Working Paper No. 2018006

Peking University HSBC Business School

Optimal Nonlinear Rates of Public Pension Contribution and Benefit

Insook Lee

Optimal Nonlinear Rates of Public Pension Contribution and Benefit

Insook Lee*

Abstract: Optimal public pension schedule is characterized in a Mirrleesian economy where the risk of outliving retirement savings is not fully insured by financial market. Optimal public pension rates have the components of the redistribution of welfare gain from providing public pension insurance as well as the welfare gain itself. Without income taxation, optimal marginal public pension rates are actuarially fair only for individuals of the highest earning ability. At the optimum, progressive benefit rate may entail progressive contribution rate. With income taxation, optimal marginal public pension rates can be actuarially fair for all individuals regardless of earning ability.

Keywords: optimal public pension insurance, progressive public pension
JEL Code: H55, H21, D60

1. Introduction

Public pension is one of the largest social insurances that have been expanding. To take an example of the US, in 2017, the total spending for Social Security (social insurance program of the US) reached 4.9% of its GDP and the spending for Social Security public pension program (Old-Age and Survivors Insurance) amounted to $813 billion, taking up about 80% of the total Social Security budget and about 25% of the total federal budget.[1] In this light, the design of public pension contribution and benefit rates is quite important; nonetheless, it has not been studied well. Thus, this paper derives optimal nonlinear rates of public pension and explores their features.

While it is run by government, public pension is a pension insurance; hence, it insures against the risk of outliving retirement savings, which differentiates it distinctly from income taxes and from any other social insurances (e.g., disability insurance and unemployment insurance). Thus, the risk of outliving retirement savings is a crucial defining property of

*Corresponding author at: Peking University HSBC Business School, University Town, Nanshan District, Shenzhen, China, 518055. Tel: +86 755 2603 2293 E-mail address: islee@phbs.pku.edu.cn

[1] For further details, refer to the report of 2017 budget review from the Social Security Administration and the 2017 long-term budget outlook from Congressional Budget Office of the US.

public pension; and, it is from longevity shock on retirement savings that finance post-retirement consumption. To incorporate this defining feature of public pension in efficient and tractable way, this paper adopts a two-period life-cycle model where individuals of unobservable earning abilities face uninsurable longevity shocks on their savings after they retire.

From this life-cycle Mirrleesian model, optimal public pension rate formula is obtained without and with income taxation, respectively. The components of optimal public pension rate include the redistribution of welfare gain from insuring individuals against the risk of outliving retirement savings as well as the welfare gain itself. Although public pension benefit is given for post-retirement consumption and depends on pre-retirement labor earnings, optimal public pension rate keeps the optimal labor supply distortion from directly affecting marginal utility of post-retirement consumption but makes it intra-temporal to minimize efficiency loss to get the redistributive welfare gain. Moreover, optimal nonlinear rates of public pension contribution and benefit that implement socially optimal allocation are not uniquely determined. Nevertheless, we can delineate the following features of all the implementable optimal nonlinear public pension rates that satisfy the obtained formula. Without income taxation, optimal marginal public pension rates are actuarially fair only for individuals of the highest earning ability, whereas optimal marginal public pension contribution and benefit rates can be strictly positive for those of the highest earning ability. With income taxation, optimal marginal public pension rates can be actuarially fair for all individuals regardless of their earning ability. Comparison of optimal public pension rates without and with income taxation reveals that introduction of income tax makes implementation of optimal public pension rates less restrictive.

Under an empirically realistic condition, optimal public pension contribution rate is progressive whenever optimal public pension benefit rate is progressive. With income taxation, this finding is extended that the sum of optimal income tax and public pension contribution rates is progressive whenever optimal public pension benefit rate is progressive. Although progressive rates entail labor supply distortion, the government can trade the distortion off with redistributive welfare gain from the progressive rates.

The rest of this paper unfolds as follows. Section 2 reviews related literature. Section 3 describes the model which is analyzed without government in Section 4 and with government in Section 5 and 6. Lastly, Section 7 concludes the paper.

2. Literature Review

In contrast to a great amount of the literature established on reforming public pension, optimal public pension rate schedule is not yet studied extensively. Diamond and Mirrlees (1978, 1986) first studied optimal public pension in terms of consumption and labor supply paths of ex-ante identical individuals with no uncertainty on longevity. In particular, Diamond and Mirrlees (1978, 1986) assumed that individuals can face a negative health shock which disables them so that they have to retire, while government cannot observe the health shock realization to tell whether individuals choose to retire or have to. While Diamond and Mirrlees (1986) assumed no saving and binary labor supply choice, Diamond and Mirrlees (1978) did not. Both found that the present value of net optimal benefit is affected by retirement date and optimal public pension benefit increases with the age of retirement. However, their finding is from optimal paths of consumption and labor supply by implicitly assuming that public pension is the only policy in effect. Neither of Diamond and Mirrlees (1978 and 1986) has any variable that explicitly corresponds to public pension contribution or benefit, although both are seminal for the theoretical research on optimal social insurance.

Due to the nature of ill-health risk, various theoretical research following Diamond and Mirrlees (1978), like Whinston (1983) and Boadway et al. (2006), addressed optimal level of disability insurance, instead of public pension insurance. Note that optimal disability insurance problem is not isomorphic to optimal public pension insurance problem. Clearly, in many crucial aspects including the relevant risk, public pension insurance is different from other social insurances like disability insurance, unemployment insurance, and public health insurance.

Employing variables of public pension contribution and benefit explicitly, numerical simulations such as İmrohoroglu, İmrohoroglu and Joines (2003) and Yew and Zhang (2009) were conducted with the US data and reported their computation results on optimal flat rates of public pension contribution and benefit. Although they are useful for the US public pension policy, their data-specific outcomes of optimal public pension rates are not applicable for any other countries or for the US with different values of data. In addition to the inevitable lack of generality, those computation outcomes themselves do not transparently show general property of optimal public pension rate or its logical underpinnings. On the other hand, Feldstein (1985) theoretically characterized optimal level of public pension benefit from a model where no uncertainty exists and labor supply is exogenously given. Feldstein (1985) regarded myopia as the source of outliving retirement savings and assumed that individuals differ only in the degree of myopia and face the same flat rate of public pension benefit. However, according to this model, the risk of outliving retirement savings is surely realized only for individuals who are myopic, which is certainly known at the moment when they are born. Most of all, none of these studies (Diamond and Mirrlees, 1978, 1986; Feldstein, 1985; İmrohoroglu, İmrohoroglu and Joines, 2003; Yew and Zhang, 2009) studied nonlinear rates of public pension, while actual public pension rates are nonlinear.

By allowing different rates for different individuals, Cremer, Lozachmeur and Pestieau (2004) addressed optimal nonlinear rates of income tax and public pension benefit for an economy with certain longevity and zero interest rate (i.e., no saving). Specifically, they assumed that individuals differ in both wage and disutility of remaining in the labor force which are unobservable to a social planner. In deriving their optimum, however, Cremer, Lozachmeur and Pestieau (2004) reduced the two rates into one net tax/transfer rate and did not explicitly analyze optimal public pension contribution or benefit rate. On the other hand, Farhi and Werning (2013), and Golosov, Troshkin and Tsyvinski (2016) stated that they explore optimal nonlinear rates of social insurance arrangement; however, they did not include public pension contribution or benefit rate in their model but analyzed labor income tax.[2] Rather, optimal labor supply distortion is characterized extensively by delineating its time-series property (Farhi and Werning, 2013) or its limiting behavior (Golosov, Troshkin and Tsyvinski, 2016). Similar to the model of this paper, Farhi and Werning (2013), and Golosov, Troshkin and Tsyvinski (2016) assumed that individuals differ in unobservable earning abilities (skills) and retire in a given age. Unlike this paper, however, in their model, earning ability of each individual is subject to frequent shocks (i.e., shock in every period), while neither longevity nor retirement savings faces any shock.

In sum, although the previous studies have provided useful frameworks for exploring optimal public pension rates, none of them derived a formula for optimal nonlinear rates of public pension contribution and benefit or demonstrated its property with incorporating the relevant risk (longevity risk on retirement savings). Presumably, the risks of low earning ability (skill) and ill-health problem are more relevant to unemployment insurance and disability insurance than to public pension insurance.

3. The Model

Let us consider a life-cycle Mirrleesian economy that is populated by a continuum of individuals who live for two periods working for the first period and then being retired for the second period. Following the canonical model of Mirrlees (1971) and its dynamic version of

[2] Similarly, while Grochulski and Kocherlakota (2010) did not analyze public pension rates but optimal history-dependent capital asset tax rates only, Grochulski and Kocherlakota (2010) referred to optimal asset taxation as social security system because its history dependency resembles Social Security of the US.

Golosov, Tsyvinski and Werning (2006), individuals differ only in their earning ability θ, as their identifier, that cannot be observed and verified by the government. In contrast, earning ability θ is known privately to each individual at the very beginning of the first period before he makes any decisions and it does not change until he dies at some *uncertain* point in the second period. The earning abilities are distributed according to a publicly known distribution $F(\theta)$ whose density function is $f(\theta)$ and support is $\Theta \equiv [\underline{\theta}, \bar{\theta}]$ with $\underline{\theta} > 0$. For period *t*, the utility function of each individual is stated as

$$u(c_t) - v(l_t) \tag{1}$$

where c_t and l_t are consumption and labor supply, respectively, of period *t*. As usual, $u' > 0$, $u'' < 0$, $v' > 0$, and $v'' > 0$. In addition, an individual of higher θ (i.e., a more able worker) has less marginal disutility of working. That is,

$$\frac{d^2 v}{dl d\theta} = \frac{dv'}{d\theta} < 0 \tag{2}$$

which can be called as single crossing property. $y_t \equiv \theta l_t$ refers to labor earnings of an ability-θ individual in period *t* and publicly known. Following the literature based on Mirrlees (1971), the government cannot observe labor supply of each individual while it can observe labor earnings of each individual. This standard assumption of unobservable amount of labor supply is necessary for individuals to keep their own innate earning ability as private information. Without this assumption, it is immediate that the government can precisely infer and verify the earning ability of each individual from the public information of the labor earnings. Since individuals are retired in the second period, $l_2 = 0$ and $y_2 = 0$. Thus, omitting the time subscripts for labor supply and labor earnings is concise without begetting confusion. Then, the life-time utility of an ability-θ individual, evaluated in the first period, is

$$u(c_1) - v(\frac{y}{\theta}) + \beta E[u(c_2)] \quad (3)$$

where β is time preference parameter. The production technology of this economy is linear so that the total output is the sum of labor earnings of all individuals. In addition, interest rate, *r*, is exogenously given.

