Social insurance and contribution density[1]

by

Salvador Valdés-Prieto[2]

Draft, September 10, 2008

Abstract

This paper makes two contributions to the welfare basis for social insurance. On the one hand, it offers a formal model where middle class individuals suffer an optimistic bias (unrealistic optimism: Weinstein, 1980). This endows social insurance with a benevolent role. On the other hand, the model allows individuals to choose a low "density of contribution," i.e. to devote a low proportion of active life to work in covered jobs. Thus, pension adequacy is not a direct choice variable for policymakers, and a mandate to contribute distorts the choice of density by the middle class. This implies that contributory pensions should be complemented with a subsidy to earnings in covered jobs. In a second stage, income inequality is added, and redistributive concerns and non-contributory subsidies for the old poor ensue. In this setting, the model finds that non-contributory pensions induce reductions in density of contribution and adequacy. The size of this reduction depends on the design of the non-contributory pension. Thus, non-contributory pensions for the old poor crowd out social insurance. The model integrates these issues and allows optimization of policy parameters. It is found that the in the absence of unrealistic optimism, the presence of non-contributory pensions does not justify a mandate to save for old age when a targeted tax-expenditure system is available. However, the main findings await the completion of the simulation stage of this research.

JEL classification: H55, H53, H24
Keywords: Social Security, Pensions, Density of contribution

---

[1] This paper was prepared for the CeRP annual conference *Financial Security in Retirement*, September 18-19, 2008, Collegio Carlo Alberto, Moncalieri, Turin. I thank comments from Robert Dur, Christian Keuschnigg and other participants at the CESifo Area Conference on Employment and Social Protection, May 23-24, 2008, and also from Andrea Repetto, UAI; I also thank Gonzalo Valdés for able research assistance. A companion paper (Valdes-Prieto 2008b) describes and evaluates the recent pension reform in Chile, which modified non-contributory pensions for the old poor. Comments are welcome.

[2] Professor of Economics at the Pontificia Universidad Católica de Chile, svaldes@faceapuc.cl

1. Introduction

For hundreds of millions of middle-class people around the world, consumption when old depends on the contributory social insurance promoted by the State, through either mandates or fiscal incentives. However, if a large share of participants chooses a low density of contribution, the median replacement rate will be inadequate, even when the benefit formula is generous for a high-density participant.

Since pension benefit formulae are defined over the time series of the participant's past contributions, the most relevant aspect of coverage is density of contribution over time, not the cross-section. [3] Bucheli et al (2007) show how to estimate density from individual panel data using survival analysis.

Low density of contribution is prevalent in many countries, despite mandates. In the case of old-age pensions, important examples are China - only 48% of urban employees contributed in 2005 (Salditt et al 2007), Poland - 68% contribute through ZUS, South Korea - 58% of the labor force contributes (World Bank, 2000), Brazil –only 49% of the employed contribute, and Mexico - only 38% of the employed contribute (Rofman and Lucchetti, 2006). In Western European countries with generous unemployment benefits, many less qualified people spend a low proportion of their active life working in covered jobs. In those segments, density and the adequacy of contributory pensions are low. However, the literature does not offer models about density and adequacy.

Uneven density of contribution across social groups raises additional challenges. For example, rural workers and the urban self-employed (most of them poor) are effectively exempt from the mandate in most countries, in addition to women engaged in home production. Some middle-income people may fall into relative poverty in old age, due to inadequate contributory pensions relative to former earnings. This situation may undercut the public's support for contributory pensions. For example, in the debate that led to the 2008 reform in Chile, inadequate replacement rates were agreed to be caused by uneven density, but the blame was placed on the contributory system (Valdes-Prieto, 2008b).

This paper offers a formal model of contribution density and adequacy, and builds around it to explore the main policy questions faced in social insurance. In this model, workers choose whether to bundle job selection and saving for old age, by taking a covered job, or to unbundle them by taking uncovered jobs or activities and saving independently in the financial market. The model acknowledges large spreads between borrowing and lending in the financial market, so a mandate to save cannot be undone with cheap consumer credit.

One result is that for intermediate values of the earnings differential between covered and other jobs and the net tax implicit in other branches of social insurance, relative to the spread in the financial market, some individuals choose an interior density of contribution. For extreme values of the earnings differential, individuals choose polar densities.

[3] Density is different from the coverage of contributions, defined as the ratio of the number of contributions to employment at a given point in time. A given coverage rate has different implications depending on whether employment is segmented between individuals that hold covered jobs permanently and others that hold uncovered jobs permanently, or employment is homogenous and many individuals alternate between covered and uncovered jobs and activities.

An optimistic bias about earnings when old (unrealistic optimism: Weinstein 1980, Weinstein and Klein 1996, Puri and Robinson 2007) reduces desired saving when young. This provides a benevolent role for social insurance. Assuring ample supply of commitment devices is not enough to overcome this bias, as it is in the hyperbolic discounting model. With optimistic bias, the young take covered jobs because of their higher salary alone, and *object* to the fact that legislation forces them to save for old age. Once in old age, after individuals come to recognize their bias, they regret previous decisions, and when social insurance is present, they are grateful for it.

This paper uses the model of unrealistic optimism in which the bias is located in the budget constraint, presented by Valdés-Prieto (2002, p. 165-173). Preferences are standard.

The paper presents several models that identify the optimal contribution rate for Bismarckian social insurance, which are analogous to the standard optimal income tax problem. For analytical transparency, the paper assumes first that only the middle class is affected by social insurance, so income inequality and non-contributory pensions for the old poor are absent. Social welfare is obtained by aggregating lifetime utility levels assessed by individuals when old, which are free from bias. Individual heterogeneity remains regarding the optimistic bias and the earnings differential, but the State is forced to mandate a uniform contribution rate. Despite this, some social insurance is likely to be desirable.

Another finding is that mandatory social insurance should be complemented with a subsidy to earnings on covered jobs, in order to counteract the labor distortion created by the mandate.

Subsequently, income inequality and a non-contributory pension are introduced. The non-contributory pension has a design that encompasses several standard designs, such as a universal flat, a minimum pension subsidy, and the intermediate designs used by the Nordic countries and recently legislated in Chile.

Our conjecture is that the socially optimal subsidy withdrawal rate is small but positive (Valdés-Prieto 2002, p. 67-71). This is suggested by previous results in optimal income tax theory by Slemrod et al (1994) and Diamond (1998), who recommend that the net marginal tax rate at low income levels be larger than the net marginal tax rate for income levels near the median of the earnings distribution. Simulations by Poblete (2005) for a simpler model also found that the socially optimal withdrawal rates would be near 20%. This paper's model is able to evaluate this intuition through simulations.

An analytic argument proves that non-contributory pensions induce reductions in density of contribution, and that the size of this effect depends on the design of the non-contributory pension. Thus, non-contributory pensions crowd out contributory pensions.

The paper also evaluates the claim that in the absence of unrealistic optimism, social insurance can still be justified as a response to the distortions caused by non-contributory pensions. However, a targeted tax-expenditure is likely to be more efficient: a lower subsidy for low earners in the active phase can be coupled with a higher non-contributory subsidy in old age, to obtain equivalence to a targeted mandate to save for old age. This targeting allows the middle class to be exempted from the mandate, which would be more efficient because density is not distorted.

The approach in this paper allows optimization of the combined "multipillar" structure, where each participant can get both noncontributory and contributory pensions in old age.

Related literature includes the one on the marginal link between contributions and benefits in contributory old-age pensions, which affects the supply of hours to the labor market (Auerbach and Kotlikoff, 1987), but was silent on the choice between covered and uncovered jobs. Another literature argues that the design of mandatory old-age pensions affects choice between covered and uncovered jobs, but is silent on the effect of first-pillar design. Hubbard, Skinner and Zeldes (1995) argue that a high rate of withdrawal of subsidies to the old poor reduces the voluntary saving of the poor, but are silent on the impact on density of contribution to second pillars.[4] Valdés-Prieto (2002, p. 241-57) showed how the presence of uncovered jobs allows workers to cap the implicit tax imposed by a second pillar, but is silent on the impact on density. The literature on optimal income tax schedules stresses the disincentive effect on supply of hours of violent withdrawal of subsidies, identifies the optimal two-bracket schedule (Slemrod et al, 1994), and identifies more general optimal tax schedules (Diamond, 1998), but is silent on job choice.

Section 2 presents the assumptions of the model and its main parameters. Section 3 presents individual optimization when choosing density and independent saving simultaneously. Section 4 models unrealistic optimism as a bias that affects the young, and identifies benevolent social insurance. Section 5 introduces income inequality and non-contributory subsidies. The final remarks are in section 6.

2. Basis for a model of contributory old-age pensions

This section presents a model of labor supply where each individual chooses between jobs covered by State-supported contributory old-age pension plans, and other jobs (uncovered jobs). Coverage of contributions is endogenous.

2.1 Labor market and tax institutions

It is assumed that the State is, in many jobs, unable to enforce a mandate to contribute based on all labor productivity. This is obvious for home production and for individuals drawing on unemployment benefits, which may be generous and permanent. Within the wage economy, self-employment facilitates underreporting of work and earnings.[5] Informal firms, defined as those that evade taxes and regulations, also evade the mandate to

[4] Although the empirical literature on the impact of tax incentives on voluntary saving finds modest elasticities for the poor (Attanasio and Browning, 1996), violent withdrawal of subsidies create unusually large implicit tax rates, so the effect on voluntary saving can still be sizable.

[5] This happens for two reasons: (i) there is no employer interested in maximizing reported labor costs in order to minimize corporate income taxes; and (ii) there is no employer interested in minimizing the penalties applied by enforcers of employment protection regulations, which include compliance with social security and labor regulations.

contribute.[6] The assumption is that a mandate to participate in contributory old-age pensions cannot be enforced by the State in all jobs.

The State may also be *unwilling* to enforce a mandate to contribute on jobs taken primarily by the poor. In most emerging economies, this has originated legal exemptions to the mandate to contribute, mainly for the self-employed, and regulatory forbearance for noncompliance. In some high-income economies, low-wage jobs are formally exempt from the mandate to contribute (Netherlands, Australia).

The premise of this model is that uncovered jobs are a significant job option, not a marginal exception. At the same time, it is assumed that some jobs are covered, specifically those at large employers. Empirical work for Chile confirms that uncovered jobs are a realistic option for many workers (Torche and Wagner, 1997). In most emerging economies uncovered jobs are closer to 50%, and in most countries in Africa and South Asia they are closer to 80% (van Ginneken, 1999).

"Contribution density" is defined as the share of (the present value of) hours in the active phase of life on which the individual contributes to some contributory pension system for old-age. For any given rate of turnover between covered jobs and other uses of time, average density falls when self-employment and informal employment expands and when activity outside the labor force (mainly home production) rises. Density may also change for a different reason: underreporting of earnings, keeping the headcount constant.

This model collapses a continuum of dates and ages into two separate periods, the active phase and old age. The labor market and tax variables are:

$y^i \equiv w \cdot x_a^i \cdot (1 - \hat{l}_a)$, $i = c, ex$. Where,

w is the wage rate per hour per efficiency unit in the economy, x^i_a is the efficiency of this individual in sector i in the active phase, and $(1 - \hat{l}_a)\ \varepsilon\ (0,1)$ is the sum of hours supplied in the active phase to all jobs, which is exogenous.

y^c = gross earnings in the covered sector per period, in the active phase, assuming full density of work in covered jobs.

y^{ex} = earnings per period in jobs that are exempt or uncovered, in the active phase, assuming zero density in covered jobs.

z^{ex} = the productivity ratio. $z^{ex} = x^{ex}/x_a^c$. $1\text{-}z^{ex}$ is the earnings differential against exempt or uncovered jobs.

