a worker ξ[19]. The hiring cost ξ can be regarded as a cost to provide training, and the marginal productivity of

[17]Note that $\overline{w} H_{it} h_{it}$ is pre-tax income (but FICA tax is already taxed) from which workers make contributions to pensions 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 FICA tax which employers pay directly to the government.

[18]Due to the fixed cost of hiring, the profit is not constant return to scale.

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

workers increase due to the effect of training. I assume that the effect of training is firm-specific, and the marginal productivity goes down to $\overline{\alpha}$ if a worker switches to a non-pension firm. The expected profit of the pension-providing firm from worker i is given by

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

where p_{it} is an indicator of worker i working for the pension providing firm at period t, ξ is the fixed cost of hiring a worker, and $\tilde{\psi}(.)$ is employers' contributions to pension plans:

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

The first term in the solid bracket in Equation (2) is what worker i produce at period t. The second term in the solid bracket is the sum of the wage paid to this worker and employers' contributions to pension plans. The last term is the fixed cost of hiring a worker. Pension providing firm j's profit maximization problem can be written as

$$\tilde{\Pi}_j = \max_{\tilde{w}(Hh)} \sum_{i \in N_j} \left\{ \sum_{t=1}^{T-1} \left(\frac{1}{1+r} \right)^{t-1} p_{it} \left[\tilde{\alpha} H_{it} h_{it} - \left(\frac{1+\tau_{fica}}{1-\tau_{fica}} \right) \left(1 + \psi I_{it} \right) \tilde{w} \left(H_{it} h_{it} \right) H_{it} h_{it} - \xi \mathbf{1} \left[q_{it} = 1 \right] \right] \right\}$$

$$\text{s.t.} \quad \tilde{V}_t(S_{it}, P_{it}, H_{it} | p_{it} = 1) > \tilde{V}_t(S_{it}, P_{it}, H_{it} | p_{it} = 0)$$

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 life-time utility after period t when worker i chooses pension/non-pension firms at period t. The human capital rental rate $\tilde{w} \left(H_{it} h_{it} \right)$ is determined in equilibrium so that the expected profit from a worker is positive, but almost zero.

I solve firms problem to estimate the model under assumption that firms with a fixed cost of hiring have an incentive to offer pension plans. However, in order to check if the participation constraint (of providing pension) is satisfied or not, I do a simulation by using the estimated model where a firm with a fixed cost of hiring do not to offer pension and choose the profit maximizing human capital rental rate function when all the other firms with a fixed cost of hiring offer pensions and equilibrium human capital rental rate function at the estimates, and workers can choose job offers with and without pension plans from employers with a fixed cost of hiring as well as the job offers without pension plans from firms without fixed cost of hiring.

Equilibrium

I assume that the economy is overlapping generations with no population growth and no aggregate shock. An equilibrium of the 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 rate $\overline{w}$ and $\tilde{w}^{*}(.)$, consumption C_{it}^{*}, hours of work h_{it}^{*}, investment to pension assets I_{it}^{*}, participation to pensions p_{it}^{*}, and job switch q_{it}^{*} maximize workers life-time utility. Allowing for workers' response, firms choose the human capital rental rate to offer. I assume that the labor market is so competitive that the firms have no way but to offer the most generous human capital rental rate with positive (but almost zero) profit the 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 the equilibrium.

Since there is no closed form solution to the pension providing firm's problem, I solve the problem numerically in the following way; Let $l_{it} \equiv H_{it}h_{it}$ be the labor supply of a worker i at time t. On the space of l, I take G_l grid points $\{l_1, l_2, \ldots, l_{G_l}\}$, and on the space of $\tilde{w}$ I take G_w grid points $\{\tilde{w}_1, \tilde{w}_2, \ldots, \tilde{w}_{G_w}\}$. For each grid point of l_j, $j \in \{1, 2, \ldots, G_l\}$, I pick up one grid point $\tilde{w}^{(j)} \in \{\tilde{w}_1, \tilde{w}_2, \ldots, \tilde{w}_{G_w}\}$, $j = 1, 2, \ldots, G_l$. By connecting those point $\left\{\left(l_j, \tilde{w}^{(j)}\right)\right\}_{j=1}^{G_l}$ by using monotone cubic spline (Fritsch and Carlson (1980)), we can construct a human capital rental rate function $\tilde{w}(l)$. Since firms have an incentive to offer better deal to workers with larger labor supply, I pick up increasing functions. When the number of grid points of l is 4 and the number of grid points of $\tilde{w}$ is 10, there are 205 candidate human capital rental rate functions. In a competitive market, the firms offer the most generous human capital rental rate with positive expected profit. Therefore, among 205 candidate human capital rental rate functions , I choose the one with the highest rate of pension participation with positive profit. Figure (2) illustrates how I compose a function $\tilde{w}(l)$.

4 Data

I use 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. The SIPP data collects a nationally representative sample of individuals aged 15 and older, and they are interviewed every 4 month over 32 to 52 month periods[20]. Even though each panel contains 8 to 13 waves, detailed information about pensions is provided only in topical modules which appears at most on 3 waves in each panel. In addition, some important variables related to pension plans are available only in 1996 panel and after. Therefore I take one wave from each of 1996, 2002, 2004, 2008 panels, and pool the sample to constitute a repeated cross section data[21].