Since typical pension is in the form of a guaranteed life annuity, the risk that general pension insurance protects insurees against is the risk of outliving their retirement savings (which is essentially longevity risk on retirement wealth). No matter how appropriately an individual saves for his retirement while working, he can end up with living longer than expected so that his retirement savings turn insufficient for post-retirement consumption. To incorporate this risk into our model in an efficient way, longevity shock is introduced on the savings that are made in the first period and used for the consumption during retirement (the second period). In particular, in the second period, an individual can face some cuts in his savings that are made in the first period, to capture the state where he lives longer than expected finding himself be left with insufficient[3] resources for his post-retirement consumption. Likewise, in the second period, the individual also can face some increase in the savings, to describe the state where he lives shorter than expected. There are a finite number of the states that can be realized in the second period. Let S denote for the set of all the possible states in the second period. Each state in S is indexed by *s* and has strictly positive probability of realization $\mu(s)$; thus, $\mu(s) \in (0,1)$ for $\forall s \in S$. Indeed, a strictly positive probability is assigned on the state where an individual lives exactly for the expected life span so that his savings from the first period (i.e., retirement savings) do not increase or decrease in the second period. From the first-period point of view, which state *s* will be realized in the second period is not certain, although the realization probability is known in

[3] In the life-cycle context of this paper, it refers to being insufficient to smooth consumption before and after retirement (across the two periods).

the first period. Moreover, the realization of any state s is independent of earning ability θ. Notice that in our model an unexpected increase or decrease in the retirement savings, which stems from the longevity shock in the second period, occurs on the principal amount of the savings itself, without a change in the rate of return r that is certain.

Moreover, unlike the innate earning ability θ, the realization of state s is known to the government, instead of being privately known to each individual. When individuals decide how much they save in the first period, they have to form an expectation on their own longevity based on the average life expectancy. In the second period, depending on the realized state s, some individuals end up with living longer than the average life expectancy (and thus finding their own retirement savings insufficient) while others do not. The average life expectancy (an aggregate variable) and post-retirement consumption and actual life span of an individual are public information. Therefore, it is feasible for the government to know which state s is realized to each individual in the second period.

Noticeably, the above-described risk structure of our model means that the risk of outliving retirement savings is not fully insured by financial market in this economy. The previous studies on optimal social insurance (e.g., Diamond and Mirrlees 1978, 1986; Farhi and Werning, 2013; Golosov, Troshkin and Tsyvinski, 2016) also assumed that private insurance market is not complete. As a matter of fact, we could find some evidence for the incompleteness of private pension insurance market. For instance, in reality, although defined-benefit pension plan provides more complete insurance against the risk of outliving retirement savings than defined-contribution pension plan does, the private provision of defined-benefit pension plan is not more than that of defined-contribution pension plan but has been declining (e.g., Butrica et al. 2009). In addition, the well-known observation of insufficient level of annuity to ensure stable post-retirement consumption, which is called as ‘annuity puzzle,’ would be another example that indicates the lack of full private insurance against the risk of outliving retirement savings provided by financial market.

The available policy instruments that the government of this Mirrleesian economy utilizes resemble and are no better than those of government in reality, to provide more relevant policy implications. Specifically, public pension contribution rates are based only on the current labor earnings (not on the past ones), whereas public pension benefit rates depend on the accumulated contributions in the past. In addition, the government also can impose labor income tax as well; and, income tax rates also depend only on the current labor earnings. That is, neither public pension contribution rates nor labor income tax rates is history-dependent, while public pension benefit rates are history-dependent. On the other hand, Weinzierl (2011), Farhi and Werning (2013), and Bastani, Blomquist and Micheletto (2013) showed that if history-dependent (age-dependent) tax on labor earnings is implemented, it can improve social welfare. Nonetheless, our model reflects the fact that history-dependent rates of income tax or public pension contribution are not yet implementable.[4] Since public pension is of our main interest, we introduce income taxation after analyzing the case where the government runs only public pension program.[5]

In contrast, however, even though individuals make savings (investment) in the first period, capital income taxation is not considered in the present analysis. Notice that, in our model, the longevity shock is realized on the amount of the savings and translated into an unexpected change in the total amount of capital income. As a result, if the government implements capital income taxation, then part of the capital income tax is levied effectively on the realized longevity shock, which is misleading and improper representation of capital income taxation. Therefore, with our model, only the cases where the government implements one or

[4] In addition, history-dependent (age-dependent) contribution for public pension and income tax may face some political and legal issue, such as age-discrimination, if government attempts to legalize them for implementation.

[5] Since our model is not overlapping generations, pay-as-you-go aspect of public pension is not fully incorporated. Nonetheless, introducing pay-as-you-go aspect into our model will simply add a government budget constraint to the social welfare maximization problems of this paper. Therefore, since the theoretical findings of public pension from our model are more general without the additional government budget constraint than with it, they can be extended for pay-as-you-go public pension.

both of public pension and labor income taxation are explored.

4. Laissez Faire Economy

To set a benchmark, let us start with characterizing optimal allocation of this economy with no government. The best-possible allocation of this decentralized economy is defined by the solution of the following constrained maximization problem: for any given $\theta \in [\underline{\theta}, \bar{\theta}]$,

$$\max u(c_1) - v(l) + \beta \sum_{s \in S} u(c_2(\theta; s))\mu(s) \text{ s.t } c_1 \le \theta l - k \text{ and } c_2 \le (1+r)k(s) \tag{4}$$

where k is the amount of savings made in the first period. Since there is no government, the information asymmetry on the earning ability θ does not impose any constraint on the problem of (4). For characterizing the laissez faire economy optimal allocation, $\{c_1^{lf}(\theta), l^{lf}(\theta), \{c_2^{lf}(\theta; s)\}_{s \in S}\}_{\theta \in \Theta}$, the intra-temporal optimality condition for (4) is stated as follows. For $\forall \theta \in [\underline{\theta}, \bar{\theta}]$,

$$\theta = \frac{v'(l^{lf}(\theta))}{u'(c_1^{lf}(\theta))} = \frac{v'(\frac{y^{lf}(\theta)}{\theta})}{u'(c_1^{lf}(\theta))} \tag{5}$$

which means no distortion in labor supply as wage rate is equated with marginal disutility of labor in terms of marginal utility of consumption (which is financed by the wage income). In other words, since there is no government intervention on the effective price of labor supply, the first-best level of labor supply is attained in this laissez faire economy.

On the other hand, however, obtaining the inter-temporal optimality condition for (4) is not as straightforward as obtaining the intra-temporal optimality condition for (4). In usual dynamic models, the principal amount of savings itself stays the same over two adjacent periods so that individuals adjust the current savings to equalize the present-value marginal utilities of consumption in both current and next periods, meeting the standard consumption-smoothing Euler equation. In contrast, however, in our model, since the principal amount of savings made in the first period does not always stay the same but may change due to the longevity shock in the second period, the standard Euler equation is not directly applicable. Facing this uncertainty, in the first period, an individual might choose his savings to equate the marginal utility of the current consumption with the present value of the marginal utility from the second-period consumption in the worst-possible state. However, there is a substantive chance that the worst-possible state is not realized in the second period, which entails excessive sub-optimal savings. This failure in the consumption smoothing still can happen even when the individual chooses his savings to equate the marginal utility of the current consumption with the present value of the marginal utility of the second-period consumption in the state of living exactly for the average life span or any other state in the set *S*. Moreover, in the unit of resources for consumption, taking a derivative with respect to consumption for the second period cannot derive one optimality condition, like (5), since there are several possible second-period consumption levels depending on the realized state.

Alternatively, we state the dual problem of (4) as follows: for any given $\theta \in [\underline{\theta}, \bar{\theta}]$,

$$\min\ c_1(\theta) + \frac{\sum_{s\in S} c_2(\theta; s)\mu(s)}{(1+r)} \quad \text{s.t}\ V^{lf}(\theta) = u(c_1^{lf}(\theta)) - v(l^{lf}(\theta)) + \beta \sum_{s\in S} u(c_2^{lf}(\theta; s))\mu(\mathrm{s}). \quad (6)$$

Now, at the optimum allocation of $\{c_1^{lf}(\theta), l^{lf}(\theta), \{c_2^{lf}(\theta; s)\}_{s\in S}\}_{\theta\in\Theta}$, let us introduce a little deviation from $u(c_1^{lf}(\theta))$ to $u(c_1^{lf}(\theta)) - \Delta$ via an infinitesimal change in the savings which is translated in the unit of utility. This perturbation entails $u(c_2^{lf}(\theta; s)) + \frac{1}{\beta}\Delta$ for $\forall s \in S$ to maintain the level of $V^{lf}(\theta)$ in (6). Moreover, since $u' > 0$, the inverse function of u, notated as u^{-1}, is uniquely defined. Then, the dual problem (6) is restated in terms of such Δ as follows: for any given $\theta \in [\underline{\theta}, \bar{\theta}]$,

$$\min\ u^{-1}(u(c_1^{lf}(\theta))-\Delta)+\frac{\sum_{s\in S}u^{-1}(u(c_2^{lf}(\theta;s))+\frac{1}{\beta}\Delta)\mu(s)}{(1+r)}\quad \text{s.t}\ \ \Delta=0. \tag{7}$$

Then, obtaining the first-order condition of (7) with respect to Δ and setting Δ as zero (meaning no deviation of Δ chosen at an optimum) yield the inter-temporal optimality condition for (4) as follows: for any given $\theta\in[\underline{\theta},\overline{\theta}]$,

$$\frac{1}{u'(c_1^{lf}(\theta))}=\frac{1}{\beta(1+r)}\sum_{s\in S}\frac{1}{u'(c_2^{lf}(\theta;s))}\mu(s). \tag{8}$$

This takes the form of the inverse Euler equation as in Golosov, Tsyvinski and Werning (2006) and Farhi and Werning (2013), although there is difference that our inverse Euler equation of (8) does not have a social planner (government) meet some incentive constraints while theirs does with the principal amount of individuals' savings being unchanged over two adjacent periods. Notice that when (8) is met, the standard Euler equation of consumption smoothing, which equalizes marginal utility of the first-period consumption with the expected present value of marginal utility of the second-period consumption, may or may not be met. In fact, in this laissez faire economy, the standard Euler equation is not met for any individual.