D = density, ε [0,1]. It is the proportion of hours in the active phase of life on which the individual contributes to some contributory old-age system.

$e_p \equiv w \cdot x_p \cdot (1 - \hat{l}_p)$, where

e_p is earnings in old age, x_p is the efficiency that the individual expects to have when old and $(1 - \hat{l}_p)\ \varepsilon\ (0,1)$ is the sum of hours supplied when old, which is exogenous.

ψ^a_p = the productivity ratio when old, as expected when active. $\psi^a_p = x_p / x_a^c$. Ageing is represented by the assumption $\psi^a_p < 1$.

[6] The informal employer loses by reporting the earnings of his employee to the State, because this betrays its presence.

t_a = net taxes on earnings, applicable only to covered jobs in the active phase. Old individuals are assumed to be exempt from paying analogous taxes.

It is assumed that the demand for labor in firms that offer both covered and uncovered jobs is infinitely elastic at the market wage rates. In an open economy this result obtains after capital adjusts through international capital flows.

The net tax rate t_a is constructed from the tax that finances other branches of social insurance. This tax is levied only on covered earnings, not on earnings from exempt jobs. Then, the marginal benefit must be subtracted. Since health insurance taxes tend to be highly redistributive, more contributions seldom attract more health services at the margin. When unemployment insurance benefits are redistributive, the marginal benefit is also smaller than the contribution rate.[7] The net tax rate t_a is thus substantial. This is independent from the level of the contribution rate for old age.

Accounting for administrative and marketing costs in both independent saving and social insurance would complicate the model. Although these costs are ignored for this reason, it is clear that small amounts of savings can be rendered inefficient by these costs.

A voluntary “third pillar” such as employer pensions, including defined benefit and 401(k) plans is ignored for simplicity. Section 5 assumes that the earnings from exempt jobs are effectively free from income tax.

2.2 The contributory old-age pension system for the middle classes

A contributory system for the middle classes is described by parameters such as the contribution rate, contribution base, minimum density to vest benefits, pension ages, the formula for the initial benefit (pension), the indexation rule for ongoing pensions and transfer values when exiting the plan at mid-career. With only two phases of life, many of these parameters do not apply.

The contribution base $CB(y^c)$ can be a nonlinear function of covered earnings y^c. The function CB can include a maximum for taxable earnings, a minimum taxable earnings and other aspects. In this section a “Bismarckian” design is assumed, i.e. that $CB(y^c) = y^c$.[8] This simplification cuts the number of parameters to two:

> θ = effective contribution rate for old age, deducted from gross earnings. This applies only in covered jobs and is uniform for all participants.[9]
>
> ρ^c = net internal rate of return (real terms) paid by the contributory system to each generation of participants. The replacement rate is $\theta\cdot(1+\rho^c)$. This return if defined as free from income taxes.[10]

---

[7] The same happens in Brazil and Mexico regarding mandatory contributions for housing.
[8] Its opposite is the “Beveridgian” design, where $CB(y^c) = A$ for any level of y^c.
[9] If there are separate contribution rates paid by workers and employers, θ is the equivalent rate that would apply if the total actual contribution were paid by the worker alone, equal to $(\theta_w + \theta_E)/(1+\theta_E)$. This formula assumes also that all contributions to health insurance and other branches of social insurance are paid by the worker, none by the employer. The general case is available in Valdés (2002) chapter 5.3.
[10] The return ρ^c is also net of the health insurance taxes levied on contributory pensions, but not levied on other income in passive life, such as earnings when old and the product of voluntary saving.

The value of ρ^c is heavily influenced by the financing method used by the contributory pension system and its degree of maturity. Full funding under efficient asset management and tax exemptions should bring ρ^c close to the pre-tax returns on financial saving. Mature pay-as-you-go finance under a rule where at least one parameter is adjusted to preserve financial independence from the fiscal budget, makes ρ^c equal to the growth rate of the real contribution revenue, which in a steady state is equal to the growth rate of real GDP. In dynamically efficient economies the growth rate of real GDP is smaller on average than the rate of return on physical capital (Demange, 2002). Of course, this difference between funding and pay-as-you-go originates in the need to pay the costs of the transfer given to the initial generations when pay-as-you-go finance was initially introduced, and cannot be eliminated without making some subset of transition generations pay this "legacy" cost (Sinn, 1998).

In any overlapping generation model with two-period lives, in a steady state:

(0) $$1+\rho^c = \left(\frac{\phi}{1+r} + \frac{1-\phi}{1+g_{CR}}\right)^{-1}; \qquad g_{CR} \equiv d\ln(\theta \cdot D \cdot y^c)/dt$$

where ϕ is the degree of funding of the plan, i.e. the proportion of liabilities backed by assets protected by property rights in favor of the plan (Valdés-Prieto 2002, 2005).

Another determinant of ρ^c is the tax treatment of the returns earned by the pension fund, if any. In most countries, these returns are partially exempt from income taxes, allowing the ρ^c for a fully-funded system to exceed the after-tax return earned by independent voluntary saving. This excess can be substantial even for individuals with a zero marginal tax rate on personal income, if corporate taxes on profits and dividends paid to voluntary savers are not fully credited towards personal income taxes, or cannot be fully claimed back, and if pension funds are allowed to claim back a larger portion of corporate taxes on profits and dividends (Feldstein and Liebman, quoted by Smetters 2005, 2002).

3. A model of labor supply in the presence of contributory pensions

3.1 Budget constraint for an individual

In each phase of life, the income identities before voluntary saving are[11]:

(1a) $$y_a(D) \equiv y^c \cdot \{D \cdot (1-\theta-t_a) + z^{ex} \cdot (1-D)\}$$

(1b) $$y_p(D) \equiv \theta \cdot (y^c \cdot D) \cdot (1+\rho^c) + e_p$$

It is instructive to obtain from (1) the marginal rate of transformation over time, between earnings in different phases of life, obtained by changing density:

(1c) $$MRT_{CS} \equiv -\frac{dy_p}{dy_a} \equiv -\frac{\partial y_p/\partial D}{\partial y_a/\partial D} = (1+\rho^c) \cdot \frac{1}{[\theta + t_a - (1-z^{ex})]/\theta}$$

[11] (1a) assumes that health insurance contributions have the same base as old-age pension contributions.

Equation (1c) reveals that the effective gross rate of return on contributions can be higher than $(1+\rho^c)$, even when the tax t_a is zero, if the earnings differential against uncovered jobs $(1-z^{ex})$ is larger than t_a.

The intuition is as follows: if the salary paid by the covered jobs is sufficiently above earnings in the uncovered sector, a worker would make only a minor sacrifice when taking the covered job, provided that the tax t_a does not compensate, and provided that ρ^c is not too small compared to the return on voluntary saving. Of course, it is also possible that the earnings differential be negative $(z^{ex} > 1)$, if the uncovered job is more productive, as would be the case for entrepreneurs. In this case the right hand factor is less than 1.

Eq. (1c) is the flipside of the well-known effect of financial returns on the value of a pension plan and thus on job choice. To formalize, define wealth as $W = y_a + e_p/(1+d)$, where d is some discount rate. From eq. (1) it follows that $\partial W/\partial D = [(1+d) - MRT_{cs}]\cdot k$, where k is a constant. When MRT_{cs} is less than $(1+d)$, as in most mature pay-as-you-go pensions, devoting more hours to covered jobs implies accepting an implicit tax on the rate of return earned by the saving bundled to that job. Since the financial return of a pension system affects its value, it also affects job choice.

To identify the impact of net tax t_a on the MRT_{CS}, consider the neutral case where $z^{ex} = 1$. The effective rate of return simplifies to $(1+\rho^c)/[1+(t_a/\theta)]$. Thus, differential taxation is a solid separate reason for an individual to reject bundling job choice with saving for old age. This impact is exacerbated when the contribution rate falls towards zero, because gains from social insurance fall to zero while the loss caused by tax t_a remains fixed.

The contribution rate θ also affects the MRT_{CS}. The sign of this influence depends on whether the net tax t_a exceeds, equals or is below the earnings differential $(1-z^{ex})$. In the central case of equality, these two effects cancel exactly. In figure 1, this central case is represented by point X1.

Now enter voluntary saving. Let F be the stock of independent and voluntary saving or dissaving in the financial market, at the end of the active phase. The real rate of return on pure saving depends on the sign of F and on taxation. Define:

$r(-)$: the rate of interest paid on consumer loans (when $F < 0$).
r: the pre-tax rate of interest paid to savers (when $F > 0$). Positive administration and marketing costs insure that $r(-) > r$.
τ_S: algebraic sum of the total tax rate levied on the return from voluntary saving, which includes the corporate tax rate and the personal income tax rate.
$r(+) \equiv r(1-\tau_S)$: the after-tax rate of interest paid to savers ($F > 0$).
$s \equiv r(-) - r(1-\tau_S)$ is defined as the net interest *spread* faced by individuals[12]. This spread is positive in most cases. However, the spread is negative when the rate of the fiscal incentive for voluntary saving for old age surpasses the sum of the corporate tax rate, the personal income tax rate and the cost of administration and marketing.

[12] This expression reflects the fact that the authorities add interest income to the taxable base, but do not allow interest paid in consumer loans to be deducted from the taxable income. It also reflects the fact that fiscal incentives for voluntary saving for old age channeled through the financial market are not affected by interest paid in consumer loans.

Because of the spread *s*, individuals cannot use cheap consumer credit to undo mandatory contributions. More generally, a positive spread reduces the consumption opportunity set (Valdés-Prieto 2002, Ch. 4)[13].

The period budget constraints are:

(2a) $c_a = y_a(D) - F - L_a$

(2b) $c_p = y_p(D) + F \cdot [1 + r(sign(F))]$

where T_a is a lump-sum tax levied in the active phase, used to balance the fiscal budget. This tax is elaborated in section 4.4.

3.2 Individual optimization in the active phase

The individual maximizes lifetime utility, which is assumed to be additive separable across phases of life. Labor supplies in both phases are assumed to be inelastic due to institutional constraints. This assumption is made to set aside possible changes in labor supply due to wealth changes, including those created by the mandate to contribute and by fiscal incentives for voluntary saving for old-age saving. Those changes would invalidate some of the graphical analysis presented below.

Still, labor supply to the covered sector remains endogenous because density *D* is endogenous. The individual solves the following program (P1) in the active phase:

$$Max_{\{D,F\}}\ U \equiv u(c_a) + v(c_p)$$

subject to

(1) $c_a = w \cdot x_a^c \cdot (1-\hat{l}_a) \cdot \left[D \cdot (1-\theta-t_a) + z^{ex} \cdot (1-D)\right] - F - L_a$

(2) $c_p = \theta \cdot (w \cdot x_a^c \cdot (1-\hat{l}_a) \cdot D) \cdot (1+\rho^c) + w \cdot \psi_p^a \cdot x_a^c \cdot (1-\hat{l}_p) + F \cdot [1 + r(sign(F))]$

(3) $0 \le D \le 1.$

As usual, it is assumed that *u'* > 0, *v'*> 0, *u''*< 0 and *v''* < 0. The Inada conditions apply.

Since the budget constraint is the result of competition between saving options that are linear, it may appear that only corner solutions will obtain for density *D*. This is correct for many parameter values, but not for an intermediate range. Rather than going through all possible combinations as in the Kuhn-Tucker conditions, corners are ordered using the following identity, obtained from (P1):

(3a) $\dfrac{\partial U}{\partial D} \equiv \left\{\dfrac{\partial U}{\partial F} \cdot \left(\theta + t_a - (1-z^{ex})\right) + v' \cdot \left[\theta(1+\rho^c) - \left(\theta + t_a - (1-z^{ex})\right) \cdot (1 + r(signF))\right]\right\} \cdot w x_a^c (1-\hat{l}_a)$

If $\left(\theta + t_a - (1-z^{ex})\right)$ is positive, (3a) can be reordered in a more revealing way as:

[13] This assumes that future pensions are not allowed to be alienated as collateral.