Since I assume that the earnings on pension assets are tax-deferred and all of the pension-

[20]Thus each panel of the SIPP has 8 to 13 waves and each wave is 4-month-long.
[21]I run a regression and eliminate the cohort effect and year effect.

providing firms make contributions to the pension plans in the model, I defined pension workers as those who have tax-deferred pension accounts and whose employer make contributions to match employees' contributions. Workers with different types of pension plans such as DB plans, non-401(K) defined contribution plans, or hybrid plans are excluded. Even though workers with pension plans in the model can choose not to make contributions, I regard workers who make contributions as those who are working for pension-providing firms. This is because workers in the sample do not provide enough information about their pension plans when they do not make contributions. In the model, the firms are supposed to be profit maximizers, so I eliminate self-employed workers and public workers. I focus on male workers, because female workers are facing different problems such as childbirth. Even though the SIPP data covers all individuals older than 15, the number of workers older than 50 with 401(K) is quite small. This is because 401(K) plans were became common since the mid-1980s. Therefore I collect the individuals age 24 to 48, and match the data moment and simulated moments of labor income and the share of pension workers for the first four periods (24 to 29, 30 to 35, 36 to 41, and 42 to 47), and match the moments of pension assets and non-pension assets at the beginning of each period for the first 5 periods (at age around 24, 30, 36, 42, and 47). There are 2791 pension workers while 3780 workers without pension plans [22]. Since workers of different cohorts are mixed, I adjust the price levels by using CPI and eliminate the effect of cohorts and years by regressing all the related variables on year and cohort dummies so that I can regard the sample to be homogeneous, and then calculate the moments from the data.

4.1 Summary Statistics

Figure (1) shows annual income from wage and salary of male full-time employed workers aged 23 to 52 in the SIPP data. It shows that workers with pension plans earn more than workers without pension plans. On the other hand, Figure (3) shows the hourly wage (annual income divided by annual hours of work) and hours worked of male employed workers aged 23-52 in the SIPP data. They indicate that the difference in annual income of workers with pension plans and workers without pension plans is explained by the difference in hourly wages and hours of work: workers with pension plans earn more per hour, and work for longer hours.

4.1.1 Simple Regression Analysis

In order to eliminate the effect of other factors, I regress annual income, hourly wage, and tenure respectively on age, age squared, dummy variables of education levels, 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. In addition, in order to get rid of the effect of employers' preference on the way of providing

[22] Since I eliminate the workers with DB pensions and workers whose employer do not provide match, the number of pension workers for estimation is smaller than the actual number of pension workers.

pecuniary/non-pecuniary compensation, I included the dummy variables of health insurance being fully financed and partly financed. The regression results suggest that workers with employer-provided pensions earn 5,200 dollars more per year, 2.5 dollars more per hour, and stay 22 months longer than workers without pension plans do. The result is summarized in Table (7).

4.1.2 Employers Matching Provision

Figure (4) is the fraction of workers whose employers link their contributions to their employees' contributions in workers with tax-deferred pension accounts. It shows that around 80% of workers with tax-deferred pension accounts have employers whose employers' contributions are entirely or partly linked to their own contributions. For those who reported that their employers' contributions entirely or partly depends on their contributions, I calculate the ratio of employers' contributions to employees' contributions. Figure (5) indicates that many of workers make contributions up to the level to which 100% match is provided.

5 Estimation

5.1 Estimation Method

I estimate the model by using the method of indirect inference (Gourieroux, Monfort and Renault (1993)). The auxiliary models are the following;

$$\begin{aligned}
Y_{it} &= \sum_{t=1}^{4} \omega_{YPt} \mathbf{1}\{p_{it}=1\} \mathbf{1}\{Age=t\} + \sum_{t=1}^{4} \omega_{YNt} \mathbf{1}\{p_{it}=0\} \mathbf{1}\{Age=t\} + \epsilon_{Yit} \\
S_{it} &= \sum_{t=1}^{5} \omega_{SPt} \mathbf{1}\{p_{it}=1\} \mathbf{1}\{Age=t\} + \sum_{t=1}^{5} \omega_{SNt} \mathbf{1}\{p_{it}=0\} \mathbf{1}\{Age=t\} + \epsilon_{Sit} \\
P_{it} &= \sum_{t=1}^{5} \omega_{PPt} \mathbf{1}\{p_{it}=1\} \mathbf{1}\{Age=t\} + \sum_{t=1}^{5} \omega_{PNt} \mathbf{1}\{p_{it}=0\} \mathbf{1}\{Age=t\} + \epsilon_{Pit} \\
q_{it} &= \sum_{t=2}^{5} \omega_{qPt} \mathbf{1}\{p_{it}=1\} \mathbf{1}\{Age=t\} + \sum_{t=2}^{5} \omega_{qNt} \mathbf{1}\{p_{it}=0\} \mathbf{1}\{Age=t\} + \epsilon_{qit} \qquad (3) \\
p_{it} &= \sum_{t=1}^{4} \omega_{pt} \mathbf{1}\{Age=t\} + \epsilon_{pit} \\
q_{it} &= \omega_0 + \omega_1 p_{it} + \omega_2 WR_{it} + \omega_3 S_{it} + \omega_4 P_{it} + \epsilon_{q2it} \qquad (4)
\end{aligned}$$

The coefficients in the four equations are the mean of annual labor income, non-pension assets, pension assets of pension, and the rate of job switch of workers with and without pension plans for the first 4 periods, and the coefficients in the fifth equation are share of workers with pension plans. Since 401(K) plans became so popular after the mid-1980s, many of older workers with 401(K) plans are those who started their career with defined benefit pensions. Therefore I match the moments of the first 4 periods for annual income and the rate of participation, and the first 5 periods for non-pension assets, and pension assets[23].