Lemma 1. At the optimum of laissez faire economy, $\{c_1^{lf}(\theta),l^{lf}(\theta),\{c_2^{lf}(\theta;s)\}_{s\in S}\}_{\theta\in\Theta}$, individuals cannot smooth consumption at ex ante point of view by meeting the standard Euler equation. The marginal utility gain unattainable due to the failure in the consumption smoothing is $\beta(1+r)\sum_{s\in S}u'(c_2^{lf}(\theta;s))\mu(s)-u'(c_1^{lf}(\theta))>0$ for any given $\theta\in[\underline{\theta},\overline{\theta}]$.

Proof. See Appendix A.

Obviously, **Lemma 1** suggests that the optimal inter-temporal allocation of consumption of laissez faire economy is not the first-best, unlike the first-best level of labor supply $\{l^{lf}(\theta)\}_{\theta\in\Theta}$. Although there is no distortion generated by a government, the incomplete financial market for insuring against the longevity shock on retirement savings leads to the failure in the consumption smoothing. Therefore, the optimal allocation of laissez faire economy leaves a room for a government to improve individuals' utility by providing insurance against the risk of outliving retirement savings, even without any redistribution among individuals. In this light, we introduce public pension insurance to this economy in the next section.

5. Economy with Public Pension

Now, a government is installed in this economy and it runs public pension program. In particular, the government collects public pension contribution from labor earnings of each individual in the first period, and public pension contribution depends only on the labor earnings for the first period. After individuals are retired (i.e., in the second period), the government gives them public pension benefits regardless of whether they end up with living longer than expected or not, to ensure the individuals against the risk of outliving their retirement savings. By the nature of pension insurance, public pension benefit is based on the total amount of the public pension contribution. Consequently, public pension benefit paid in the second period also depends on the first-period labor earnings. The amount of public pension contribution collected from an ability-θ individual is denoted as $S(y(\theta))$ whereas the amount of public pension benefit paid to the ability-θ individual is $\{B(y(\theta);s)\}_{s\in S}$. Assume that both $S(y(\theta))$ and $\{B(y(\theta);s)\}_{s\in S}$ are continuously differentiable for $\forall\theta\in[\underline{\theta},\bar{\theta}]$. $S'(y)$ and $B'(y;s)$ may vary for different earning abilities, instead of being assumed to be constant (flat rate).

Given the public pension program $\{S(y(\theta)),\{B(y(\theta);s)\}_{s\in S}\}_{\theta\in\Theta}$, for any given $\theta\in[\underline{\theta},\bar{\theta}]$, an ability-$\theta$ individual maximizes his life-time utility of

$$u(c_1)-v(\frac{y(\theta)}{\theta})+\beta\sum_{s\in S}u(c_2(\theta;s))\mu(s) \tag{9}$$

subject to the following budget constraints

$$c_1 \le y - S(y) - k \ \text{ and } \ c_2 \le (1+r)k(s) + B(y). \tag{10}$$

The optimality condition for his labor supply is

$$u'(c_1(\theta))(1-S') + \beta \sum_{s \in S} u'(c_2(\theta;s))B'\mu(s) = v'(\frac{y(\theta)}{\theta})\frac{1}{\theta}. \tag{11}$$

Comparison of (11) and (5) reveals that, with public pension, labor supply is less than the first-best level. Markedly, the labor supply distortion is inter-temporal affecting not only the marginal utility of the current consumption but also the marginal utility of the future consumption after retirement.

In fact, the asymmetry information on the innate earning ability can beget another labor supply distortion. In light of (2), marginal disutility of working is decreasing in the earning ability. As a consequence, individuals have incentive to reduce their own labor supply for pretending to be of lower earning ability than their true earning ability. To see this, consider an ability-θ' individual with $\theta' > \theta$. Due to (2) and (11), $u'(c_1(\theta))(1-S') + \beta \sum_{s \in S} u'(c_2(\theta;s))$ $B'\mu(s) = v'(\frac{y(\theta)}{\theta})\frac{1}{\theta} > v'(\frac{y(\theta)}{\theta'})\frac{1}{\theta'}$. Hence, the ability-$\theta'$ individual can have a gain from reducing his labor supply to earn $y(\theta)$, instead of $y(\theta')$, so that public pension rate intended for ability-θ individuals is wrongfully applied to him. This is feasible since the government cannot tell an ability-θ' individual from an ability-θ individual if both earn the same level of labor income. On the other hand, due to (2), individuals do not have incentive to pretend to have higher earning ability by taking excessive disutility of over-working. To prevent this type of labor supply distortion, in designing public pension rates, the government needs to give each individual incentive not to pretend to be of other earning ability. Thus, the government faces the following incentive constraint: for $\forall \hat{\theta} \in [\underline{\theta}, \overline{\theta}]$

$$u(c_1(\theta))-v(\frac{y(\theta)}{\theta})+\beta\sum_{s\in S}u(c_2(\theta;s))\mu(s)\geq u(c_1(\hat{\theta}))-v(\frac{y(\hat{\theta})}{\theta})+\beta\sum_{s\in S}u(c_2(\hat{\theta};s))\mu(s)\,. \quad (12)$$

Above all, optimal public pension rate schedule that the government, as a social planner, seeks to design is obtained from solving the following constrained maximization of social welfare.

$$\max\int_{\underline{\theta}}^{\bar{\theta}}\{u(c_1(\theta))-v(\frac{y(\theta)}{\theta})+\beta\sum_{s\in S}u(c_2(\theta;s))\mu(s)\}g(\theta)dF(\theta) \quad (13)$$

subject to the incentive constraint (12) and the aggregate resource constraint[6]

$$\int_{\underline{\theta}}^{\bar{\theta}}c_1(\theta)+\frac{1}{1+r}\sum_{s\in S}c_2(\theta;s)\mu(s)dF(\theta)\leq\int_{\underline{\theta}}^{\bar{\theta}}y(\theta)dF(\theta) \quad (14)$$

where $g(\theta)$ is social weight given to an ability-θ individual and $g(\theta)>0$ for $\forall\theta\in[\underline{\theta},\bar{\theta}]$. Then, optimal schedule of public pension contribution and benefit can be characterized by the first-order conditions of the Lagrangian of this constrained maximization of the social planner government. The incentive constraint (12) is not replaced with the first-order conditions of individuals' maximization, since the budget sets of individuals under a public pension program may be non-convex to make those first-order conditions insufficient for replacing the incentive constraint (12). The optimal allocation of this economy with public pension which solves the above constrained social welfare maximization problem is notated as $\{c_1^{pp}(\theta),\,l^{pp}(\theta),\{c_2^{pp}(\theta;s)\}_{s\in S}\}_{\theta\in\Theta}$.

Proposition 1. Optimal rates of public pension contribution and benefit are defined as follows: for any given $\theta\in[\underline{\theta},\bar{\theta}]$,

$$S'(\theta)-\frac{\sum_{s\in S}B'(\theta;s)\mu(s)}{1+r}=1-\frac{\mathrm{g}(\theta)+\phi(\hat{\theta},\theta)-\int_{\theta}^{\bar{\theta}}\phi(\theta,\hat{\theta})\frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}}{\mathrm{g}(\theta)+\phi(\hat{\theta},\theta)-\int_{\theta}^{\bar{\theta}}\phi(\theta,\hat{\theta})\frac{\theta}{\hat{\theta}}\frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}}=1-\frac{v'(\frac{y^{pp}(\theta)}{\theta})\frac{1}{\theta}}{u'(c_1^{pp}(\theta))} \quad (15)$$

[6] Due to Walras law, once the aggregate resource constraint is met, the total government budget constraint is always met. Thus, for avoiding redundant budget constraints, the total government budget constraint is not additionally imposed on the social welfare maximization problems in Section 5 and 6.

where $\phi(\hat{\theta},\theta)$ is the Lagrangian multiplier of the incentive constraint (12). Under optimal public pension rate schedule, all individuals smooth consumption at ex ante point of view. That is, for any given $\theta \in [\underline{\theta},\bar{\theta}]$,

$$u'(c_1^{pp}(\theta)) = \beta(1+r)\sum_{s\in S} u'(c_2^{pp}(\theta;s))\mu(s) . \tag{16}$$

Proof. See Appendix B.

As revealed from comparing (16) in **Proposition 1** and **Lemma 1** from the laissez faire economy, individuals are benefited from public pension insurance that enables them to smooth their consumption. Through the consumption smoothing with optimal public pension, the optimal labor supply distortion, the right-hand side of (15), turns intra-temporal, instead of inter-temporal as in (11). Markedly, optimal public pension makes the current labor supply distortion no longer directly affect the marginal utility of the future consumption, while public pension benefit that also depends on the current labor earnings finances the future consumption. In Farhi and Werning (2013) and Golosov, Troshkin and Tsyvinski (2016), at their optimum, the intra-temporal labor supply distortion, which is the same as the right-hand side of (15), is equated with the optimal marginal rate applied for the current period while the left-hand side of (15) includes the marginal rate applied for the future period as well. However, both Farhi and Werning (2013) and Golosov, Troshkin and Tsyvinski (2016) are based on a life-cycle Mirrleesian model, like ours. In addition, the present value of the future public pension benefit rate is *currently* taken into account for the current labor supply decision even though public pension benefit rate is not applied in the current period but after retirement in the future. In this sense, our characterization of optimal public pension rates of (15) which balances the present value of net marginal contribution rate against the current labor supply distortion is consistent with their characterization.