(3b) $$\frac{\partial U}{\partial D} \equiv \left\{ \frac{\partial U}{\partial F} + v' \cdot \left[MRT_{CS} - (1 + r(signF)) \right] \right\} \cdot \left(\theta + t_a - (1 - z^{ex}) \right) \cdot wx_a^c (1 - \hat{l}_a)$$

Equations (3) show that the relationship between independent saving (selection of F) and job choice (selection of density *D*) depends on the sign of two terms: $\left(\theta + t_a - (1 - z^{ex})\right)$, which is the relative net productivity of the uncovered sector as seen by the individual in the active phase, and the term in the square bracket of (3b).

If expression $\left(\theta + t_a - (1 - z^{ex})\right)$ is negative or zero, meaning that covered jobs dominate uncovered jobs despite taxes and contributions, both terms in the square brackets would be positive in equation (3a). If in addition independent saving *F* is an interior solution, so that $\partial U/\partial F$ is zero, the positive square bracket in (3a) makes $\partial U/\partial D$ positive always. Thus, *D** must be in the corner with *D* = 1* in this case. This is because a zero or negative $\left(\theta + t_a - (1 - z^{ex})\right)$ makes covered jobs dominant.

For cases where expression $\left(\theta + t_a - (1 - z^{ex})\right)$ is strictly positive, version (3b) can be used. If independent saving *F* is in an interior solution (so that $\partial U/\partial F$ is zero), the sign of $\partial U/\partial D$ is governed by the sign of the term in square brackets of (3b). This term is the difference between the return on saving through the contributory system (MRT_{CS}, eq. 1c) and the return on saving through the financial market. Either return may dominate. When this difference is negative, saving through the mandatory system is inferior to voluntary saving and the labor optimum is at the corner with the lowest possible density (*D*= 0*, case 1 below). If the term in the square bracket is positive, the labor optimum is at the corner with *D* = 1* (case 3 below). In a range of intermediate cases, an interior solution for *D** applies ($D^* \varepsilon\ [0,1]$, case 2 below). These results are summarized by:

PROPOSITION 1: In Bismarckian social insurance $(CB(y^c) = y^c)$ and in the absence of noncontributory subsidies for the old, the individual's optimum must be in one of only three situations, labeled 1 to 3 and characterized as follows:

a) 1: The productivity of uncovered jobs is high enough to make the individual prefer to unbundle saving from job choice. *D* = 0* and the sign and size of *F** depends on preferences. The condition for this case is $z^{ex} > 1 - t_a - \theta \cdot [1 - (1+\rho^c)/(1+r \cdot (1-\tau_S))]$.
b) 2: The productivity of uncovered jobs is intermediate. The condition for this case is[14]:

(4) $$s > MRT_{cs} - (1 + r \cdot (1 - \tau_S)) > 0$$

Depending on preferences, there are three subcases: 2a, where *D* = 0* and *F* < 0* (the individual issues consumer debt); 2b, where *D** is interior and *F* = 0* (he avoids the financial market); and 2c, where *D* = 1* and *F* > 0* (he saves in the financial market in addition to saving in the contributory system).

c) 3: The productivity of uncovered jobs is low enough. *D* = 1* and the sign and size of *F** depends on preferences. The condition is $z^{ex} < 1 - t_a - \theta \cdot [1 - (1+\rho^c)/(1+r \cdot (1-\tau_S)+s)]$.

[14] Note that MRT_{CS} depends on z^{ex}.

Proof: Based on figure 1. This figure shows possible consumption opportunity sets. The point labeled (D = 1, θ = 0) represents a case where there is no contributory pension system and the worker chooses the covered job (D = 1). The heavy dashed arrow labeled "θ", from point (D = 1, θ = 0) to the point labeled (D = 1, θ > 0) presents the impact on the endowment of a mandate to contribute, if density remained at 1. The change in endowment contracts the consumption opportunity set by the area shown in grey. This contraction has two distinct sources: first, figure 1 assumes that $\rho^c < r \cdot (1-\tau_s)$, which is likely under pay-as-you-go finance. Second, the presence of the spread $s>0$ cuts out the Southeastern portion of the consumption opportunity set. The overall contraction is a measure of the cost imposed on individuals that would have preferred a saving rate smaller than θ.

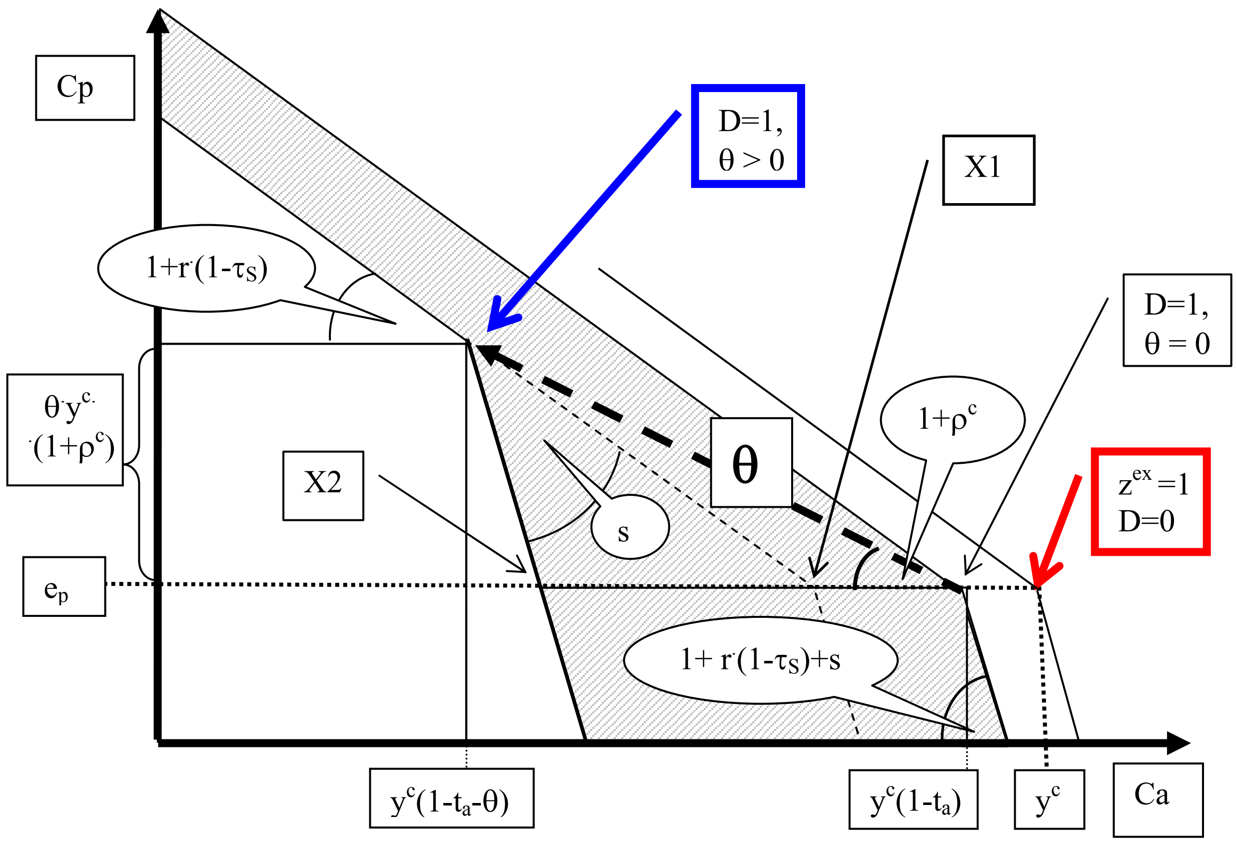

Figure 1: Consumption opportunity sets with and without a State-supported contributory pension plan, and with zero density in covered jobs at different earnings differentials.

Now consider the choice of density D. Choosing zero density, i.e. the exempt sector and unbundled saving for old age, yields an endowment located in the horizontal line that includes X1 and X2. These points differ only by z^{ex}, i.e. by the size of the earnings differential ($1- z^{ex}$). As the earnings differential rises, the point that represents that job (D = 0) moves to the left along the horizontal. When the earnings differential is zero, ($z^{ex} =1$), the endowment is the point thus labeled. In that case the consumption opportunity set is larger than in (D=1, θ=0) due to the presence of tax t_a. In contrast, choosing full density takes the endowment to the point labeled (D = 1, θ > 0).

Figure 2 presents the set of possible budget constraints, for different values of z^{ex}. The A points are endowments obtained by choosing zero density. Each thick line that starts in B and goes down to one of the A points, represents one set of endowments that the individual can achieve by varying density. The slope of each thick line AB is the net return of mandatory saving as compared with uncovered jobs, i.e. MRT_{CS}.

These endowment lines are supplemented by lines that represent borrowing and lending opportunities in the financial market. In figure 2, opportunities for financial saving are represented by the dashed lines that go Northwest. The individual may also issue consumer debt, represented by the dotted lines that go Southeast.

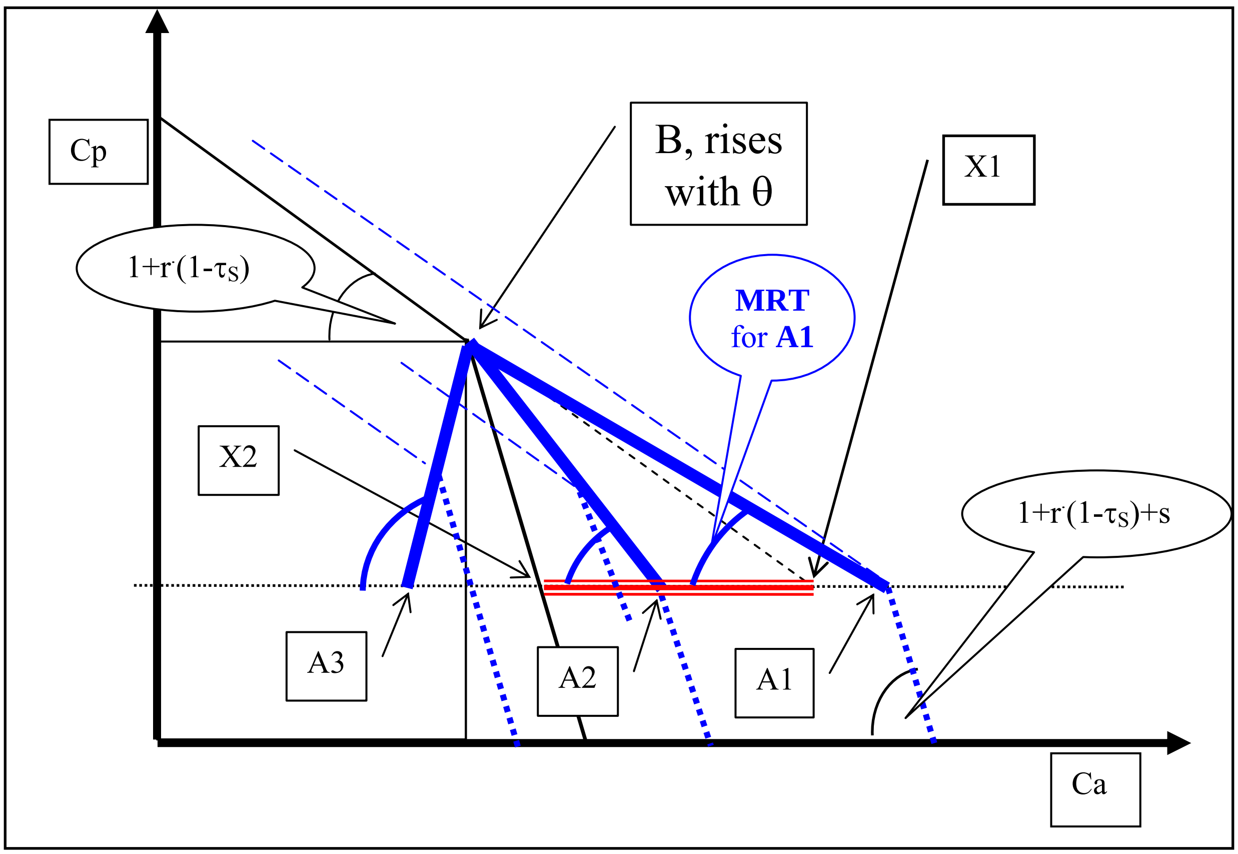

Figure 2: Budget constraints in a mandatory contributory pension plan (B = full density in covered jobs; A = zero density)

If the productivity of the uncovered job z^{ex} is high enough to push point A1 to the right of point X1, line AB has a smaller slope than $(1+r(1-\tau_S))$ and the consumption opportunity set is maximized by choosing D* = 0. This can be seen in that the financial lines would create a smaller consumption opportunity set if they started from any point along A1B, which have D > 0. The intuition is that any saving is made more efficiently through voluntary vehicles. The consequence is that uncovered jobs are dominant and density is zero.