Let θ be the parameters to be estimated, $\omega(\theta, \tilde{w}(l))$ be the vector of the coefficients in auxiliary model calculated from the simulated data where

$$\omega(\theta, \tilde{w}(l)) = \left\{ \left\{ \{\omega_{Yjt}\}_{t=1}^{4}, \{\omega_{Sjt}\}_{t=1}^{5}, \{\omega_{Pjt}\}_{t=1}^{5}, \{\omega_{qjt}\}_{t=2}^{5}, \{\omega_{qjt}\}_{t=2}^{5}, \right\}_{j\in\{P,N\}}, \{\omega_{pt}\}_{t=1}^{4}, \{\omega_{i}\}_{i=1}^{4} \right\}$$

, and ϖ be the vector of the corresponding coefficients calculated by using the SIPP data. The vector of parameters to be estimated, θ, consists of parameters of workers' preference, parameters in the transition equation of human capitals, marginal product of labor of firms with a fixed cost of hiring $\tilde{\alpha}$, and the fixed cost of hiring ξ.

Estimation Procedure

The estimation takes the following procedure: for a candidate parameter vector $\tilde{\theta} \in \Theta$, I solve the individuals' problem for each candidate $\tilde{w}(l)$ by backward induction to get policy functions of workers. By using the policy functions, I run a simulation and calculate the vector of simulated moments $\omega\left(\tilde{\theta}, \tilde{w}(l)\right)$ and the expected value of profit of pension-providing firms for each $\tilde{w}(l)$. Then I pick up $\tilde{w}^*\left(l, \tilde{\theta}\right)$ which maximize the number of workers who work for pension-providing firms with positive expected profit at $\tilde{\theta}$, and match the simulated moments $\omega\left(\tilde{\theta}, \tilde{w}^*\left(l, \tilde{\theta}\right)\right)$with the moments ϖ calculated from simulated data. The objective function is given by

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

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

5.2 Identification

In general, the identification of the discount factors and the relative risk aversion coefficient is ambiguous when they are estimated by using the life-cycle path of asset accumulation or consumption. However, my model 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, and demand for pension assets depend on preference for future consumption. Therefore those parameters can be identified by matching the life-cycle path of pension and non-pension asset accumulation.

[23]Hours of work is the hours of work during each period, while the other variables are state at the beginning of each period. So I include fifth period only for wage rates, non-pension assets and pension assets.

The parameters in the human capital accumulation equation are identified from the life-cycle profiles of annual labor income. The cost of hiring and marginal product of labor jointly determine the rate of participation to pension plans, but the marginal productivity of labor is related to wages. So 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 and δ_1). Since the disutility from working is also related to hours of work and the depreciation rate is related to wages, $\underline{\nu}$, $\overline{\nu}$, δ_0, and δ_1 are identified from the life-cycle path of hours of work, annual labor income, and the rate of changing employers.

The discount factor $\overline{\beta}$ and $\underline{\beta}$ can be identified from the life-cycle profile of hours of work, non-pension assets and pension assets of pension workers and non-pension workers if low discounters and high discounters are perfectly sorted into pension jobs and non-pension jobs. However, the workers are sorted not only according to time preference, but also according to human capital levels. It is possible that a model with homogeneous β generate the moments which is close to the data moments (the mean of hours worked, non-pension assets, and pension assets), because higher human capital can result in higher hours of work (because of substitution effect), and larger assets. This can cause a serious identification problem with respect to the heterogeneity of the discount factor β. In order to avoid this identification problem, I added the last equation to the auxiliary model: Since pension plans in my model do not have an incentive effect, the difference in the probability of job switch between workers with and without pension plans can only be explained by the sorting/screening by pensions if other related variables are controlled. In other words, the coefficient on the indicator of working for pension-providing firms p_{it} is so large as is observed in the data only when low discounters are sorted into pension jobs. Therefore $\overline{\beta}$ and $\underline{\beta}$ are separately identified by using the last equation of the auxiliary model.

Note that the fixed cost of hiring a worker ξ cannot be pinned down at one value without zero-profit condition when 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 to pension plans is the higher than the rate of participation to pension plans at the other candidate human capital rental rate functions. Therefore, $\tilde{w}^*(l)$ can be an equilibrium human capital rental rate function even if the profit is positive, and $\tilde{w}^*(l)$ keeps being an equilibrium human capital rental rate as long as ξ stays within the boundary $\left[\underline{\xi}^*, \overline{\xi}^*\right]$: if $\xi < \underline{\xi}^*$another human capital rental rate function with larger rate of participation to pension plans will be an equilibrium human capital rental rate function. If $\xi > \overline{\xi}^*$, firms with fixed cost of hiring make negative profits and $\tilde{w}^*(l)$ will not be the equilibrium human capital rental rate anymore. The estimates of the model and the simulated moments of the model does not change as long as ξ stays within this boundary, but the result of the counter factual simulation can change. Therefore I pin down the cost of hiring a worker ξ at $\xi = \overline{\xi}^*$ so that pension firms' profits at equilibrium are zero.

6 Estimation Results and Counter-factual Simulation

6.1 Model Fit

The estimated parameters are given in Table (3). The estimates of discount factor $\overline{\beta}$ and $\underline{\beta}$ are 0.967 and 0.924 in annual terms. The data moments and corresponding simulated moments are given in Figure (6).