Moreover, to be more informative, the optimal public pension rates of (15) in **Proposition 1** can be restated in the unit of utility as follows.

Proposition 2. The optimal rates of public pension contribution and benefit can be decomposed into the components of (A) welfare gain from providing public pension insurance, (B) redistribution of the welfare gain, and (C) efficiency loss from labor supply distortion. That is, for any given $\theta \in [\underline{\theta}, \bar{\theta}]$,

$$S'(\theta) - \frac{\sum_{s\in S} B'(\theta;s)\mu(s)}{1+r} = \frac{1}{\xi(\theta)}\{A(\theta)+B(\theta)+C(\theta)\} \tag{17}$$

$$A(\theta) = \frac{\beta(1+r)\sum_{s\in S} u'(c_2^{lf}(\theta;s))\mu(s) - u'(c_1^{lf}(\theta))}{-u''(c_1^{pp}(\theta))}\gamma(\theta)\,\mathrm{g}(\theta) - \frac{\lambda}{u'(c_1^{pp}(\theta))} \tag{18}$$

$$B(\theta) = \mathrm{g}(\theta) - \frac{\beta(1+r)\sum_{s\in S} u'(c_2^{lf}(\theta;s))\mu(s) - u'(c_1^{lf}(\theta))}{-u''(c_1^{pp}(\theta))}\gamma(\theta)\,\mathrm{g}(\theta) \tag{19}$$

$$C(\theta) = \int_{\theta}^{\bar{\theta}} [v'\mathrm{g}'(\theta) - \lambda(1+\frac{1}{\varepsilon(\theta)})]\frac{1}{v'}d\theta \tag{20}$$

where $\xi(\theta) = \mathrm{g}(\theta) + \phi(\hat{\theta},\theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta,\hat{\theta})\frac{\theta}{\hat{\theta}}\frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}$; λ is the Lagrangian multiplier of the resource constraint (14); $\gamma(\theta) = -\frac{u''(c_1^{pp}(\theta))}{u'(c_1^{pp}(\theta))}$ is the degree of the risk aversion; and, $\varepsilon(\theta)$ $= v'(\frac{y^{pp}(\theta)}{\theta})[v''(\frac{y^{pp}(\theta)}{\theta})\frac{1}{\theta}]^{-1}$ is the labor supply elasticity at the level of $y^{pp}(\theta)$.

Proof. See Appendix C.

Since public pension insures the risk of the longevity shock on retirement savings, individuals now can smooth their consumption, as shown in (16) of **Proposition 1**. Thus, the marginal utility gain of $\beta(1+r)\sum_{s\in S} u'(c_2^{lf}(\theta;s))\mu(s) - u'(c_1^{lf}(\theta)) > 0$, which is unattainable in the laissez faire economy (**Lemma 1**), is now obtained at the cost of public funds whose marginal value is λ. The marginal social welfare gain from public pension insurance increases in the degree of the risk aversion of individuals, as shown in the first term of (18). Through risk-sharing among heterogeneous individuals, public pension insurance achieves such a welfare gain, which is consistent with the previous studies on insurance nature of public pension such as Krueger and Kubler (2006) and Gottardi and Kubler (2011). On the other hand, for some individuals (low net gainers) marginal utility of consumption per period after introducing public pension, $u'(c_1^{pp}(\theta))$, can be lower than $\beta(1+r)\sum_{s\in S}u'(c_2^{lf}(\theta;s))\mu(s)$ $-u'(c_1^{lf}(\theta))$ that is obtained at the expense of public funds, while $u'(c_1^{pp}(\theta))$ can be higher than $\beta(1+r)\sum_{s\in S}u'(c_2^{lf}(\theta;s))\mu(s)-u'(c_1^{lf}(\theta))$ for others (high net gainers). In this light, (17) and (19) suggest that the welfare gain from providing public pension insurance is redistributed to the low net gainers. With the current level of generality, it is not feasible to clearly establish whether (19) increases in the earning ability. In this regard, with the earning history data of the Social Security Administration of the US, Gustman and Steinmeier (2001) found that the Social Security redistributes from individuals with high life-time income to those with low life-time income. Such redistribution would not be achieved if pension insurance were provided by financial market. Notably, as a benevolent social planner, the government pursues and attains social welfare improvement by redistribution of the welfare gain from public pension insurance. Lastly, (20) describes the marginal utility changes from the forgone resources due to the labor supply distortion and from the resulting reduction in disutility of working.

While the socially optimal allocation $\{c_1^{pp}(\theta), l^{pp}(\theta), \{c_2^{pp}(\theta;s)\}_{s\in S}\}_{\theta\in\Theta}$ uniquely defines the optimal labor supply distortion, the right-hand side of (15), it does not uniquely define optimal public pension contribution or benefit rate of the left-hand side of (15). For any given $\theta\in[\underline{\theta},\overline{\theta}]$, there can exist an infinite number of pairs of public pension contribution and benefit rates which meet the optimal public pension rate formula (15) with its right-hand side being uniquely defined. As long as the difference between marginal public pension contribution rate and the present value of the expected marginal public pension benefit rate is equal to the optimal labor supply distortion for $\forall\theta\in[\underline{\theta},\overline{\theta}]$, all of such pairs of public pension contribution and benefit rates can implement the same socially optimal allocation. As noted by Golosov, Tsyvinski and Werning (2006), generally, tax systems that implement an optimal allocation may not be uniquely determined. That is, there are various rate schedules that government can choose to achieve an optimal level of social welfare with socially optimal allocation, which gives government flexibility in designing public pension rates but clearly throws challenges for identifying optimal rates of public pension contribution and benefit. Nonetheless, we still can identify the following unique property of optimal public pension rates of this economy without income taxation.

Proposition 3. Without income taxation, optimal marginal public pension rates are not actuarially fair except for individuals of the highest earning ability $\overline{\theta}$.

Proof. See Appendix D.

In terms of (15), **Proposition 3** means that the optimal labor supply distortion is zero only for individuals of the highest earning ability and is strictly positive for anyone else. Thus, **Proposition 3** shows no distortion at the top. Consequently, except for individuals of the highest earning ability $\overline{\theta}$, the expected present value of marginal public pension benefit rate is not equal to but lower than marginal public pension contribution rate. In this regard, Diamond and Mirrlees (1978; 1986) found that optimal public pension is not actuarially fair; however, their definition of actuarial fairness is quite different from general one like ours. Diamond and Mirrlees (1978; 1986) defined that pension insurance is actuarially fair if the present value of net transfer is not affected by retirement date in their model. On the other hand, it should be noted that **Proposition 3** is concerned with actuarial fairness/unfairness at the *margin*. Therefore, **Proposition 3** does not mean that the total amounts of public pension contribution and benefit are actuarially fair for individuals of the highest earning ability.

As well known, no distortion at the top in optimal income taxation of a Mirrleesian economy without public pension means *zero* marginal income tax rate on the top earners, unlike our result of actuarial fairness of marginal public pension rate on the top earners. The zero marginal income tax rate on the top earners is regarded as having 'limited practical relevance' (Mankiw, Weinzierl and Yagan, 2009). In contrast, according to **Proposition 3**, notice that both optimal marginal public pension contribution rate and optimal marginal public pension benefit rate can be strictly positive for the top earners to be actuarially fair (i.e., $S'(\bar{\theta}) = \frac{\sum_{s \in S} B'(\bar{\theta}; s)\mu(s)}{1+r}$), although both rates can be zero as well. Notably, in many countries like the US, Austria, and Germany, the government sets the maximum amount[7] of labor earnings above which public pension contribution is no longer counted. Above the ceiling of the maximum, an increase in the labor earnings no longer increases contribution for public pension to entail an increase in public pension benefit paid later. This ceiling on public pension insurance effectively means zero marginal rates of public pension contribution and benefit, which is optimal if the ceiling is equal to the highest earnings, based on **Proposition 3**.

With the current level of generality, the property of optimal public pension rates for individuals whose earning abilities are not the highest is not able to be characterized as concretely as that for individuals of the highest earning ability (**Proposition 3**). Nonetheless, **Proposition 1** implies that if the optimal labor supply distortion, the right-hand side of (15), varies across different earning abilities instead of staying constant, then both optimal public

[7] Taking example of the US, in 2018, the maximum amount of taxable earnings (before full retirement age) subject to the Social Security tax is $128,400. The taxable income above this ceiling made in 2018 is not included for the total amount of Social Security contributions which is the basis for calculating Primary Insurance Amount later to determine Social Security benefit. The maximum is adjusted according to consumer price index. For details, https://www.ssa.gov/planners/maxtax.html.

pension contribution rate and optimal public pension benefit rate are not flat and at least one of the two rates should be nonlinear. In this regard, with the Social Security Administration panel data (over 1978-2011) that provides approximated distribution of earning ability, Golosov, Troshkin and Tsyvinski (2016) found that the optimal labor supply distortion estimated over the different earning levels (which approximate different earning abilities) is U-shaped, resonating with other previous studies like Diamond (1998) and Saez (2001). Since our model and the model of Golosov, Troshkin and Tsyvinski (2016) are alike in that both are a life-cycle Mirrleesian model with a given retirement date and both equate optimal marginal rate with optimal labor supply distortion, although the risk structure and total number of periods are different from each other. Thus, their empirical finding on the distribution of the optimal labor supply distortion across different earning abilities can be applied to this study. This implies that the optimal labor supply distortion of (15) does vary across different earning abilities and increases at least over a subset of the support $[\underline{\theta},\bar{\theta}]$. Let Λ denote for such a subset of the support over which the optimal labor supply distortion, $1-\frac{v'(\frac{y^{pp}(\theta)}{\theta})\frac{1}{\theta}}{u'(c_1^{pp}(\theta))}$, increases in θ for $\forall\theta\in\Lambda$. As mentioned above, optimal rates of public pension contribution and benefit cannot be uniquely identified without further assumptions or restrictions. Nevertheless, if the government chooses progressive public pension benefit rate (i.e., higher marginal public pension benefit rate for lower earning ability individuals for each state) over the subset of Λ, we can identify whether optimal public pension contribution rate is progressive (i.e., higher marginal public pension contribution rate for higher earning ability individuals) or not.