At the other extreme, if the productivity of the uncovered job z^{ex} is low enough to push point A3 to the left of point X2, line AB has a higher slope than $(1+r'(1-\tau_S)+s)$ and the consumption opportunity set is maximized by choosing D* =1. Again the financial lines would create a smaller consumption opportunity set if they started from any point along A3B, which have D < 1.

The set of intermediate positions for A, in which the slope of line AB is larger than $(1+r'(1-\tau_S))$ but smaller than $(1+r'(1-\tau_S)+s)$, can create intermediate job choices. Now the financial lines that start from intermediate points along A2B do not dominate each other, and their outer envelope defines the consumption opportunity set. Depending on preferences, the optimal density may be interior in A2B, or in one of the extremes, with the financial market providing the required extra saving or dissaving.

Three common myths can be dispelled at this stage: First, it has been asserted that since the current needs of the poor outweigh any gain from saving for the long term, the poor would always minimize density of contribution (Titelman and Uthoff 2005, Van Ginneken 2007). Second, that illiquidity of pension rights, implying subvaluation of contributory pensions, always leads to a reduced density of contribution (Diamond and Valdes-Prieto 1994). Third, that any worker that (partially) neglects old age would choose a reduced density of contribution.

However, figure 2 shows that if the productivity of uncovered jobs is low enough, so that point A is located to the left of X2, the individual still finds the covered job preferable, and chooses full density of contribution despite any of these three traits. Thus, the presence of any of these three traits does not imply zero density of contribution.

4. Optimal social insurance for the middle class

This section introduces a benevolent role for social insurance for the middle classes, in the absence of noncontributory subsidies for the old. Redistributive concerns are assumed to be taken care of by other institutions, such as the tax and transfer system. It is well-known that when individuals visualize old age correctly and act as planned, they would always save adequately for old age in the absence of social insurance. If that case were prevalent enough, there would be no benevolent role for a public policy that mandates saving for old age in some covered sector, nor for policies that give fiscal incentives to save for old age, as opposed to incentives for general saving.[15] In the United States, Social Security and 401(k) plans would be superfluous.[16]

4.1 A benevolent role for social insurance

A possible - but ultimately unconvincing- benevolent explanation for social insurance is that the financial system is less efficient than the State in some areas. In the context of

[15] Incentives for general savings can be a method to transform an income tax into a consumption tax. But this approach should object to targeting savings incentives to favor specific uses, such as old age.

[16] The socially optimal social insurance policy would be $\theta^* = 0$ and τ_S^* equal to the total effective tax rate on general savings.

saving, one claim is that the financial market for voluntary annuities fails due to adverse selection. In the Chilean pension system for the middle class, participants are allowed to choose at pension age between an annuity and a programmed withdrawal, so adverse selection could be substantial. However, the empirical impact of adverse selection is weak, if any (Rocha and Thorburn 2007, p. 138), and is clearly smaller than the impact of sales effort by annuity companies.[17] In any case, if adverse selection were significant, the optimal intervention would be a mandate to purchase an annuity, when reaching pension age, of a size linked to an individual wealth assessment (as in car insurance), and not a mandate to save in certain "covered" jobs.

Another claim about the inefficiency of financial markets is that they are incomplete because intergenerational risk is not traded (Demange 2002, Ball and Mankiw 2008). However, the proposition that the State is capable of providing efficient intergenerational insurance ignores political risk. The existence of revolutions, wars and other events that put an end to States suggests that political risk rises fast with the length of the horizon. Simulations for the U.S. find that the value of intergenerational insurance is modest, even when adding the value of insurance between the labor and capital shares of GNP (Krueger and Kubler, 2006).

The valuable and growing literature on time inconsistency does not offer a rationale for social insurance. Time-inconsistency, for example in the form proposed by Laibson (1997), creates a voluntary demand for commitment devices to save for general purposes. This demand is voluntary and not limited to saving for old age. Since the private supply of commitment devices is plentiful, if each individual is rational when assessing the value of commitment devices for herself, the market equilibrium should be efficient. Thus, time inconsistency by itself does not justify mandates nor fiscal incentives to use commitment devices.

Consumer inertia and devices that modify the default option can have a major impact on market equilibria (Madrian and Shea, 2001). Inertia is also important for the performance of quasi-markets for merit goods, defined as markets where the State gives individuals vouchers to purchase some predefined service, at regulated quality and price levels. However, when individuals are rational, they all value being relieved from their own inertia. Thus, there exists a voluntary demand for jobs or services bundled with anti-inertia devices. Indeed, the first experiments on these anti-inertia devices in the U.S. were made at a private firm without State action. Thus, the State's role may be limited to popularizing the supply of anti-inertia devices. Another role for the State may be to create anti-inertia devices such as the one attached to Kiwisaver in New Zealand (the State created a default option serviced by suppliers chosen by auction). However, if demand is rational when suppliers of anti-inertia devices charge nor more than average cost, the State does not have a further benevolent role. In that setting, fiscal incentives to save for old age, such as the generous ones created by New Zealand in 2007 and proposed for Britain from 2012, are not justified, and the same applies to a mandate.

[17] Moreover, the adverse selection hypothesis predicts that if an individual who correctly estimates his risk chooses one insurance contract over another offering more coverage, then it must be true that his loss probability under the contract chosen is strictly lower than the per unit premium of the additional coverage offered by the other contract. This is not true in the data. The puzzle disappears if some individuals underestimate their risk, as in Koufopoulos (2005), but then unrealistic optimism reappears.

The econometric evidence presented by Torche and Wagner (1997) using Chilean individual data for 1990 shows that on average, workers that are indifferent between covered and uncovered jobs consider that about *half* of their mandatory contribution for old age is a pure tax[18]. This negative valuation of the mandate obtains in a pension system that is relatively efficient: it is fully funded (market returns), it enjoys tax preferences larger than those offered to normal financial saving (in the sample period), charges lower administrative fees than most local financial saving vehicles, has access to a menu of financial assets as large as the menu accessible through voluntary saving, has a modest contribution rate (10%) and suffers from modest political risk (in the sample period). Despite all this, the perceived burden of this mandate to save is large (half of the contribution). This is incompatible with a substantial positive valuation of anti-inertia devices and commitment devices.

Two explanations for the Torche-Wagner evidence remain: (a) the perceived burden is the outcome of the illiquidity of mandatory saving; and (b) there exits at least one more layer of bias in the demand for long-term saving by young individuals.

In the illiquidity hypothesis the old agree with the young in detesting the mandate to save, on a lifetime basis.[19] An optimistic bias is also consistent with young individuals detesting the mandate to participate in social insurance. However, individuals do recognize their own bias gradually over time, and correct it. This occurs slowly enough to eliminate significant saving possibilities. Older individuals regret the saving amounts chosen when young, as too skimpy. Critically, older individuals become grateful for social insurance. This reversal of opinion is unique to the hypothesis of optimistic bias when young.

An extensive literature in psychology documents a high prevalence of “unrealistic optimism” in the population (Weinstein 1980, Weinstein and Klein 1996, Puri and Robinson 2007). Surveys find that workers raise their assessment of the best age to retire considerable as they age (Lehr 1980, p. 228). Other studies show that the cognitive ability to represent and feel a future that does not (yet) exist varies considerably in the population (Feuerstein, 1980). Research in economics and finance has begun to use the unrealistic optimism hypothesis intensively (Camerer and Lovallo 1999, Malmendier and Tate 2005).

The optimistic bias can take several forms regarding old age saving, which may operate simultaneously. One is that relative labor productivity in old age may be smaller than hoped for when young. Another is that the duration of old age may be underestimated, and a third is that the out-of-pocket expenses caused by the decline in health may be underestimated (Diamond, 1977). Evidence for Chile shows that women aged 20-65 underestimate their life expectancy by about 6 years, while men aged 30-60 do so by about 2 years (PACPR 2006, p. 94). Another form is that the speed at which time inconsistency is overcome may

[18] Later confirmed by further econometric work by Edwards and Cox-Edwards (2002), with data for 1994. Gruber (1997) found the opposite but uses aggregate data at the plant level for 1979-1986, which is less reliable. In addition, that paper builds variables in a way that creates a systematic spurious correlation, as acknowledged by the author.

[19] A difference also emerges for fiscal incentives to save for old age: Since saving with fiscal incentives is voluntary and the stock of such saving as substantially liquid (if design is efficient), the illiquidity hypothesis implies 100% take-up of subsidies to saving when young and no regret when old. In contrast, the hypothesis of optimistic bias when young implies reluctance and limited take-up when young, and regret when old because own take-up is now regarded as insufficient. The 401(k) data for the U.S. favors the latter.

be overestimated (O'Donoghue and Rabin 1999). In any case, substantial heterogeneity in optimistic bias has been documented (Lusardi, 2000). Unrealistic optimism may be more relevant for irreversible long-term decisions that are experienced only once per lifetime.