Figure (9) illustrates the rate of taking jobs with pension plans by human capital levels across low/high discounters. Note that the majority of workers have human capital stocks of 15 to 30. It shows that low discounters are likely to take jobs with pension plans, and more productive workers (workers with larger human capital stocks) are more likely to take the jobs with pension plans. Figure (10) illustrates the hours worked of pension workers and workers without pension plans by human capital levels. It shows that if we compare the hours of work of individuals with the same human capital level, workers with pension plans work harder than workers without pension if the human capital level (or the wage rate of non-pension job) is higher than 15. It might appear to be odd that hours of work is decreasing in human capital levels, but it is possible when the income effect is so strong.

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. Even though the difference in the estimated discount factors is only 1.5% in annual terms, the behavior of low discounters and high discounters are surprisingly different. Workers with higher β, low discounters, work more hours, accumulate more human capital, and more likely to take jobs with pension plans. The probability of a high discounters to take a pension job is half as low as the probability of a low discounter to take a pension job, and low discounters are more willing to save. Those observations are mostly consistent with economic intuition.

6.3 Human Capital Rental Rate Function

Figure 8 illustrates the human capital rental rate function $\tilde{w}^*(l)$. It shows that the human capital rental rate that the pension firms offer is lower than the marginal productivity if the labor supply over six years is less than 250,000. This means that, if there is a worker who would earn \$20 per hour and work for 2000 hours per year at a non-pension job, the wage rate of a pension job offered to this worker by pension firms should be lower than the wage rate offered by the current employer.

6.4 Robustness Check

I assume that the rate of matching contribution is given. In order to see how sensitive is the model to this assumption, I tried the following two plans;

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

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

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

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

Firms are allowed to change human capital rental rates according to the matching rules. Figure (11) shows that the model is quite sensitive to the parametric assumption of $\psi(.)$. So the model is not robust, the last panel shows that the rate of participation to pension is highest with the original matching rule. This means original one is most likely to be accepted by workers and firms are likely to choose in the equilibrium, because firms try to be as generous as possible in the equilibrium.

6.5 The Importance of the Tax Benefit

In order to see the importance of tax deferral of the earnings on pension assets, I did a counterfactual simulation where the government imposes income tax on contribution to pension plans by workers and firms. In addition, the government imposes income tax on earnings accrued in pension accounts and individual retirement accounts, but do not impose tax when the pension benefit become available. This is very similar to pension plans called "Roth 401(K)". The simulation result is summarized in Figure (12). The fraction of workers who work for pension-providing firms goes up to 80%. It seems there might be coding error.

7 Concluding Remarks

To Be Added.

References

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

[5] Gourieroux, C., Monfort, A., & Renault, E. (1993). Indirect inference. Journal of applied econometrics, 8(S1), S85-S118.

[6] 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.

[7] Heiland, F., and 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).

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

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

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

[11] Kusko, A. L., Poterba, J. M., & Wilcox, D. W. (1998). Employee Decisions with Respect to. Living With Defined Contribution Pensions: Remaking Responsibility for Retirement, 98.

[12] 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.

[13] 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

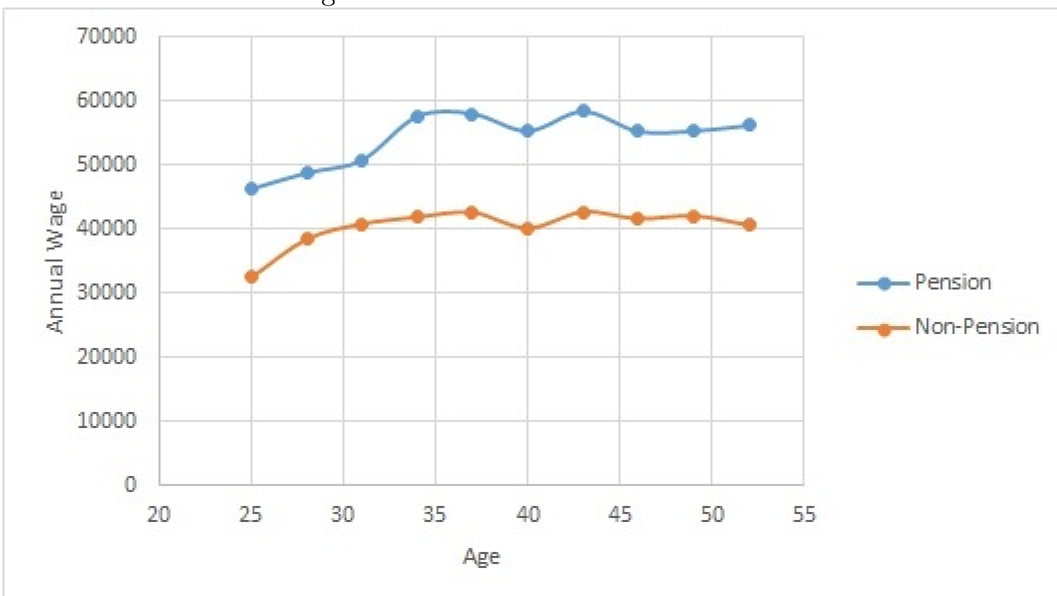

Figure 2: How to Find $\tilde{w}(.)$

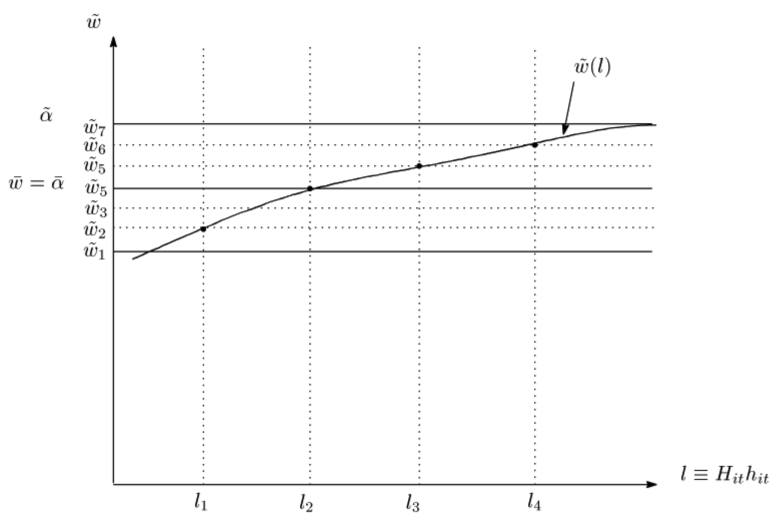

Figure 3: Pension and Hourly Wage (left), Pension and Hours Worked (right)