Proposition 4. Over the Λ, whenever optimal public pension benefit rate is progressive, optimal public pension contribution rate is also progressive.

Proof. See Appendix E.

Owing to the empirical finding of Golosov, Troshkin and Tsyvinski (2016), the condition of **Proposition 4** (i.e., existence of the Λ) is realistic rather than implausible. The social welfare gain to the redistributive government from providing public pension insurance may increase at the cost of an increase in the labor supply distortion. Intuitively, the more social welfare gain from public pension insurance the government obtains, the government can afford higher efficiency cost of labor supply distortion that public pension insurance brings, which gives the government more room for redistribution. Thus, as the optimal labor supply distortion increases, the government can choose progressive contribution and benefit rates of public pension by designating more generous (less generous) rates of public pension contribution and benefit for individuals of lower (higher) earning ability. Admittedly, to implement the socially optimal allocation, it is not necessary that the government always chooses progressive public pension benefit rate when the optimal labor supply distortion increases. However, **Proposition 4** shows that once progressive public pension benefit rate is chosen over the Λ, optimal public pension contribution rate is progressive if income tax is not levied. Moreover, according to **Proposition 1**, *net* public pension contribution rate should always be progressive at the optimum over the Λ, even when public pension benefit rate is not progressive.

With our focus on optimal public pension, we so far assume that the government only runs public pension program. The next section will extend the current analysis to the case where the government runs public pension program and imposes labor income tax at the same time.

6. Economy with Public Pension and Income Tax

In addition to public pension, the government imposes tax on the labor earnings in this Mirrleesian economy to finance a given public expenditure whose total amount is E at the present value in the first period. Let $T(y(\theta))$ denote for the amount of income tax from an ability-θ individual. Since labor is supplied only in the first period, $T(y(\theta))$ is collected only in the first period. Assume that $T(y(\theta))$ is continuously differentiable for $\forall\theta\in[\underline{\theta},\bar{\theta}]$. $T'(y)$ lies between 0 and 1 and may vary across different earning abilities.

Given the public pension and income tax $\{S(y(\theta)),\{B(y(\theta);s)\}_{s\in S},T(y(\theta))\}_{\theta\in\Theta}$, for any given $\theta\in[\underline{\theta},\bar{\theta}]$, an ability-$\theta$ individual maximizes his life-time utility of (9) subject to the following budget constraints

$$c_1\le y-T(y)-S(y)-k \ \text{ and } \ c_2\le(1+r)k(s)+B(y). \tag{21}$$

The optimality condition for his labor supply is

$$u'(c_1(\theta))(1-T'-S')+\beta\sum_{s\in S}u'(c_2(\theta;s))B'\mu(s)=v'(\frac{y(\theta)}{\theta})\frac{1}{\theta}. \tag{22}$$

Similar to (11), with public pension and income tax, labor is supplied less than the first-best level and the labor supply distortion is inter-temporal. Above all, the social planner government obtains optimal rate schedule of public pension and income tax by maximizing the social welfare function of (13) with meeting the incentive constraint of (12) and the following aggregate resource constraint

$$\int_{\underline{\theta}}^{\bar{\theta}}c_1(\theta)+\frac{1}{1+r}\sum_{s\in S}c_2(\theta;s)\mu(s)dF(\theta)+E\le\int_{\underline{\theta}}^{\bar{\theta}}y(\theta)dF(\theta). \tag{23}$$

Except for income tax and public expenditure *E*, the government solves the same problem that it solves in Section 5. The optimal allocation of this economy with public pension and labor income tax is from solving this constrained social welfare maximization problem and is notated as $\{c_1^{pl}(\theta),l^{pl}(\theta),\{c_2^{pl}(\theta;s)\}_{s\in S}\}_{\theta\in\Theta}$.

Proposition 5. Optimal rates of public pension and income tax are defined as follows: for any given $\theta\in[\underline{\theta},\bar{\theta}]$,

$$T'(\theta)+S'(\theta)-\frac{\sum_{s\in S}B'(\theta;s)\mu(s)}{1+r}=1-\frac{\mathrm{g}(\theta)+\phi(\hat{\theta},\theta)-\int_{\theta}^{\bar{\theta}}\phi(\theta,\hat{\theta})\frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}}{\mathrm{g}(\theta)+\phi(\hat{\theta},\theta)-\int_{\theta}^{\bar{\theta}}\phi(\theta,\hat{\theta})\frac{\theta}{\hat{\theta}}\frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}}=1-\frac{v'(\frac{y^{pl}(\theta)}{\theta})\frac{1}{\theta}}{u'(c_1^{pl}(\theta))}. \tag{24}$$

Under optimal public pension and income tax rate schedule, all individuals smooth consumption at ex ante point of view. That is, for any given $\theta\in[\underline{\theta},\bar{\theta}]$,

$$u'(c_1^{pl}(\theta))=\beta(1+r)\sum_{s\in S}u'(c_2^{pl}(\theta;s))\mu(s). \tag{25}$$

Proof. See Appendix F.

The benefit of public pension insurance that enables individuals to smooth consumption is not altered by introducing income tax. Moreover, the middle part and the right-hand side of (24) take the same form as the corresponding parts of (15). This implies that as in **Proposition 2**, the optimal rates of public pension and income tax of (24) also can be decomposed into (A) welfare gain from public pension insurance, (B) redistribution of the welfare gain, and (C) efficiency loss from labor supply distortion. Moreover, like the right-hand side of (15), the optimal labor supply distortion of (24) is intra-temporal and uniquely defined for any given $\theta\in[\underline{\theta},\bar{\theta}]$, whereas optimal rates of public pension contribution and benefit and optimal income tax rate are not uniquely determined. Notably, public pension contribution and income tax alike are on the same labor earnings and thus affect the same labor supply margin. For any given $\theta\in[\underline{\theta},\bar{\theta}]$, there exist more than one trio of public pension contribution and benefit rates and income tax rate which satisfies (24).

With a larger number of policy instruments available, the government has more options to implement the socially optimal allocation. Consequently, now allowing income tax rate to take different values across different earning abilities, **Proposition 3** of actuarially fair marginal public pension rate only for individuals of the highest earning ability $\bar{\theta}$ can be extended for the other individuals of lower earning abilities.

Proposition 6. With income taxation, optimal marginal public pension rates can be actuarially fair for all individuals regardless of earning ability.

Proof. See Appendix G.

By adjusting marginal income tax rate to implement the socially optimal allocation $\{c_1^{pl}(\theta), l^{pl}(\theta), \{c_2^{pl}(\theta;s)\}_{s\in S}\}_{\theta\in\Theta}$, optimal marginal public pension contribution rate can always be equated with the expected present value of optimal marginal public pension benefit rate, even when the optimal labor supply distortion, the right-hand side of (24), is strictly positive and varies by earning ability. Among all the implementable optimal public pension rates that meet (24) of **Proposition 5**, if the government chooses marginal public pension rates to be actuarially fair for all (i.e., $S'(\theta) = \frac{\sum_{s\in S} B'(\theta;s)\mu(s)}{1+r}$ for $\forall\theta\in[\underline{\theta},\bar{\theta}]$), then optimal marginal income tax rate is zero for the top earners because of no labor supply distortion at the top. On the other hand, however, if the government does not, then (24) of **Proposition 5** implies that optimal marginal income tax rate on individuals of the highest earning ability can be strictly positive with net optimal marginal public pension benefit rate being strictly positive, even if the optimal labor supply distortion is still zero for those of the highest earning ability.

Moreover, notice that by replacing marginal public pension contribution rate of (15) with the sum of marginal income tax rate and marginal public pension contribution rate of (24), both (15) and (24) can become identical. In this light, with income taxation available, **Proposition 4** entails the corollary that the sum of optimal income tax and public pension contribution rates is progressive whenever optimal public pension benefit rate is progressive over the range where the optimal labor supply distortion $1-\frac{v'(\frac{y^{pl}(\theta)}{\theta})\frac{1}{\theta}}{u'(c_1^{pl}(\theta))}$ is increasing. In fact, under the Social Security public pension program of the US, public pension contribution rate is flat (non-progressive) while both public pension benefit rate and labor income tax rate are progressive. The corollary of **Proposition 4** for economy with public pension and income taxation may suggest that non-progressive flat rate of the Social Security public pension contribution paired with progressive income tax rate might not necessarily be sub-optimal even when the Social Security public pension benefit rate is progressive.

In a nutshell, introducing income taxation does not fundamentally alter the main findings of optimal public pension rates without income taxation (Section 5) while it makes the feasible implementations of optimal public pension rates be less restrictive but continues to gain social welfare improvement from providing public pension insurance.

7. **Concluding Remarks**

This paper studies optimal nonlinear rates of public pension contribution and benefit without and with income taxation in a life-cycle Mirrleesian model where private pension market is not complete. In particular, the risk of outliving retirement savings (the longevity risk on retirement savings) — the defining feature of public pension insurance — is incorporated into the model. From constrained social welfare maximization under this model, optimal public pension contribution and benefit rates are derived without and with income taxation. The optimal public pension has the components of the redistribution of welfare gain from providing public pension insurance as well as the welfare gain. With no income taxation, optimal marginal public pension rates are actuarially fair only for individuals of the highest earning ability, although optimal marginal public pension contribution and benefit rates can be strictly positive for individuals of the highest earning ability. On the other hand, with income tax being levied as well, optimal marginal public pension rates can be actuarially fair for all individuals irrespective of earning ability.

Furthermore, without income taxation, progressive public pension benefit rate entails progressive public pension contribution rate, under a realistic condition on the distribution of labor supply distortion across different earning abilities. Similarly, with income taxation, the sum of income tax and public pension contribution rates should be progressive whenever public pension benefit rate is progressive under the same condition.