An optimistic bias among young individuals does not imply a large consumption surprise at the date of retirement. Surprises would occur gradually over time, say starting at age 40, so consumption drops should be small in any given year. The empirical literature on this "retirement consumption puzzle" (Banks et al 1998, Bernheim, Skinner and Weinberg 2001, Aguiar and Hurst 2005, Hurd et al 2008) evaluates an excessively specialized version of the optimistic bias hypothesis, where almost all the surprises occur at the very end of active life. Gradual awareness about the high cost of old age still implies that many individuals lose significant saving opportunities in the absence of social insurance.[20]

4.2 A simple model of unrealistic optimism among the young

In the model of this section, when reaching old age and assessing his own lifetime welfare backwards, each individual recognizes she suffered an optimistic bias when young: her own former assessment of her productivity ratio when old ψ^{a}_{p} is relabed as ψ^{biased}_{p}. Now, at old age, she realizes that the correct productivity ratio ψ^{p}_{p} is smaller:

(5) $0 \leq \psi^{p}_{p} < \psi^{biased}_{p} < 1$

In lives with just two periods, recognition of this bias can only occur in old age. At that age, the only possible adjustment is a cut in consumption when old. This would not be the case with many periods, because saving and labor supply could adjust smoothly as recognition advanced. With two-period lives there is no need to model the learning process about one's own unrealistic optimism. Mathematically, this model is represented by problem P2, where V is lifetime welfare as assessed by the old individual. P2 is:

$$V \equiv u(c_a) + v(c_p) \qquad where$$

$$(A) \qquad c_a = w \cdot x^c_a \cdot (1-\hat{l}_a) \cdot \left[D^{biased} \cdot (1-\theta-t_a) + z^{ex} \cdot (1-D^{biased}) \right] - F^{biased} - L_a$$

$$(B) \qquad c_p = \theta \cdot (w \cdot x^c_a \cdot (1-\hat{l}_a) \cdot D^{biased}) \cdot (1+\rho^c) + w \cdot \psi^p_p \cdot x^c_a \cdot (1-\hat{l}_p) + F^{biased} \cdot [1 + r(sign(F^{biased}))]$$

$$(C) \qquad (D^{biased}, F^{biased}) = Arg\,Max_{\{D,F\}} \left\{ u(c_a) + v(c^{biased}_p) \right\}$$

$$subject\ to$$

$$(1) \qquad c_a = w \cdot x^c_a \cdot (1-\hat{l}_a) \cdot \left[D \cdot (1-\theta-t_a) + z^{ex} \cdot (1-D) \right] - F - L_a$$

$$(2) \qquad c^{biased}_p = \theta \cdot (w \cdot x^c_a \cdot (1-\hat{l}_a) \cdot D) \cdot (1+\rho^c) + w \cdot \psi^{biased}_p \cdot x^c_a \cdot (1-\hat{l}_p) + F \cdot [1 + r(sign(F))]$$

$$(3) \qquad 0 \leq D \leq 1.$$

---

[20] The simulation by Scholz et al (2006) finds that only 20% of the age group between 51 and 61 were at risk of inadequate replacement rates in 1992. However, other data for 2004 suggests that 32% of that age group was at risk of inadequate replacement rates, mainly due to drop in coverage of DB pensions since 1992 (Munnell et al 2007). More important, the simulation by Scholz et al (2006) does not model financial markets realistically: by assuming zero intermediation costs it allows individuals to obtain long-term consumer credit at 4% real interest rate after tax. This generous credit line, which can be used contingently, reduces the precautionary demand for saving artificially. Ignoring uncertainty in returns also reduces the precautionary demand for savings. This leads to low "projected optimal" savings stock as of retirement age. Given observed savings stocks, this reduces artificially the proportion of the age group at risk.

Note that problem P1 is embedded within problem P2. In P1, expected future earnings when old, given in equation (2), are based on $\psi^{biased}{}_p$, the productivity ratio for old age expected when active, which is biased upwards. In contrast, equation (C) reflects recognition and elimination of the bias, so earnings when old are given by the actual productivity ratio for old age, $\psi^p{}_p$. Due to the bias, consumption when old is smaller or equal than the consumption for old age that was planned when young:

(6) $$b \equiv c_p^{biased} - c_p = \left[\psi_p^{biased} - \psi_p^p\right] \cdot w \cdot x_a^c \cdot (1 - \hat{l}_p) \geq 0$$

where b is the bias in absolute terms. Since $v'' < 0$ it follows that $v'(c_p) > v'(b_p)$.

Figure 3 presents problem P2 graphically, and is taken from Valdes-Prieto (2002, p. 167-9). It reveals that this particular case of optimistic bias does not rely on modifying preferences, but on modifying the lifetime budget constraint.

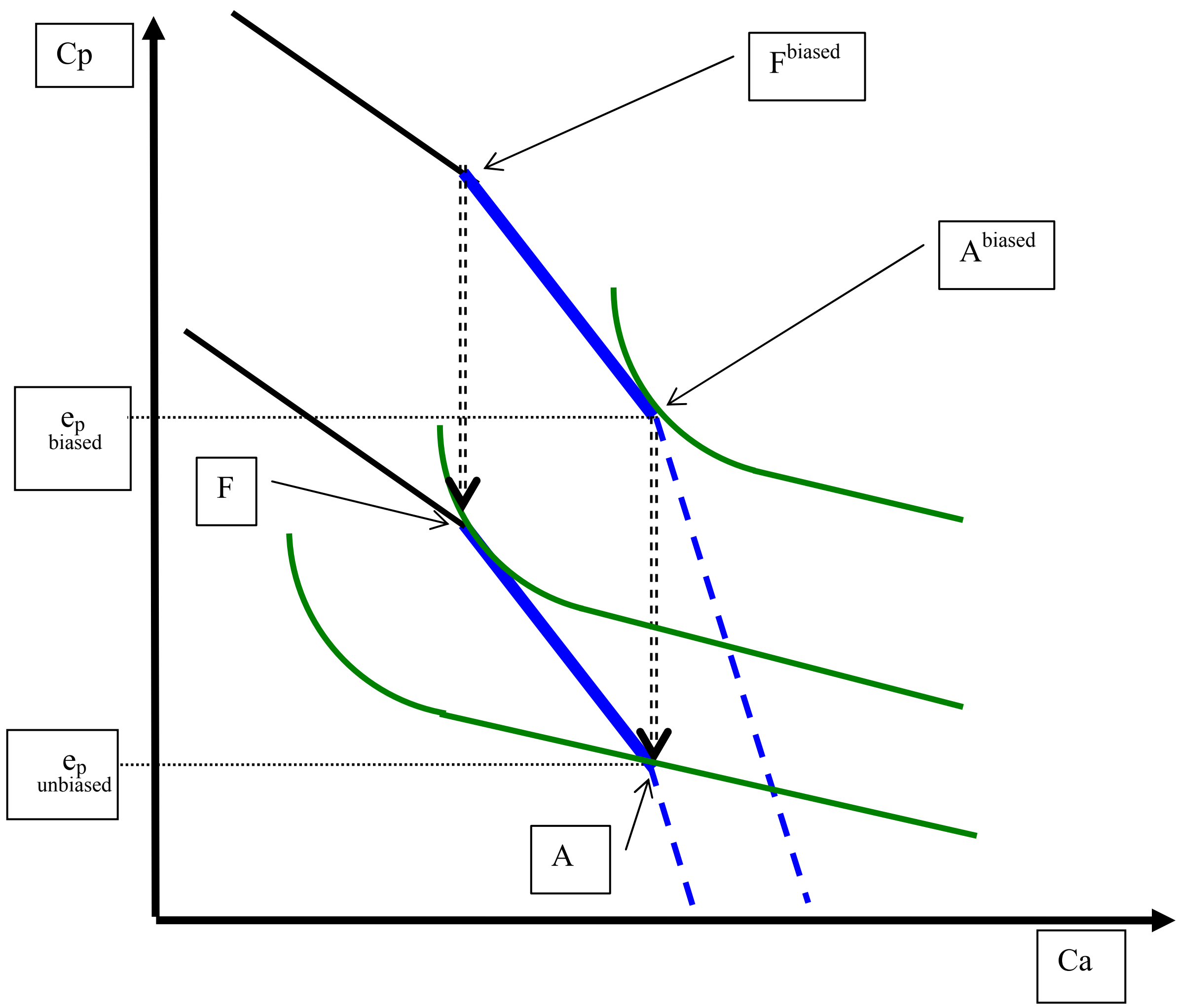

Figure 3: Impact of the optimistic bias on consumption and welfare assessed when old

4.3 The best personalized social insurance

This section analyzes the welfare problem that a benevolent planner would face is she could take orders from the individual when old, i.e. after the bias is recognized and eliminated, and applied them to the young individual. In this problem, the planner chooses the Bismarckian contribution rate θ, to be applied when the individual is active. The planner is time-consistent.

Since the planner takes orders from a single individual, social insurance becomes personalized, i.e. adjusted to each individual's traits such as the productivity differential between the covered and exempt jobs available to him.

In this section the impact on the fiscal balance is not taken into account, because each individual is atomistic with respect to the budget. Problem P3 is:

$$Max_{\{\theta\}} V(\theta) \equiv u(c_a) + v(c_p)$$

where

$$(A) \qquad c_a = w \cdot x_a^c \cdot (1-\hat{l}_a) \cdot \left[D^{biased} \cdot (1-\theta-t_a) + z^{ex} \cdot (1-D^{biased}) \right] - F^{biased} - L_a$$

$$(B) \qquad c_p = \theta \cdot (w \cdot x_a^c \cdot (1-\hat{l}_a) \cdot D^{biased}) \cdot (1+\rho^c) + w \cdot \psi_p^p \cdot x_a^c \cdot (1-\hat{l}_p) + F^{biased} \cdot [1 + r(sign(F^{biased}))]$$

$$(C) \qquad (D^{biased}, F^{biased}) = Arg\, Max_{\{D,F\}} \left\{ u(c_a) + v(c_p^{biased}) \right\}$$

subject to

$$(1) \qquad c_a = w \cdot x_a^c \cdot (1-\hat{l}_a) \cdot \left[D \cdot (1-\theta-t_a) + z^{ex} \cdot (1-D) \right] - F - L_a$$

$$(2) \qquad c_p^{biased} = \theta \cdot (w \cdot x_a^c \cdot (1-\hat{l}_a) \cdot D) \cdot (1+\rho^c) + w \cdot \psi_p^{biased} \cdot x_a^c \cdot (1-\hat{l}_p) + F \cdot [1 + r(sign(F))]$$

$$(3) \qquad 0 \le D \le 1.$$

As shown in Proposition 1, P1 has several corner solutions. In situation 1, the best choice of density is D*= 0, so only uncovered jobs are taken. Since changes in θ do not affect welfare, the planner is *unable to attenuate* the impact of the optimistic bias.

At the same time, the prevalence of situation 1 regarding variations in z^{ex}, the relative productivity of uncovered jobs, does depend on θ: as shown by Proposition 1, a higher θ makes situation 1 more likely if $\rho^c < r(1-\tau_S)$ and less likely if $\rho^c > r(1-\tau_S)$.

Now move to situation 3. In this case the best choice of density is D* = 1, i.e. only covered jobs are taken and changes in θ affect welfare directly. If F* is interior $u'(c_a)$ can be eliminated with the FOC of P1 and a first-order Taylor expansion of $v'(b_p)$ in terms of $v'(c_p)$ allows simplification to:

$$(7) \quad \frac{\partial V}{\partial \theta} \approx \left[\rho^c - r(signF) \right] \cdot v'(c_p) + \left[\psi_p^{biased} - \psi_p^p \right] \cdot w \cdot x_a^c \cdot (1-\hat{l}_p) \cdot \left\{ -v''(c_p) \right\} \cdot \left(1 + r(signF) + \frac{\partial F^{biased}}{\partial \theta} \right)$$

Equation (7) identifies two impacts of a change in θ for situation 3: first, the well-known difference in after-tax returns between the mandatory saving and independent saving.

Second, the optimistic bias, which is always positive if only the first term of the Taylor expansion around $v'(c_p)$ is considered. Therefore, the positive sign of the second term applies to "small" biases only.

The positive sign of the term for the optimistic bias implies that the socially optimal rate for social insurance is positive. There exists a benevolent role for social insurance. Conversely, if there is no bias (if $\psi^a_p = \psi^a_p$) and if the return on saving in the mandatory pension system equals the return on independent saving (if $\rho^c = r(1-\tau_S)$), then the socially optimal contribution rate is θ = 0 and there is no role for social insurance.

As before, the prevalence of situation 3 also depends on θ: a higher θ makes situation 3 less likely if $\rho^c < r(1-\tau_S)+s$, and more likely if $\rho^c > r(1-\tau_S)+s$.

In situation 2, there are three subcases regarding the choice of density. In case 2a, the choice is D*=0, so the impact of changes in θ is the same as in situation 1: only the prevalence of situation 2a is affected. In case 2c, the choice of density is D*=1, so the impact of changes in θ is the same as in situation 3: the socially optimal rate for social insurance is positive and there exists a benevolent role for social insurance. The prevalence of situation 2c is affected too.