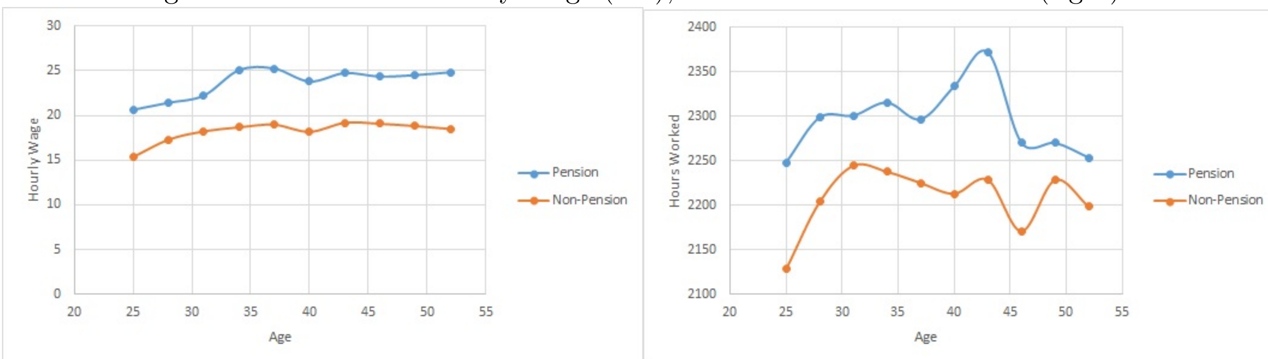

Table 1: Annual, Hourly Wage and Pension

	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.

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 4: The Link between the Contributions of Employers and Employees