Appendix

A. Proof for **Lemma 1**

Consider a strictly convex function $\psi(x)=\frac{1}{x}$. Then, due to Jensen's inequality, for any given $\theta\in[\underline{\theta},\overline{\theta}]$,

$$\psi\Big(\sum_{s\in S}\frac{1}{\beta(1+r)u'(c_2^{lf}(\theta;s))}\mu(s)\Big)<\sum_{s\in S}\psi\Big(\frac{1}{\beta(1+r)u'(c_2^{lf}(\theta;s))}\Big)\mu(s). \tag{A1}$$

Since $\mu(s)\in(0,1)$ for $\forall s\in S$, the equality between the right-hand and left-hand sides of (A1) does not hold. Moreover, in light of (8) and by the definition of ψ, (A1) is restated as

$$u'(c_1^{lf}(\theta))<\beta(1+r)\sum_{s\in S}u'(c_2^{lf}(\theta;s))\mu(s) \tag{A2}$$

which means that the standard Euler equation of $u'(c_1^{lf}(\theta))=\beta(1+r)E[u'(c_2^{lf}(\theta;s))]=\beta(1+r)\sum_{s\in S}u'(c_2^{lf}(\theta;s))\mu(s)$ cannot be met for any individuals, entailing failure in their consumption smoothing across the two periods at the first-period point of view. As a result, for any given θ, the marginal utility gain from the consumption smoothing for an ability-θ individual, which cannot be obtained at the optimum of laissez faire economy, is the difference between the right-hand and left-hand sides of the standard Euler equation $\beta(1+r)\sum_{s\in S}u'(c_2^{lf}(\theta;s))\mu(s)-u'(c_1^{lf}(\theta))>0$. ■

B. Proof for **Proposition 1**

At the outset, the Lagrangian of the maximization of the social planner government is

$$L^{pp}=\int_{\underline{\theta}}^{\overline{\theta}}\{u(c_1(\theta))-v(\frac{y(\theta)}{\theta})+\beta\sum_{s\in S}u(c_2(\theta;s))\mu(s)\}g(\theta)f(\theta)d\theta+\lambda\int_{\underline{\theta}}^{\overline{\theta}}\{y(\theta)-c_1(\theta)-\frac{1}{1+r}\sum_{s\in S}c_2(\theta;s)\mu(s)\}f(\theta)d\theta+\int_{\underline{\theta}}^{\overline{\theta}}\phi(\hat{\theta},\theta)\{[u(c_1(\theta))-v(\frac{y(\theta)}{\theta})+\beta\sum_{s\in S}u(c_2(\theta;s))\mu(s)]-[u(c_1(\hat{\theta}))-v(\frac{y(\hat{\theta})}{\theta})+\beta\sum_{s\in S}u(c_2(\hat{\theta};s))\mu(s)]\}f(\theta)d\theta. \tag{A3}$$

By the revelation principle, the government can focus on direct mechanisms where, for any given θ, public pension contribution and benefit rates that are intended for ability-θ individuals are correctly applied to the ability-θ individuals (not to others) as if the government can observe the earning ability, since individuals do not pretend to have lower earning ability (by reducing their labor supply) with the incentive constraint (12) being met. Moreover, because of the single crossing property (2), for any given θ, only those whose earning ability is higher than θ have incentive to pretend to be an ability-θ individual. That is, those whose earning ability is lower than θ have no incentive to pretend to be an ability-θ individual by over-working. Therefore, the first-order conditions for optimal public pension rates are as follows: for any given $\theta \in [\underline{\theta}, \bar{\theta}]$ and any given $s \in S$,

$$\frac{dL^{pp}}{dc_1} = \{u'(c_1^{pp}(\theta))\,\mathrm{g}(\theta) - \lambda + u'(c_1^{pp}(\theta))[\phi(\hat{\theta},\theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta,\hat{\theta}) \frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}]\} f(\theta) = 0 \qquad \text{(A4)}$$

$$\frac{dL^{pp}}{dc_2} = \{\beta u'(c_2^{pp}(\theta;s))\,\mathrm{g}(\theta) - \frac{\lambda}{1+r} + \beta u'(c_2^{pp}(\theta;s))[\phi(\hat{\theta},\theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta,\hat{\theta}) \frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}]\} f(\theta) = 0 \quad \text{(A5)}$$

$$\frac{dL^{pp}}{dy} = \{-v'(\frac{y^{pp}(\theta)}{\theta})\frac{1}{\theta}\mathrm{g}(\theta) + \lambda - v'(\frac{y^{pp}(\theta)}{\theta})\frac{1}{\theta}[\phi(\hat{\theta},\theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta,\hat{\theta}) \frac{\theta}{\hat{\theta}} \frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}]\} f(\theta) = 0 \,. \quad \text{(A6)}$$

From (A4) and (A5),

$$u'(c_1^{pp}(\theta)) = \beta(1+r)u'(c_2^{pp}(\theta;s)) \,. \qquad \text{(A7)}$$

Then, (11) and (A7) imply that

$$(1 - S'(\theta)) + \frac{\sum_{s\in S} B'(\theta;s)\mu(s)}{1+r} = \frac{1}{u'(c_1^{pp}(\theta))} v'(\frac{y^{pp}(\theta)}{\theta})\frac{1}{\theta} \,. \qquad \text{(A8)}$$

On the other hand, (A4) and (A6) imply that

$$\frac{1}{u'(c_1^{pp}(\theta))} v'(\frac{y^{pp}(\theta)}{\theta})\frac{1}{\theta} = \frac{[\mathrm{g}(\theta) + \phi(\hat{\theta},\theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta,\hat{\theta}) \frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}]}{[\mathrm{g}(\theta) + \phi(\hat{\theta},\theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta,\hat{\theta}) \frac{\theta}{\hat{\theta}} \frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}]} > 0 \qquad \text{(A9)}$$

since $u' > 0$, $v' > 0$, and $\underline{\theta} > 0$. Therefore, from (A8) and (A9),

$$S'(\theta) - \frac{\sum_{s\in S} B'(\theta;s)\mu(s)}{1+r} = 1 - \frac{\mathrm{g}(\theta) + \phi(\hat{\theta},\theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta,\hat{\theta}) \frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}}{\mathrm{g}(\theta) + \phi(\hat{\theta},\theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta,\hat{\theta}) \frac{\theta}{\hat{\theta}} \frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}} = 1 - \frac{v'(\frac{y^{pp}(\theta)}{\theta})\frac{1}{\theta}}{u'(c_1^{pp}(\theta))} \,.$$

Moreover, (A7) holds for an arbitrarily given state $s \in S$ and $\theta \in [\underline{\theta}, \bar{\theta}]$, which implies that, for any given $\theta \in [\underline{\theta}, \bar{\theta}]$, the present value of the marginal utility of the second-period consumption of an ability-θ individual is

$$\beta \sum_{s \in S} u'(c_2^{pp}(\theta; s))\mu(s) = \beta \sum_{s \in S} \frac{u'(c_1^{pp}(\theta))}{\beta(1+r)} \mu(s) = \frac{u'(c_1^{pp}(\theta))}{(1+r)} \tag{A10}$$

since $\sum_{s \in S} \mu(s) = 1$. Since (A10) is valid for an arbitrarily given $\theta \in [\underline{\theta}, \bar{\theta}]$, all individuals smooth consumption at ex ante point of view by meeting $u'(c_1^{pp}(\theta)) = \beta(1+r)\sum_{s \in S} u'(c_2^{pp}(\theta; s))\mu(s)$. ■

C. Proof for **Proposition 2**

Rearranging the terms in (A4) yields

$$\lambda = u'(c_1^{pp}(\theta))[\mathrm{g}(\theta) + \phi(\hat{\theta}, \theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta, \hat{\theta}) \frac{f(\hat{\theta}) d\hat{\theta}}{f(\theta)}]. \tag{A11}$$

With (A11), the optimal public pension contribution and benefit rates of (15) for an arbitrarily given $\theta \in [\underline{\theta}, \bar{\theta}]$ are restated as

$$S'(\theta) - \frac{\sum_{s \in S} B'(\theta; s)\mu(s)}{1+r} = \frac{1}{\xi(\theta)} \{\mathrm{g}(\theta) + \phi(\hat{\theta}, \theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta, \hat{\theta}) \frac{\theta}{\hat{\theta}} \frac{f(\hat{\theta}) d\hat{\theta}}{f(\theta)} - \frac{\lambda}{u'(c_1^{pp}(\theta))}\} \tag{A12}$$

where $\xi(\theta) = \mathrm{g}(\theta) + \phi(\hat{\theta}, \theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta, \hat{\theta}) \frac{\theta}{\hat{\theta}} \frac{f(\hat{\theta}) d\hat{\theta}}{f(\theta)}$. On the other hand, (A6) implies that

$$\frac{\lambda}{v'(\frac{y^{pp}(\theta)}{\theta})\frac{1}{\theta}} - \mathrm{g}(\theta) = [\phi(\hat{\theta}, \theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta, \hat{\theta}) \frac{\theta}{\hat{\theta}} \frac{f(\hat{\theta}) d\hat{\theta}}{f(\theta)}]. \tag{A13}$$

Taking derivative of (A13) with respect to θ, we get

$$\frac{\lambda}{v'(\frac{y^{pp}(\theta)}{\theta})}(1 + \frac{1}{\varepsilon(\theta)}) - \mathrm{g}'(\theta) = \frac{d[\phi(\hat{\theta}, \theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta, \hat{\theta}) \frac{\theta}{\hat{\theta}} \frac{f(\hat{\theta}) d\hat{\theta}}{f(\theta)}]}{d\theta} \tag{A14}$$

where $\varepsilon(\theta) = v'(\frac{y^{pp}(\theta)}{\theta})[v''(\frac{y^{pp}(\theta)}{\theta})\frac{1}{\theta}]^{-1}$. Then, integrating (A14) back yields

$$-\int_{\theta}^{\bar{\theta}} \frac{\lambda}{v'(\frac{y^{pp}(\theta)}{\theta})}(1 + \frac{1}{\varepsilon(\theta)}) - \mathrm{g}'(\theta) d\theta = \phi(\hat{\theta}, \theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta, \hat{\theta}) \frac{\theta}{\hat{\theta}} \frac{f(\hat{\theta}) d\hat{\theta}}{f(\theta)}. \tag{A15}$$