In case 2b, the optimal density is interior and $F^{biased} = 0$. Since D^{biased} is interior $u'(c_a)$ is eliminated with the FOC of P1 and a first-order Taylor expansion of $v'(b_p)$ in terms of $v'(c_p)$ allows simplification to:

$$\text{(8)}\quad \frac{\partial V}{\partial \theta}\frac{1}{y^c} \approx \left[(1+\rho^c) - MRT_{cs}\right]\cdot D^{biased}\cdot v'(c_p) + \left[\psi_p^{biased} - \psi_p^p\right]\cdot w\cdot x_a^c\cdot(1-\hat{l}_p)\cdot\left\{-v''(c_p)\right\}\cdot\left(D^{biased}\cdot MRT_{CS} + \theta\cdot\frac{\partial D^{biased}}{\partial \theta}\cdot(1+\rho^c)\right)$$

The sign of the first term is restricted in situation 2. Eq. (4) requires the term in the square bracket to remain between $[\rho^c - (r(1-\tau_S)+s)]$ and $[\rho^c - r(1-\tau_S)]$, so in most cases the first term is negative.[21] The first term represents the fact that in situation 2b an interior density allows the individual to raise consumption at a smaller cost that the interest of consumer credit, even if the return on saving in the mandatory pension system equals the return on independent saving (if $\rho^c = r(1-\tau_S)$). If in addition the pension plan offers a below-market return, the first term is even more negative. The second term is proportional to the optimistic bias and is positive for θ = 0. Therefore, the sum of both terms can have any sign in situation 2b, and the socially optimal rate for social insurance is zero or positive. These results are summarized in:

PROPOSITION 2: When a benevolent planner takes orders from the individual when old, after the optimistic bias is recognized and eliminated:

a) The personalized optimal contributory rate for individualized Bismarckian social insurance is, following the situations identified by Proposition 1: (i) zero in situations 1 and

[21] An exception is the introductory phase of pay-as-you-go finance, where ρ^c is very high due to implicit intergenerational subsidies. Of course, in that situation individuals beg to be incorporated to social insurance.

2a; (ii) ambiguous in situation 2b; and (iii) positive in situations 2c and 3. There exists a benevolent role for personalized Bismarckian social insurance in case (iii).

b) The optimal Bismarckian contribution rate differs across individuals. It is a function of the degree of optimistic bias, the earnings differential against uncovered jobs, the interest rate faced in independent saving, marginal income tax rates and preferences.

The second part of Proposition 2 is proven by observation of eqs. (7) and (8). They show that the personalized optimal rate Bismarckian insurance depends on the personal parameters identified there.

4.4 Optimal uniform social insurance for the Bismarckian case

Due to asymmetric information between the State and individuals about his heterogeneity, the State can only use a uniform contribution rate. This uniformity creates additional inefficiencies, which are expected to reduce the optimal size of social insurance. This section presents the basis for simulations that identify the optimal size of a uniform Bismarckian contribution rate, which takes into account the extra restriction imposed by asymmetric information.

To focus on social insurance for the middle class, labor earning is assumed to be the same for all individuals, both when active and in old age without bias. Thus, in this section $x^c{}_a$ = 1 for all individuals in the same generation ($x^c{}_a$ does grow over time, see eq. (0)). This eliminates the demand for progressive redistribution.

The heterogeneity that remains is bidimensional. First, individuals differ in the size of the optimistic bias given by the difference $b \equiv (\psi^{biased}{}_p - \psi^p{}_p)$. Second, individuals differ in the productivity of uncovered job options, z^{ex}. The density function is $f(z^{ex}, b)$.

General equilibrium demands attention to the fiscal balance and a framework for the time dimension. An overlapping generations setting is assumed, with lives lasting two periods of equal duration. Each period lasts about 30 years and "old age" begins at age 50.

The revenue of three taxes must be considered for fiscal balance. First, the net tax rate t_a on earnings from covered jobs in the active phase. Second, the tax levied on the return from voluntary saving. Third, there exists a tax L_a , levied only in the active phase, whose role is to balance the fiscal budget. Since the labor supply and savings response to changes in contribution rates affects aggregate tax bases, tax revenue is affected too. If an adjustment tax were not considered, the amount of resources available to individuals after tax would change in response to changes in θ, and this would bias the social evaluation.

For analytical transparency, the adjustment tax L_a is assumed to be a lump-sum, of which λ% is dissipated in efficiency losses and tax administration. This approach avoids modeling the distortion associated to the adjustment tax. The advantage is that the identity of those distortions can affect the optimal level of θ. Because of this, the optimal contribution rates found by this paper finds should be seen as informative, but subject to further adjustments in response to the details of the tax system.

The timing of the adjustment tax is chosen to prevent intergenerational redistributions that may reduce analytical transparency. If the adjustment tax applied in old age, changes in that tax would affect the welfare of the generation that is old at the date of the change. This impact would have to be distributed across all subsequent generations to preserve fiscal balance. This repercussion is avoided by locating the adjustment tax in the active phase.[22]

Now consider the time framework. The tax t_a applies to active individuals, but τ_S is levied on the old. The respective tax bases differ in proportion to the growth rate of the economy between these two overlapping generations. The fiscal balance condition is:

$$(9) \qquad R = \iint \left[t_a \cdot y^c \cdot D^{biased} + \tau_S \cdot \frac{Max(0, r \cdot F^{biased})}{1+g_{CR}} + L_a \cdot (1-\lambda) \right] \cdot f(z^{ex}, b) dz^{ex} db$$

where R is required revenue by the government expressed as a proportion of the aggregate earnings in the active phase, which is a constant, g_{CR} is the growth rate of tax bases, which is also the growth rate of contribution revenue[23], and λ is the proportion of the lump sum tax paid by individuals that is dissipated in inefficiencies and tax administration costs.

The fiscal revenue requirement must be positive when the other branches of social insurance are not financed from general taxation, i.e. if $L_a = 0$. Thus, two alternative values for R should be considered:

(10)

$$R \equiv \begin{cases} 0 & \textit{if general taxes finance other branches} \\ \iint \left[t_a \cdot y^c \cdot D^{biased}\big|_{\theta=0} + \tau_S \cdot \dfrac{Max(0, r \cdot F^{biased}\big|_{\theta=0})}{1+g_{CR}} \right] \cdot f(z^{ex}, b) dz^{ex} db & \textit{if not} \end{cases}$$

In the second case, the tax bases are the ones that would remain if the contribution rate θ for old age pensions were zero, i.e. in the absence of social insurance for old age.

Social welfare is defined as the simple sum of the lifetime utilities achieved by individuals, as assessed by those same individuals when old. Inequality aversion is not taken into account because all individuals belong to the same middle-class group: all enjoy the same labor productivity. Although some individuals have higher productivity in exempt jobs and are somewhat richer, the difference is modest.

In this context, the planner's problem is P4:

---

[22] The economy considered is always in a steady state, because the assumption of an open economy includes assuming that international capital flows keep the actual capital stock in its desired level in each period.

[23] $g_{CR} \equiv d\ln(\theta \cdot D \cdot y^c)/dt = d\ln w/dt + d\ln x_a^c/dt$ in a steady state where θ and D are constant over time.

$$Max_{\{\theta\}}W \equiv \iint \left[u(c_a)+v(c_p)\right]\cdot f(z^{ex},b)dz^{ex}db$$

subject to :

$$(A)\qquad c_a = w\cdot x_a^c \cdot (1-\hat{l}_a)\cdot\left[D^{biased}\cdot(1-\theta-t_a)+z^{ex}\cdot(1-D^{biased})\right]-F^{biased}-L_a$$

$$(B)\qquad c_p = \theta\cdot(w\cdot x_a^c\cdot(1-\hat{l}_a)\cdot D^{biased})\cdot(1+\rho^c)+w\cdot\psi_p^p\cdot x_a^c\cdot(1-\hat{l}_p)+F^{biased}\cdot[1+r(sign(F^{biased}))]$$

$$(C)\qquad (D^{biased},F^{biased}) = ArgMax_{\{D,F\}}\left\{u(c_a)+v(c_p^{biased})\right\}$$

subject to

$$(1)\qquad c_a = w\cdot x_a^c\cdot(1-\hat{l}_a)\cdot\left[D\cdot(1-\theta-t_a)+z^{ex}\cdot(1-D)\right]-F-L_a$$

$$(2)\qquad c_p^{biased} = \theta\cdot(w\cdot x_a^c\cdot(1-\hat{l}_a)\cdot D)\cdot(1+\rho^c)+w\cdot(\psi_p^p+b)\cdot x_a^c\cdot(1-\hat{l}_p)+F\cdot[1+r(sign(F))]$$

$$(3)\qquad 0\le D\le 1.$$

$$(D)\ R = \iint\left[t_a\cdot w\cdot x_a^c\cdot(1-\hat{l}_a)\cdot D^{biased}+\tau_S\cdot\frac{Max(0,r\cdot F^{biased})}{1+g_{CR}}+L_a\cdot(1-\lambda)\right]\cdot f(z^{ex},b)dz^{ex}db$$

Note that the contribution rate θ is allowed to take negative values.

4.5 A subsidy to restore pension adequacy

In the previous section, an increase in the contribution rate induces some individuals to switch to exempt jobs and reduce density. This reduces pension adequacy and weakens the ability of the mandate to counteract unrealistic optimism. Thus the policy tool used to restore savings in section 4.4 distorts the labor market too.

A possible policy response is to complement the mandate with a subsidy to earnings on covered jobs. To restore the density of contribution, this subsidy should be targeted to covered jobs alone. This idea is not new: (Corsetti et al (1997) analize a subsidy to covered earnings and find that it should equal the tax rate perceived by the young.

Such a subsidy can be presented in two other ways: First, as a subsidy to contributions. a fiscal co-payment enhances each contribution. Second, it can be presented as a cut in the net tax rate t_a levied to finance other branches of social insurance. This works only because t_a applies only in covered jobs. A cut to t_a requires that some of the financing for those other branches of social insurance be moved to general tax revenue. In this second presentation, the bundle offered by covered jobs is improved by higher net non-wage benefits provided by other branches of social insurance.

These subsidies may seem to create incentives to raise density to socially *excessive* levels among the young that exhibit a small or zero bias about old age earnings (no unrealistic optimism). However, the net tax rate t_a is defined as the tax that finances other branches of social insurance, minus the marginal benefit from those branches. Since this difference tends to be substantial, the density decision is biased by t_a in the direction of *lower* density even when θ is zero. The case of excessive density is implausible.

Of course, subsidies consume fiscal revenue and thus may raise tax distortions. According to equation (9), a reduction in t_a necessitates an increase in the lump sum tax L_a. A reduction in t_a modifies the tax bases as well, because it induces individuals to modify increase density and reduce the stock of voluntary savings.

When the social planner is allowed to optimize t_a for a given level of θ, a static fiscal policy issue appears: the net tax rate t_a can be adjusted until its marginal social cost equals the deviation λ between the private and social cost of funds. Recall that λ is the proportion of the lump sum tax that is dissipated in inefficiencies and tax administration costs.

This is important because this static optimization can raise welfare even when the size of social insurance is kept constant. For analytical transparency, this static fiscal issue should be isolated from the question of the optimal size of social insurance.