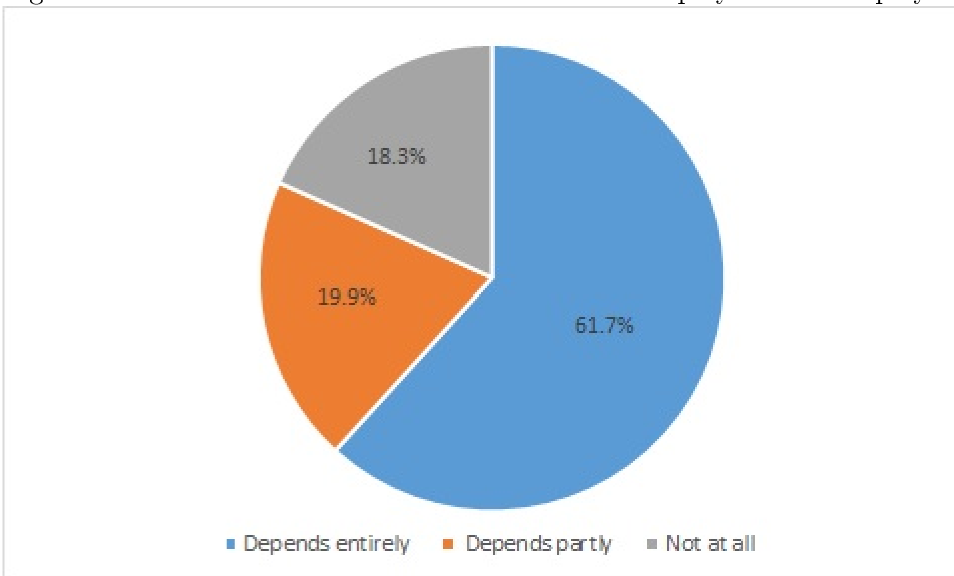

Figure 5: The Ratio of Employers' Contribution to Employees' Contribution

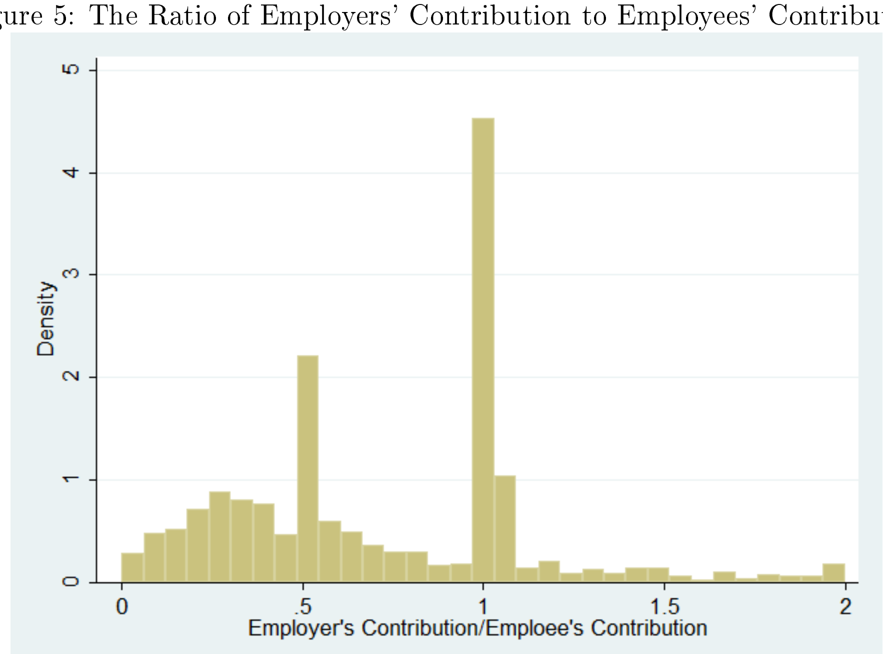

Table 3: Estimates of Parameters

Parameters		Estimates
Utility Function		
Discount Factor (high)	$\overline{\beta}$	0.818 (0.967 in annual terms)
Discount Factor (low)	$\underline{\beta}$	0.623 (0.924 in annual terms)
Weight of disutility from working (low)	$\underline{\nu}$	0.857
Weight of disutility from working (high)	$\overline{\nu}$	1.014
Minimum Consumption at $t = 1$	λ_1	132.339
Coefficient of Increase in $\underline{C}$	λ_2	21.362
Marginal Utility at Consumption at $t = T$	μ	0.688
CRRA coefficient	γ	0.531
Human Capital Accumulation		
Effect of Learning	ζ	0.984
Threshold of Learning	$\underline{h}$	11.839 (1973 hours in annual terms)
Standard Deviation of Shock	σ_h	0.164
Human Capital Depreciation		
Initial Distribution of Human Capital		
The Depreciation at $t = 1$	δ_0	0.984
The Rate of Depreciation Rate Depreciate	δ_1	0.016
Mean of $\ln H_{i1}$of High Type	μ_H	2.998
Pension Firms' Variable		
Marginal Productivity	$\tilde{\alpha}$	1.063
Hiring Cost for Pension Firms	ξ	24.909