Then, with (A11), the optimal public pension rate of (15) for an arbitrarily given $\theta \in [\underline{\theta}, \bar{\theta}]$

is restated as

$$S'(\theta)-\frac{\sum_{s\in S}B'(\theta;s)\mu(s)}{1+r}=\frac{1}{\xi(\theta)}\{\mathrm{g}(\theta)-\frac{\lambda}{u'(c_1^{pp}(\theta))}+\int_{\theta}^{\bar{\theta}}\mathrm{g}'(\theta)-\frac{\lambda}{v'(\frac{y^{pp}(\theta)}{\theta})}(1+\frac{1}{\varepsilon(\theta)})d\theta\}\ .\text{(A16)}$$

Moreover, due to **Lemma 1**, for any given $\theta\in[\underline{\theta},\bar{\theta}]$, the marginal utility gain from consumption smoothing which is now attained by public pension insurance, as shown in (16), is $\beta(1+r)\sum_{s\in S}u'(c_2^{lf}(\theta;s))\mu(s)-u'(c_1^{lf}(\theta))>0$. Thus, since $g(\theta)>0$ for $\forall\theta\in[\underline{\theta},\bar{\theta}]$

$$\frac{\beta(1+r)\sum_{s\in S}u'(c_2^{lf}(\theta;s))\mu(s)-u'(c_1^{lf}(\theta))}{-u''(c_1^{pp}(\theta))}(-\frac{u''(c_1^{pp}(\theta))}{u'(c_1^{pp}(\theta))})\,\mathrm{g}(\theta)>0\,. \tag{A17}$$

Adding (A17) to and then subtracting (A17) from $\{\mathrm{g}(\theta)-\frac{\lambda}{u'(c_1^{pp}(\theta))}+\int_{\theta}^{\bar{\theta}}\mathrm{g}'(\theta)-\frac{\lambda}{v'(\frac{y^{pp}(\theta)}{\theta})}$ $(1+\frac{1}{\varepsilon(\theta)})d\theta\}$ complete the proof. ■

D. Proof for **Proposition 3**

The optimal public pension contribution and benefit rate of (15) for an arbitrarily given $\theta\in[\underline{\theta},\bar{\theta}]$ is restated as

$$S'(\theta)-\frac{\sum_{s\in S}B'(\theta;s)\mu(s)}{1+r}=\frac{\int_{\theta}^{\bar{\theta}}\phi(\theta,\hat{\theta})[1-\frac{\theta}{\hat{\theta}}]\frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}}{\mathrm{g}(\theta)+\phi(\hat{\theta},\theta)-\int_{\theta}^{\bar{\theta}}\phi(\theta,\hat{\theta})\frac{\theta}{\hat{\theta}}\frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}}\,. \tag{A18}$$

Due to (A13), the denominator of the right-hand side of (A18) is strictly positive since λ (the marginal value of public funds) is strictly positive and $v'>0$, $\underline{\theta}>0$. That is,

$$\frac{\lambda}{v'(\frac{y^{pp}(\theta)}{\theta})\frac{1}{\theta}}=[\mathrm{g}(\theta)+\phi(\hat{\theta},\theta)-\int_{\theta}^{\bar{\theta}}\phi(\theta,\hat{\theta})\frac{\theta}{\hat{\theta}}\frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}]>0\,. \tag{A19}$$

Therefore, the sign of the left-hand side of (A18) is determined by the sign of the numerator of its right-hand side. The numerator of the right-hand side of (A18) is zero only when $\theta=\bar{\theta}$ whereas it is strictly positive for $\forall\theta\in[\underline{\theta},\bar{\theta})$ since, with an arbitrarily given $\theta\in[\underline{\theta},\bar{\theta})$, $\phi(\theta,\hat{\theta})>0$ and $1-\frac{\theta}{\hat{\theta}}>0$ for $\forall\hat{\theta}\in(\theta,\bar{\theta})$. This implies that, under the optimal public pension rate schedule of (15) in **Proposition 1**, $S'(\theta)-\frac{\sum_{s\in S}B'(\theta;s)\mu(s)}{1+r}=0$ (i.e., actuarially fair) only when $\theta=\bar{\theta}$ whereas $S'(\theta)-\frac{\sum_{s\in S}B'(\theta;s)\mu(s)}{1+r}>0$ (i.e., not actuarially fair) for $\forall\theta\in[\underline{\theta},\bar{\theta})$. ■

E. Proof for **Proposition 4**

By definition, for $\forall\theta\in\Lambda\subset[\underline{\theta},\bar{\theta}]$, the optimal labor supply distortion $1-\frac{v'(\frac{y^{pp}(\theta)}{\theta})\frac{1}{\theta}}{u'(c_1^{pp}(\theta))}$ is increasing in θ. Now, suppose that optimal public pension benefit rate is progressive over the Λ. Then, for θ^h and θ^l arbitrarily picked from the Λ with $\theta^h>\theta^l$,

$$B'(\theta^l)>B'(\theta^h) \text{ for } \forall s\in S \text{ and } \frac{v'(\frac{y^{pp}(\theta^l)}{\theta^l})\frac{1}{\theta^l}}{u'(c_1^{pp}(\theta^l))}-1>\frac{v'(\frac{y^{pp}(\theta^h)}{\theta^h})\frac{1}{\theta^h}}{u'(c_1^{pp}(\theta^h))}-1. \tag{A20}$$

Notice that the optimal public pension contribution and benefit rate of (15) can be restated as $\frac{\sum_{s\in S}B'(\theta;s)\mu(s)}{1+r}-S'(\theta)=\frac{v'(\frac{y^{pp}(\theta)}{\theta})\frac{1}{\theta}}{u'(c_1^{pp}(\theta))}-1$, which implies from (A20) that

$$S'(\theta^h)+\frac{\sum_{s\in S}B'(\theta^l;s)\mu(s)}{1+r}-\frac{\sum_{s\in S}B'(\theta^h;s)\mu(s)}{1+r}>S'(\theta^l). \tag{A21}$$

Since $B'(\theta^l)>B'(\theta^h)$ for $\forall s\in S$, (A21) implies that $S'(\theta^h)>S'(\theta^l)$. This shows that optimal public pension contribution rate is also progressive over the Λ. ■

F. Proof for **Proposition 5**

At the outset, the Lagrangian of the social welfare maximization of the government with public pension and income taxation is

$$L^{pl}=\int_{\underline{\theta}}^{\bar{\theta}}\{u(c_1(\theta))-v(\frac{y(\theta)}{\theta})+\beta\sum_{s\in S}u(c_2(\theta;s))\mu(s)\}g(\theta)f(\theta)d\theta+\lambda\int_{\underline{\theta}}^{\bar{\theta}}\{y(\theta)-c_1(\theta)-\frac{1}{1+r}\sum_{s\in S}$$
$$c_2(\theta;s)\mu(s)-E\}f(\theta)d\theta+\int_{\underline{\theta}}^{\bar{\theta}}\phi(\hat{\theta},\theta)\{[u(c_1(\theta))-v(\frac{y(\theta)}{\theta})+\beta\sum_{s\in S}u(c_2(\theta;s))\mu(s)]-[u(c_1(\hat{\theta}))-$$
$$v(\frac{y(\hat{\theta})}{\theta})+\beta\sum_{s\in S}u(c_2(\hat{\theta};s))\mu(s)]\}f(\theta)d\theta. \tag{A22}$$

Due to the revelation principle, the first-order conditions for optimal rate schedule of public pension and income tax are as follows: for any given $\theta \in [\underline{\theta}, \bar{\theta}]$ and any given $s \in S$,

$$\frac{dL^{pl}}{dc_1} = \{u'(c_1^{pl}(\theta))\,\mathrm{g}(\theta) - \lambda + u'(c_1^{pl}(\theta))[\phi(\hat{\theta}, \theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta, \hat{\theta}) \frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}]\} f(\theta) = 0 \qquad \text{(A23)}$$

$$\frac{dL^{pl}}{dc_2} = \{\beta u'(c_2^{pl}(\theta; s))\,\mathrm{g}(\theta) - \frac{\lambda}{1+r} + \beta u'(c_2^{pl}(\theta; s))[\phi(\hat{\theta}, \theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta, \hat{\theta}) \frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}]\} f(\theta) = 0 \quad \text{(A24)}$$

$$\frac{dL^{pl}}{dy} = \{-v'(\frac{y^{pl}(\theta)}{\theta})\frac{1}{\theta}\mathrm{g}(\theta) + \lambda - v'(\frac{y^{pl}(\theta)}{\theta})\frac{1}{\theta}[\phi(\hat{\theta}, \theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta, \hat{\theta}) \frac{\theta}{\hat{\theta}} \frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}]\} f(\theta) = 0\,. \quad \text{(A25)}$$

From (A23) and (A24),

$$u'(c_1^{pl}(\theta)) = \beta(1+r)u'(c_2^{pl}(\theta; s))\,. \qquad \text{(A26)}$$

Then, (22) and (A26) imply that

$$(1 - T'(\theta) - S'(\theta)) + \frac{\sum_{s\in S} B'(\theta; s)\mu(s)}{1+r} = \frac{1}{u'(c_1^{pl}(\theta))} v'(\frac{y^{pl}(\theta)}{\theta})\frac{1}{\theta}\,. \qquad \text{(A27)}$$