One way to achieve this isolation is to set the value of the required revenue R at the level R^* that emerges from P5 when only t_a is optimized and old-age social insurance is absent (θ = 0). In turn, R* should be the revenue required by other branches of social insurance in an initial situation in which substitution of t_a for L_a was not allowed because L_a was zero, and where $t_a = t_a^{\,0}$.

$$(11)\quad R^* \equiv \iint \left[t_a^0 \cdot y^c \cdot D^{biased}\big|_{\theta=0} + \tau_S \cdot \frac{Max(0, r\cdot F^{biased}\big|_{\theta=0})}{1+g_{CR}} \right] \cdot f(z^{ex}, b)dz^{ex}db$$

In this context, the planner's problem is P5:

$$Max_{\{\theta,t_a\}} W \equiv \iint \left[u(c_a) + v(c_p) \right] \cdot f(z^{ex}, b)dz^{ex}db$$

subject to:

$$(A)\qquad c_a = w\cdot x_a^c \cdot (1-\hat{l}_a)\cdot \left[D^{biased}\cdot(1-\theta-t_a) + z^{ex}\cdot(1-D^{biased}) \right] - F^{biased} - L_a$$

$$(B)\qquad c_p = \theta\cdot(w\cdot x_a^c\cdot(1-\hat{l}_a)\cdot D^{biased})\cdot(1+\rho^c) + w\cdot\psi_p^p\cdot x_a^c\cdot(1-\hat{l}_p) + F^{biased}\cdot[1+r(sign(F^{biased}))]$$

$$(C)\qquad (D^{biased}, F^{biased}) = Arg\,Max_{\{D,F\}}\left\{ u(c_a) + v(c_p^{biased}) \right\}$$

subject to

$$(1)\qquad c_a = w\cdot x_a^c\cdot(1-\hat{l}_a)\cdot\left[D\cdot(1-\theta-t_a) + z^{ex}\cdot(1-D) \right] - F - L_a$$

$$(2)\qquad c_p^{biased} = \theta\cdot(w\cdot x_a^c\cdot(1-\hat{l}_a)\cdot D)\cdot(1+\rho^c) + w\cdot(\psi_p^p + b)\cdot x_a^c\cdot(1-\hat{l}_p) + F\cdot[1+r(sign(F))]$$

$$(3)\qquad 0 \le D \le 1.$$

$$(D)\quad R^* = \iint \left[t_a\cdot w\cdot x_a^c\cdot(1-\hat{l}_a)\cdot D^{biased} + \tau_S\cdot\frac{Max(0, r\cdot F^{biased})}{1+g_{CR}} + L_a\cdot(1-\lambda) \right]\cdot f(z^{ex}, b)dz^{ex}db$$

The differences between P5 and P4 are that in P5 t_a is also a choice variable for the planner, and that the revenue requirement is set at the level R* that obtains in a previous stage.

5. Redistributive concerns and non-contributory pensions

This section extends the model to incorporate noncontributory subsidies for the old poor ("first pillar pensions"). The standard rationale for these subsidies is progressive redistribution. Therefore, the model must be extended to incorporate substantial income inequality: This can is done by allowing labor productivity to vary across individuals, both when active and in old age without bias. Thus, in this section the productivity parameter $x^c{}_a$ has a distribution, and is an additional source of heterogeneity.

This extension is important for policy because non-contributory old-age subsidies can affect density choices. This section shows that non-contributory subsidies crowd-out contributory pensions ("second pillar pensions") and reduce pension adequacy. This effect is larger among those on lower incomes, because the amount of non-contributory subsidies must be attractive for their presence to make a difference. This is not the case for the higher middle classes unless the non-contributory subsidy is large and universal.

More generally, since equilibrium density depends on the design of the combined "multipillar" structure, optimal social insurance should be designed jointly with the non-contributory subsidies.

5.1 Extending the model to include a first pillar

This section introduces a flexible family of noncontributory subsidies, that encompasses flat universal pensions, minimum pensions and even assistance pensions. Valdés-Prieto (2002, pp. 57 and 70) summarized this design with the following formula:

$$\text{(12a)} \quad NCS(D) = \max\left[0 \quad ; \quad BP - \gamma \cdot CP(D)\right]$$

where CP(D) is the contributory pension financed by the participant, BP is a basic pension or subsidy given to those that never contributed to the contributory and mandatory system (second pillar), and γ is the subsidy's "withdrawal rate", with $\gamma \varepsilon (0,1)$. This withdrawal rate can also be defined as γ = BP/MCPWS, where MCPWS is the "maximum contributory pension with subsidy".

Equation (12a) encompasses a number of policies. A minimum pension subsidy obtains by making $\gamma = 1$ (and assuming that no vesting requirements apply). A flat universal pension obtains with $\gamma = 0$ (or MCPWS = ∞), as in Denmark and New Zealand.

Intermediate values for γ create linear subsidies with gradual withdrawal. In 1957 Finland introduced the first noncontributory pension of this type, with a withdrawal rate of 0% in a small initial range (up to the equivalent to 2% of the average wage), and then in a second and dominant tranche with a withdrawal rate of 50%, applied to the amount of the contributory pension that exceeds that small initial amount (Antolin et al 2001). About 80% of all Finnish pensioners obtain some supplement from this first pillar.

The new Swedish scheme, applied since 2003, differs from (12a) in that two withdrawal rates apply, ordered in a convex way. The noncontributory subsidy is withdrawn at a rate of 100% in the initial range, which is substantial. In a second tranche, the withdrawal rate is 48% in Sweden (Scherman, 1999). The Swedish design was imitated in the Norwegian reform of 2005, to be applied from 2010. In Norway the withdrawal rates are 100% for the initial range and 60% for the second tranche (Pedersen, 2005).

The Chilean reform of 2008 will introduce gradually a new unified subsidy with a single withdrawal rate in response to the individual's mandatory funded pension. It differs from the Finnish one in that the withdrawal rate is substantially lower, about 32% (Valdes-Prieto, 2008b). Another difference is that the new subsidy is also withdrawn according to a means test, based on the average income per capita in the beneficiary's houisehold.

In the model of section 3, eq. (12a) turns into:

$$\text{(12b)} \quad NCS \equiv \max\left[0 \quad ; \quad BP - \gamma \cdot (y^c \cdot D \cdot \theta \cdot (1+\rho^c)\right]$$

The individual's optimization program P1 must be modified to incorporate the first-pillar pension: the budget constraint for old age is modified to:

$$\text{(2b')} \qquad c_p = y_p(D) + F \cdot [1 + r(sign(F))] + NCS(D, BP, \gamma)$$

For interior solutions for D, when voluntary saving is held constant:

$$\text{(13)} \quad \left.\frac{\partial c_p}{\partial D}\right|^{S=constant} = \frac{\partial y_p}{\partial D} + 0 + \frac{\partial NCS^{SW}}{\partial D} = (1-\gamma) \cdot y^c \theta (1+\rho^c)$$

Equation (13) indicates that consumption in the old age increases in proportion to $(1-\gamma)$ when density rises. Each \$1 of additional contributory pension is used by the state to withdraw its subsidy by \$$\gamma$ cents, leaving \$$(1-\gamma)$ cents to the participant, to increase his total pre-tax pension (which combines subsidies and contributory pensions). In present discounted value terms, the implicit tax rate on covered earnings is $\gamma\ \theta(1+\rho^c)/(1+r)$.

As before, local analysis is insufficient. In cases 1 and 2a, where D* = 0, the non-contributory subsidy is BP, and its presence reduces the incentive to contribute even more. In cases 3 and 2c, where D* =1, the noncontributory subsidy is minimized by the high density. If γ is large enough relative to the spread s, the individual may be drawn to switch to case 2b.

Figure 4 shows the income opportunity set faced by an individual in case 2b, where D* can be interior. When choosing the contribution density that leads to a consumption plan like point B', the individual considers both contributory and non-contributory pensions.

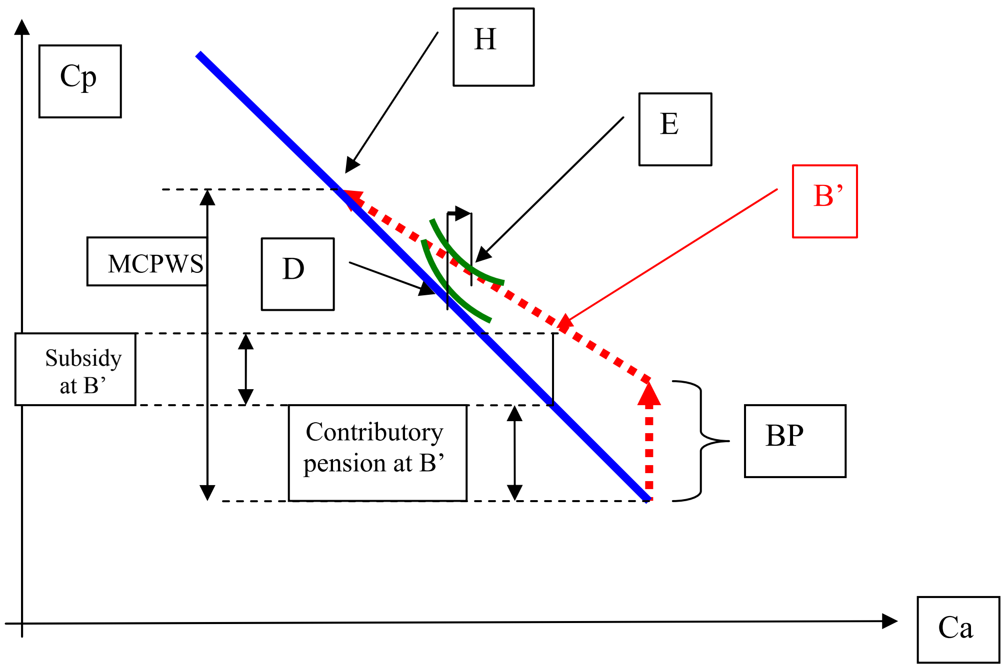

Figure 4: Budget constraint with a non-contributory pension (case 2b)

Figure 4 shows that for preferences in this range, the introduction of this noncontributory subsidy reduces density from point D to E. Therefore, the non-contributory subsidy crowds out contributory pensions. Figure 4 also shows that the reduction in density is *smaller* when the subsidy withdrawal rate γ is smaller. As the withdrawal rate is reduced towards zero, the crowding out of the contributory system is alleviated by more.

On the other hand, a smaller γ raises the fiscal cost, and this forces the planner to choose between cutting the basic pension *BP* to avoid an increase in the fiscal cost, and increasing tax rates, raising other distortions.

This result is not valid for all individuals. Figure 4 reveals that the non contributory subsidy creates a non-convexity in the budget set in point H. Almost no individual chooses point H. For individuals with preferences in the range of H, the optimal response to a small cut in γ can trigger a discrete cut in density. Cuts in the subsidy withdrawal rate can have the unintended result of reducing density.

In addition, the value of $(1-\gamma)$ and *BP* also affects the frontier between cases 2a and 2b, and also the frontier between cases 2b and 2c. This affects the tax bases.

5.2 The "moral hazard" justification for mandatory social insurance

An alternative justification for benevolent social insurance, which does not rely on unrealistic optimism, can be evaluated in this setting. This hypothesis starts by recognizing that a demand for progressive redistribution justifies the presence of non-contributory subsidies for the old poor, and that non-contributory subsidies distort labor supply and voluntary saving.

This hypothesis posits that to limit the losses from these distortions, a benevolent government should mandate everybody to save for old age (Becker et al XXX). The intuition is that individuals whose decisions have been distorted by the non-contributory pension would be forced by the mandate to save the efficient amount. This intervention takes place despite the absence of unrealistic optimism.

Thus, non-contributory subsidies for the old would *crowd-in* mandates to contribute for old age, through the benevolent policymaker's reaction function. The justification for social insurance would be the need to counteract "moral hazard", i.e. the change in behavior triggered by the non-contributory subsidies.