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

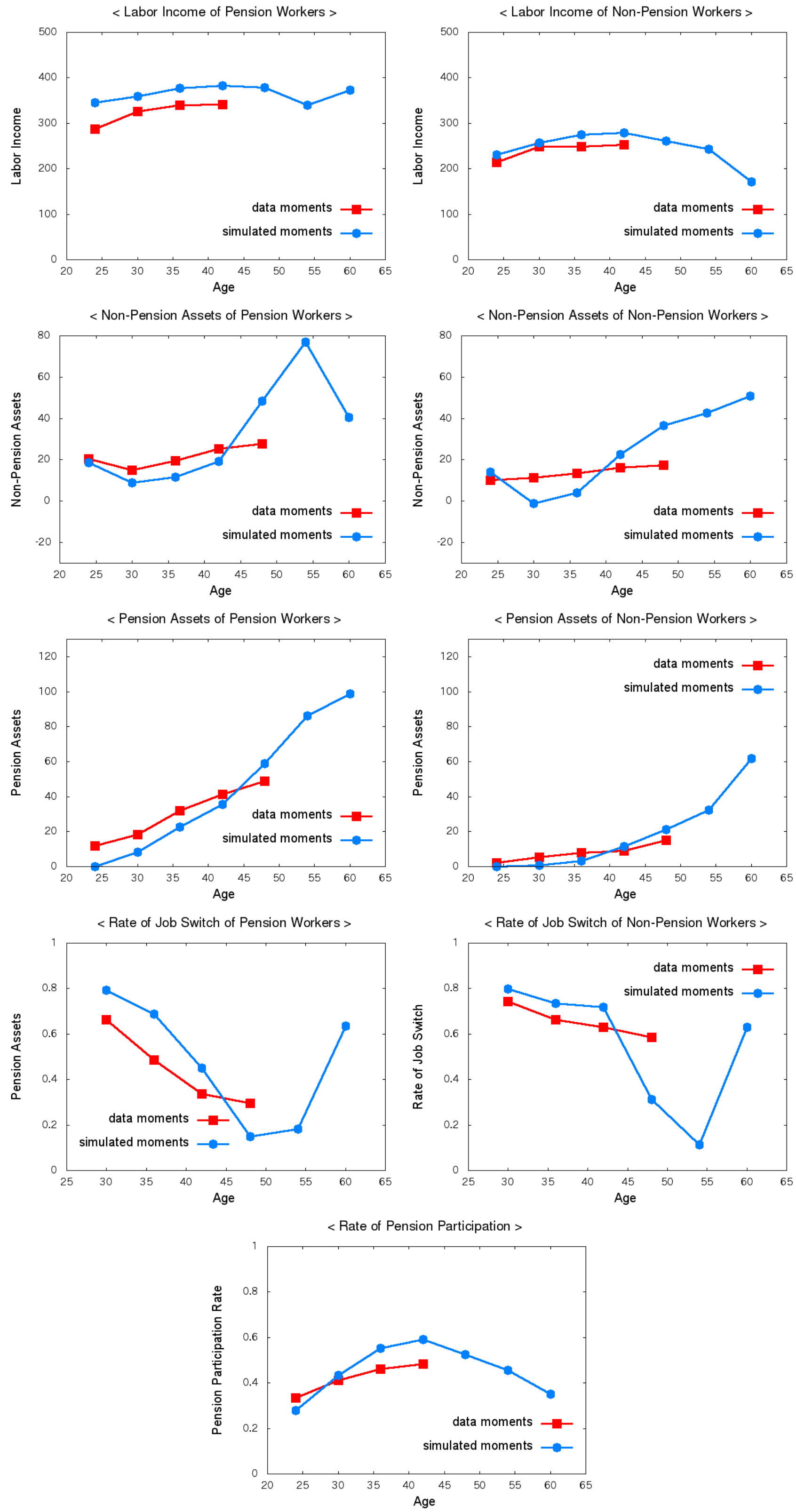

Figure 7: Simulated Moments of High/Low Discounters

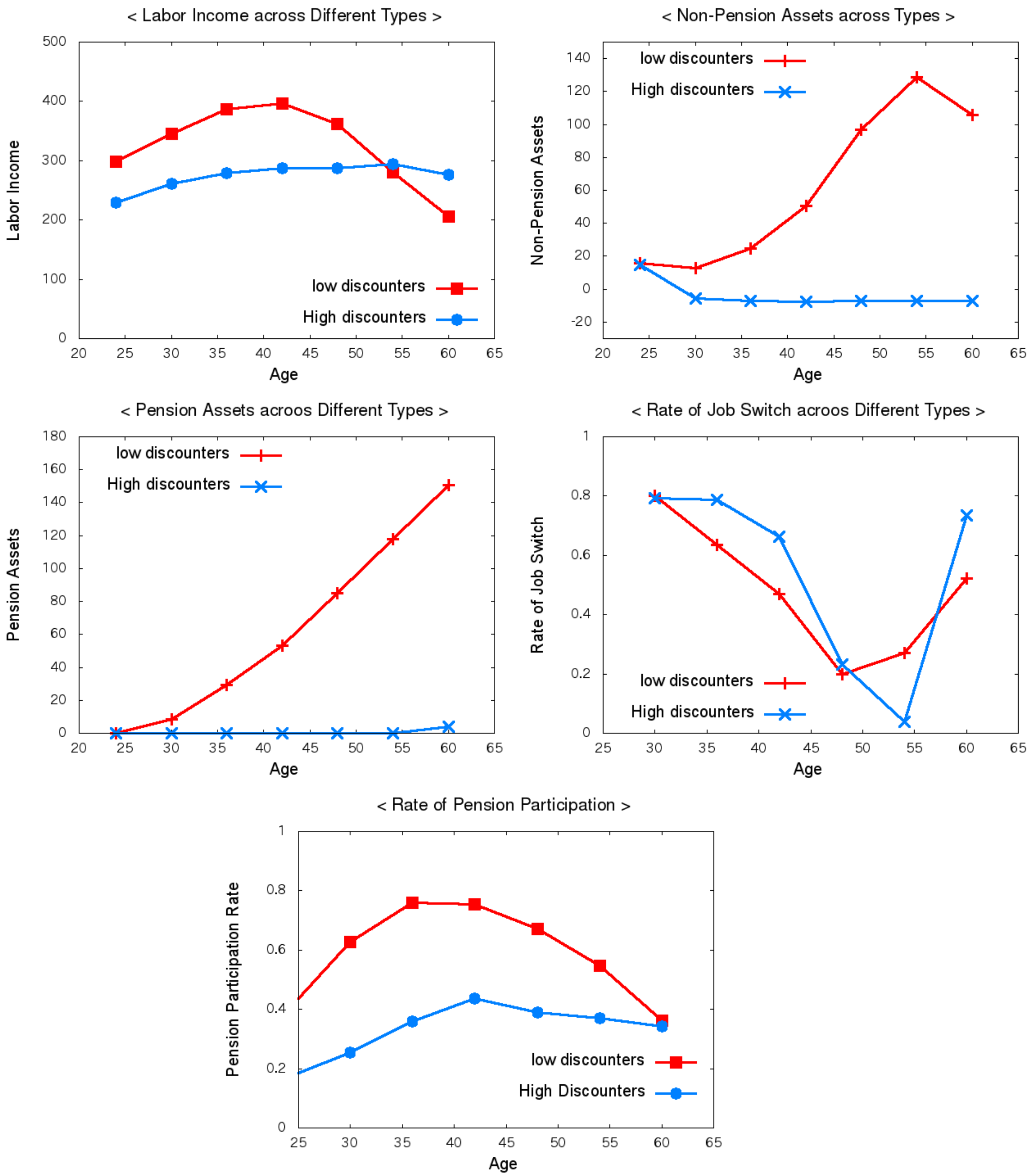

Figure 8: Labor Supply And Human Capital Rental Rate

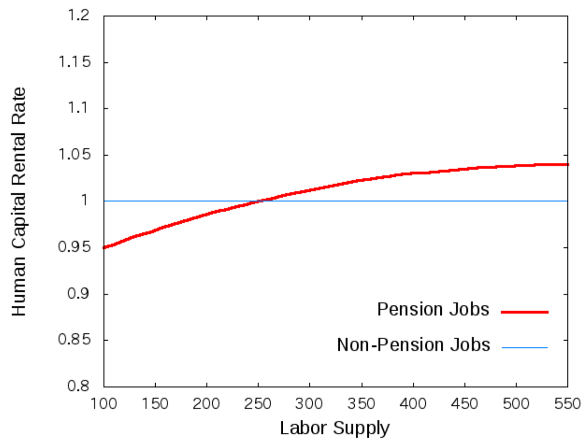

Figure 9:

< Human Capital and the Rate of Participation to Pension Plans by Type >

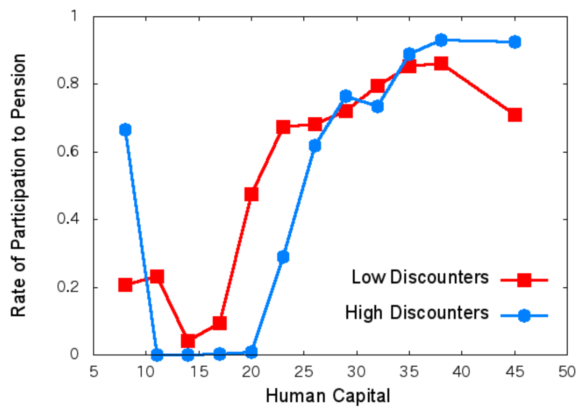

Figure 10:

< Human Capital and Hours Worked of Pension/Non-Pension Workers >

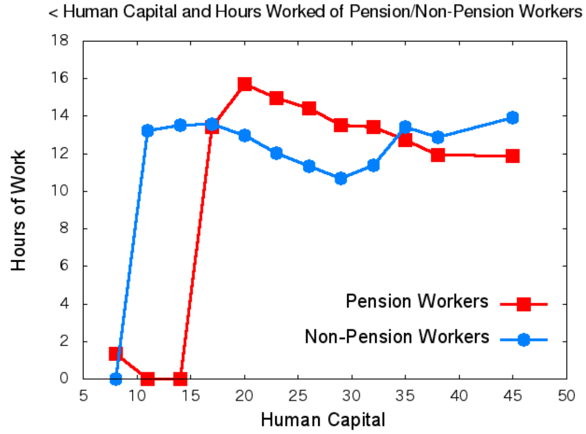

Figure 11: Counter-factual Simulation 1: Different Matching Contribution

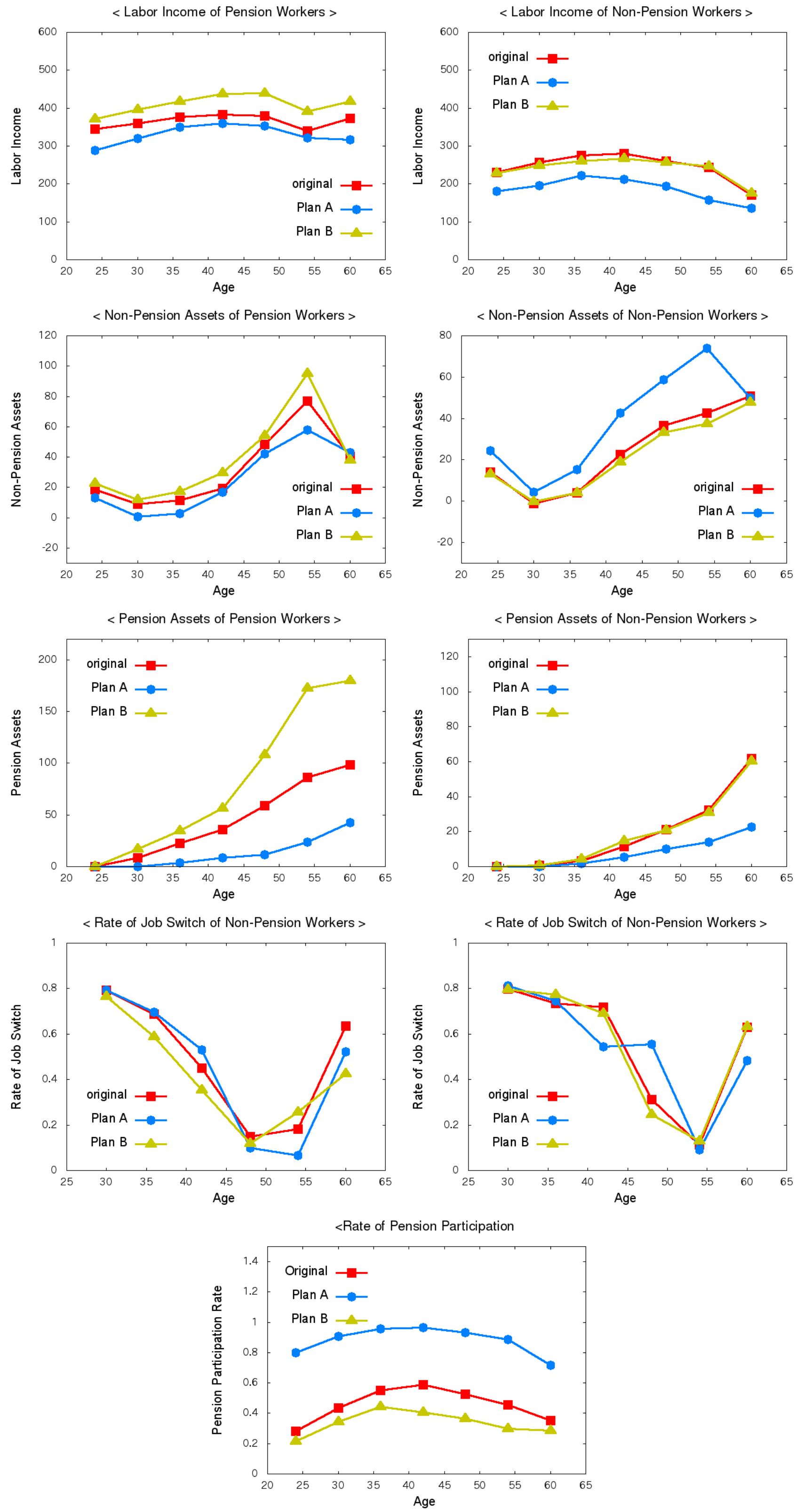

Figure 12: Counter-factual Simulation 2: Contributions and Earnings on Pension Assets are Taxed (Roth 401(K))