On the other hand, (A23) and (A25) imply that

$$\frac{1}{u'(c_1^{pl}(\theta))} v'(\frac{y^{pl}(\theta)}{\theta})\frac{1}{\theta} = \frac{[\mathrm{g}(\theta) + \phi(\hat{\theta}, \theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta, \hat{\theta}) \frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}]}{[\mathrm{g}(\theta) + \phi(\hat{\theta}, \theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta, \hat{\theta}) \frac{\theta}{\hat{\theta}} \frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}]} > 0 \qquad \text{(A28)}$$

since $u' > 0$, $v' > 0$, and $\underline{\theta} > 0$. Therefore, from (A27) and (A28),

$$T'(\theta) + S'(\theta) - \frac{\sum_{s\in S} B'(\theta; s)\mu(s)}{1+r} = 1 - \frac{\mathrm{g}(\theta) + \phi(\hat{\theta}, \theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta, \hat{\theta}) \frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}}{\mathrm{g}(\theta) + \phi(\hat{\theta}, \theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta, \hat{\theta}) \frac{\theta}{\hat{\theta}} \frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}} = 1 - \frac{v'(\frac{y^{pl}(\theta)}{\theta})\frac{1}{\theta}}{u'(c_1^{pl}(\theta))}\,.$$

Moreover, (A26) holds for an arbitrarily given state $s \in S$ and $\theta \in [\underline{\theta}, \bar{\theta}]$, which implies that

$$\beta(1+r)\sum_{s\in S} u'(c_2^{pl}(\theta; s))\mu(s) = \beta(1+r)\sum_{s\in S} \frac{u'(c_1^{pl}(\theta))}{\beta(1+r)}\mu(s) = u'(c_1^{pl}(\theta)) \qquad \text{(A29)}$$

for an arbitrarily given $\theta \in [\underline{\theta}, \bar{\theta}]$. Since (A29) is valid for an arbitrarily given $\theta \in [\underline{\theta}, \bar{\theta}]$, all individuals smooth consumption at ex ante point of view by meeting $u'(c_1^{pl}(\theta)) = \beta(1+r)$ $\sum_{s\in S} u'(c_2^{pl}(\theta; s))\mu(s)$. ■

G. Proof for **Proposition 6**

To begin, consider a public pension whose marginal rates are actuarially fair for all individuals regardless of earning ability. Such a public pension is defined by

$$S'(\theta) - \frac{\sum_{s\in S} B'(\theta;s)\mu(s)}{1+r} = 0 \quad \text{for} \quad \forall\theta \in [\underline{\theta},\bar{\theta}]. \tag{A30}$$

Then, we need to show that with (A30) being met there always exists a marginal income tax rate T' satisfying (24), which is the formula of optimal rates of public pension and income tax, for any given $\theta \in [\underline{\theta},\bar{\theta}]$. To this end, for an arbitrarily given $\theta \in [\underline{\theta},\bar{\theta}]$, we always can find such an optimal marginal income tax rate by plugging (A30) into (24) as follows:

$$T'(\theta) = 1 - \frac{\mathrm{g}(\theta) + \phi(\hat{\theta},\theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta,\hat{\theta}) \frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}}{\mathrm{g}(\theta) + \phi(\hat{\theta},\theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta,\hat{\theta}) \frac{\theta}{\hat{\theta}} \frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}} = 1 - \frac{v'(\frac{y^{pl}(\theta)}{\theta})\frac{1}{\theta}}{u'(c_1^{pl}(\theta))}. \tag{A31}$$

Thus, marginal rates of public pension contribution and benefit of (A30), which are actuarially fair for all, and marginal income tax rate of (A31) are optimal, according to **Proposition 5**.

In addition, if T' of (A31) does not lie between 0 and 1, it is not a marginal *income tax* rate. Thus, we also need to show that T' of (A31) lies between 0 and 1. Firstly, to show that T' of (A31) is not negative, (A31) is restated as

$$T'(\theta) = \frac{\int_{\theta}^{\bar{\theta}} \phi(\theta,\hat{\theta})[1-\frac{\theta}{\hat{\theta}}] \frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}}{\mathrm{g}(\theta) + \phi(\hat{\theta},\theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta,\hat{\theta}) \frac{\theta}{\hat{\theta}} \frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}} = 1 - \frac{v'(\frac{y^{pl}(\theta)}{\theta})\frac{1}{\theta}}{u'(c_1^{pl}(\theta))} \tag{A32}$$

for an arbitrarily given $\theta \in [\underline{\theta},\bar{\theta}]$. Based on (A25), the denominator of the middle part of (A32) is restated as

$$\frac{\lambda}{v'(\frac{y^{pl}(\theta)}{\theta})\frac{1}{\theta}} = [\mathrm{g}(\theta) + \phi(\hat{\theta},\theta) - \int_{\theta}^{\bar{\theta}} \phi(\theta,\hat{\theta}) \frac{\theta}{\hat{\theta}} \frac{f(\hat{\theta})d\hat{\theta}}{f(\theta)}] \tag{A33}$$

which is strictly positive due to $\lambda > 0$, $v' > 0$, and $\underline{\theta} > 0$. And, the numerator of the middle part of (A32) is zero only when $\theta = \bar{\theta}$ and is strictly positive for $\forall\theta \in [\underline{\theta},\bar{\theta})$ since, with an arbitrarily given $\theta \in [\underline{\theta},\bar{\theta})$, $1-\frac{\theta}{\hat{\theta}} > 0$ for $\forall\hat{\theta} \in (\theta,\bar{\theta}]$.

Secondly, to show that T' of (A31) is not greater than 1 by way of contradiction, suppose not. Then, (A31) implies that

$$T'(\theta) = 1 - \frac{v'(\frac{y^{pl}(\theta)}{\theta})\frac{1}{\theta}}{u'(c_1^{pl}(\theta))} > 1 \tag{A34}$$

which entails that

$$0 > \frac{v'(\frac{y^{pl}(\theta)}{\theta})\frac{1}{\theta}}{u'(c_1^{pl}(\theta))}. \tag{A35}$$

This is contradictory since $u' > 0$, $v' > 0$, and $\underline{\theta} > 0$. Therefore, T' of (A31) lies between 0 and 1. Taking all together, with income taxation, by (A30) and (A31), optimal marginal public pension rates can be actuarially fair for all individuals regardless of earning ability. ■

References

Bastani, Spencer, Sören Blomquist and Luca Micheletto. (2013) “The Welfare Gains of Age-related Optimal Income Taxation.” *International Economic Review*, 54(4): 1219-1249.

Boadway, Robin, Manuel Leite-Monteiro, Maurice Marchand and Pierre Pestieau. (2006) “Social Insurance and Redistribution with Moral Hazard and Adverse Selection.” *Scandinavian Journal of Economics*, 108(2): 279-298.

Butrica, Barbara A., Howard M. Iams, Karen E. Smith and Eric J. Toder. (2009) “The Disappearing Defined Benefit Pension and Its Potential Impact on the Retirement Incomes of Baby Boomers.” *Social Security Bulletin*, 69(3): 1-27.

Cremer, Helmuth, Jean-Marie Lozachmeur and Pierre Pestieau. (2004) “Social Security, Retirement Age and Optimal Income Taxation.” *Journal of Public Economics*, 88(11): 2259-2281.

Diamond, Peter A. and James A. Mirrlees. (1978) “A Model of Social Insurance with Variable Retirement.” *Journal of Public Economics*, 10(3): 295-336.

Diamond, Peter A. and James A. Mirrlees. (1986) “Payroll-Tax Financed Social Insurance with Variable Retirement.” *Scandinavian Journal of Economics*, 88(1): 25-50.

Diamond, Peter A. (1998) “Optimal Income Taxation: An example with a U-shaped Pattern of Optimal Marginal Tax Rates.” *American Economic Review*, 88(1): 83-95.

Farhi, Emmanuel and Iván Werning. (2013) “Insurance and Taxation over the Life Cycle.” *Review of Economic Studies*, 80(2): 596-635.

Feldstein, Martin. (1985) “The Optimal Level of Social Security Benefits.” *Quarterly Journal of Economics*, 100(2): 303-320.

Golosov, Mikhail, Aleh Tsyvinski and Ivan Werning. (2006) “New Dynamic Public Finance: A User's Guide.” *NBER Macroeconomics Annual* 21: 317-388.

Golosov, Mikhail, Maxim Troshkin and Aleh Tsyvinski. (2016) “Redistribution and Social Insurance.” *American Economic Review*, 106.2: 359-86.

Gottardi, Piero and Felix Kubler. (2011) “Social Security and Risk Sharing.” *Journal of Economic Theory*, 3: 1078-1106.

Grochulski, Borys and Narayana Kocherlakota. (2010) “Nonseparable Preferences and Optimal Social Security Systems.” *Journal of Economic Theory*, 145(6): 2055-2077.

Gustman, Alan L. and Thomas L. Steinmeier. (2001) “How Effective Is Redistribution under The Social Security Benefit Formula?” *Journal of Public Economics*, , 82(1): 1-28.

İmrohoroglu, Ayşe, Selahattin İmrohoroglu and Douglas H. Joines. (1995) “A Life Cycle Analysis of Social Security.” *Economic Theory*, 6(1): 83-114.

Krueger, Dirk and Felix Kubler. (2006) “Pareto-Improving Social Security Reform When Financial Markets Are Incomplete!?” *American Economic Review*, 96(3): 737-755.

Mankiw, N. Gregory, Matthew Weinzierl and Danny Yagan. (2009) “Optimal Taxation in Theory and Practice.” *Journal of Economic Perspectives*, 23(4): 147-174.

Mirrlees, James A. (1971) “An Exploration in the Theory of Optimum Income Taxation.” *Review of Economic Studies*, 38(2): 175-208.

Weinzierl, Matthew. (2011) “The Surprising Power of Age-dependent Taxes.” *Review of Economic Studies*, 78(4): 1490-1518.

Whinston, Michael. (1983) “Moral Hazard, Adverse Selection, and the Optimal Provision of Social Insurance.” *Journal of Public Economics*, 22(1): 49-71.

Yew, Siew Ling and Jie Zhang. (2009) “Optimal Social Security in a Dynastic Model with Human Capital Externalities, Fertility and Endogenous Growth.” *Journal of Public Economics*, 9: 605-619.