Is this hypothesis consistent? The mandate does alleviate savings distortions and may substitute for taxation externalities. However, the mandate distorts density (the labor decision) as well, and this distortion is not repaired. Indeed, a larger distortion of density substitutes for a smaller distortion in total saving (Valdés-Prieto 2002, Ch. 5). Moreover, a mandate justified in this way should be targeted to individuals whose behavior is in fact distorted by the non-contributory subsidies, and therefore should exempt the middle class and higher earners. Contributory social insurance does the opposite, as it forces all earners to save for old age. A mandatory contribution rate reduces welfare for the middle class, because a uniform contribution rate does not take into account individual heterogeneity, and because large spreads between borrowing and lending defeat attempts to undo the mandate to save with cheap consumer credit..

To evaluate this hypothesis, consider a situation with income inequality, such as the one analyzed in this section. There is a demand for progressive redistribution, so a non-contributory pension is present. It distorts saving and density, which a benevolent planner would want to counteract. Assume also that unrealistic optimism is absent, so $\psi^{p}{}_{p} = \psi^{biased}{}_{p}$ for all: no young individual has a biased expectation about old age. Is the socially optimal contribution rate positive? Assume further that the rate of return on contributions ρ^c is equal to market interest rates, so $\rho^c = r$. If $\theta^* > 0$, the moral hazard hypothesis would be supported.

When the planner has many policy tools, such as progressive income taxes and means-tested subsidies for the young poor, the tax-expenditure system is more efficient and taxation externalities are smaller. This reduces the scope for the hypothesis to be supported. To the contrary, if the planner has fewer policy tools, the hypothesis is more likely to succeed this test because taxation externalities are larger.

As a first approximation, consider the following income tax-expenditure system, which is relatively simple[24]:

[24] A negative income tax is not feasible in this paper's economy because the earnings of exempt workers is not observed by the tax authorities.

$$(14)\quad IT_a = Max\left(0 \quad ; \quad t_R \cdot (y^c \cdot D \cdot (1-\theta - t_a) - Q)\right)$$
$$(15)\quad IT_p = Max\left(0 \quad ; \quad t_R \cdot (y_p(D) + r(sign(F)) \cdot F - Q)\right)$$
$$(16)\quad S_a = Max\left(0 \quad ; \quad \gamma \cdot (Q - c_a)\right)$$
$$(17)\quad NCS \equiv \max\left[0 \quad ; \quad \gamma \cdot (Q - y^c \cdot D \cdot \theta \cdot (1+\rho^c))\right] \qquad (BP = \gamma Q)$$

There is a single marginal tax rate t_R charged to the high-income individuals. Taxable income is defined as covered earnings that exceed Q, in the active phase. For those in old age, taxable income is defined as contributory pensions plus earnings in excess of Q. Interest income is taxed at rate τ_S, as before. Parameters t_R and Q are age-independent. The authorities are assumed to run a proxy means test that measures consumption in the active phase, and this allows targeting a subsidy for the young poor. To simplify, the withdrawal rate for this subsidy is set at the same value as the withdrawal rate of the non-contributory pension. To simplify further, the maximum consumption and contributory pension at which subsidies disappear are also set at Q.

The critical feature of this tax-expenditure system is that it allows the social planner to manage the distortions caused by non-contributory pensions in a targeted manner: a lower subsidy for the low earners in the active phase, when coupled with a higher non-contributory subsidy in old age, is equivalent to a targeted mandate to save for old age. The critical aspect is targeting, because this allows the middle class to be exempted from the mandate. If this is more efficient, the optimal size of social insurance should be zero.

The fiscal balance depends also on the revenue requirement, which is labeled now as R**. This amount is set as in section 4.5. R** is be the revenue required by other branches of social insurance in an initial situation in which substitution of t_a for L_a was not allowed because L_a was zero, and where $t_a = t_a^0$.

The final aspect to consider is the social welfare function. Since there is income inequality, the issue of social aversion to inequality emerges. This paper follows the literature and assumes a constant inequality aversion parameter v, as in Slemrod et al (1994). The social planner solves problem P6, which is:

$$Max\, W(\theta, t_a, \tau_R, \gamma, Q) \equiv (1/1-v) \iint \left[u(c_a) + v(c_p) \right]^{1-v} \cdot f(z^{ex}, x_a^c) dz^{ex} dx_a^c$$

subject to:

$$(A) \qquad c_a = w \cdot x_a^c \cdot (1-\hat{l}_a) \cdot \left[D \cdot (1-\theta-t_a) + z^{ex} \cdot (1-D) \right] - F - L_a + S_a - IT_a$$

$$(B) \qquad c_p = \theta \cdot (w \cdot x_a^c \cdot (1-\hat{l}_a) \cdot D) \cdot (1+r) + w \cdot \psi_p^p \cdot x_a^c \cdot (1-\hat{l}_p) + F \cdot [1 + r(sign(F))] + NCS - IT_p$$

$$(C) \qquad (D, F) = ArgMax_{\{D,F\}} \left\{ u(c_a) + v(c_p) \right\}$$

$$(D)\ R^{**} = \iint \left[t_a \cdot w \cdot x_a^c \cdot (1-\hat{l}_a) \cdot D - S_a + IT_a + \frac{\tau_S \cdot Max(0, r \cdot F) - NCS + IT_p}{1 + g_{CR}} + L_a \cdot (1-\lambda) \right] \cdot f(z^{ex}, x_a^c) dz^{ex} dx_a^c$$

If the solution to this problem yields $\theta^* > 0$, the moral hazard hypothesis is supported. If not, it should be discarded. In that case, unrealistic optimism remains as an important benevolent justification of social insurance.

5.3 The socially optimal withdrawal rate under unrealistic optimism

In this section the paper returns to the assumption that unrealistic optimism is present and faces the question of the optimal design of a two-pillar pension system, jointly with a tax-expenditure system. Since income inequality is acknowledged, there are three dimensions of individual heterogeneity: z^{ex}, b and x_a^c.

Our review of international experience shows that an important policy question is the optimal size of the old-age subsidy withdrawal rate γ. The question is whether intermediate values for γ are socially superior to both minimum pensions ($\gamma = 1$) and to universal flat pension ($\gamma = 0$).

Another important question is whether the contributory old-age pension system should incorporate some progressive redistribution in its benefit formula. This would imply abandoning the Bismarckian character of social insurance. Alternatively, the tax-expenditure system would be able to satisfy redistributive concerns alone, and a progressive benefit formula would not be justified. Another way to pose this question is whether a special progressive tax on contributory pensions is superfluous. The special tax on contributory pensions that characterizes non-Bismarckian redistributive social insurance can be modeled as:

$$(18)\ T_p^{RSI} = Max\left(0 \quad ; \quad \tau^{RSI} \cdot \left\{ \theta \cdot (w \cdot x_a^c \cdot (1-\hat{l}_a) \cdot D) \cdot (1+\rho^C) \right\} - P \right)$$

where τ^{RSI} is the marginal tax rate and P is the threshold above which this tax applies.

This special tax provides extra flexibility for overall redistribution, because its tax base is different from those of other taxes. This may add some value, but it also causes additional distortions, as it moves away from the tradition of maximizing the tax base to minimize tax rates. The optimal value of this tax rate could be zero, or the optimal threshold P could be zero. In these cases this tax would be superfluous and the notion that contributory pensions should be purely Bismarckian would be supported.

These questions can be answered by a social planner that solves problem P7:

$$Max\, W(\theta, t_a, \gamma, t_R, Q, \tau^{RSI}, P) \equiv (1/1-\nu)\iiint \left[u(c_a) + v(c_p)\right]^{1-\nu} \cdot f(z^{ex}, b, x_a^c) dz^{ex} db dx_a^c$$

subject to:

$$(A) \qquad c_a = w \cdot x_a^c \cdot (1-\hat{l}_a) \cdot \left[D^{biased} \cdot (1-\theta-t_a) + z^{ex} \cdot (1-D^{biased})\right] - F^{biased} - L_a + S_a - T_a$$

$$(B) \qquad c_p = \theta \cdot (w \cdot x_a^c \cdot (1-\hat{l}_a) \cdot D^{biased}) \cdot (1+\rho^c) + w \cdot \psi_p^p \cdot x_a^c \cdot (1-\hat{l}_p) + F^{biased} \cdot [1 + r(sign(F^{biased}))] + NCS - T_p - T_p^{RSI}$$

$$(C) \qquad (D^{biased}, F^{biased}) = ArgMax_{\{D,F\}} \left\{u(c_a) + v(c_p^{biased})\right\}$$

subject to

$$(1) \qquad c_a = w \cdot x_a^c \cdot (1-\hat{l}_a) \cdot \left[D \cdot (1-\theta-t_a) + z^{ex} \cdot (1-D)\right] - F - L_a + S_a - T_a$$

$$(2) \qquad c_p^{biased} = \theta \cdot (w \cdot x_a^c \cdot (1-\hat{l}_a) \cdot D) \cdot (1+\rho^c) + w \cdot (\psi_p^p + b) \cdot x_a^c \cdot (1-\hat{l}_p) + F \cdot [1 + r(sign(F))] + NCS - T_p - T_p^{RSI}$$

$$(3) \qquad 0 \le D \le 1.$$

$$(D) \;\; R = \iiint \left[t_a \cdot w \cdot x_a^c \cdot (1-\hat{l}_a) \cdot D^{biased} - S_a + T_a + \frac{\tau_S \cdot Max(0, r \cdot F^{biased}) - NCS + T_p + T_p^{RSI}}{1 + g_{CR}} + L_a \cdot (1-\lambda) \right] \cdot f(z^{ex}, b, x_a^c) dz^{ex} db dx_a^c$$

If the solution to this problem involves a subsidy withdrawal rate at intermediate levels, where $0 < \gamma^* < 1$, the Nordic approach to non-contributory pensions would obtain support. Similarly, if the solution to this problem involves $\tau^{RSI} = 0$, i.e. a zero marginal tax rate for the special tax on contributory pensions, the Bismarckian design would be supported.

6. Final remarks

This paper offers a model where there is a benevolent role for social insurance for the middle classes, due to unrealistic optimism. In contrast, non-contributory subsidies cater to a social demand for redistribution towards the old poor. This model offers a framework to guide an integrated design of “multipillar” pension systems (World Bank, 1994).

In this model contribution density is endogenous, allowing for the modest coverage rates observed in the main emerging economies, which together with mobility and churning between covered and uncovered jobs, implies serious adequacy problems

It is found that the withdrawal rate of the subsidy affects the rate of return of the combined “two-pillar” system. Therefore, contribution density can be improved by streamlining the design of noncontributory subsidies, to minimize this crowding-out effect. A subsidy with a small withdrawal rate is likely to be superior to both minimum pensions and to universal flat pensions.

An important topic for future research is whether the withdrawal rate should be the same in response to different types of income when old. One such income comes from mandatory pensions, another source of income comes from pure saving, a third comes from family transfers (spouses, children) and a final type of income comes from labor earnings (when old). The proper order of these withdrawal rates is important for policy. Another topic for research is the relative performance of mandates to save for old age, and fiscal incentives for voluntary saving for old-age, say at a flat rate for all income levels.

References

Abel, A., G. Mankiw, L. Summers and R. Zeckhauser (1989) "Assessing dynamic efficiency", *The Review of Economics Studies*, Vol. 56 (1) No. 185, January, p. 1-20.

Aguiar, M. and Hurst, E. (2005) "Consumption versus Expenditure", *Journal of Political Economy,* Vol 113, nº 5 (October).

Andolfatto, D. (2002) "A theory of inalienable property rights", *Journal of Political Economy*, 110, 2, April, p. 382-393.

Antolin, P., H. Oxley and W. Suyker (2001), "How Will Ageing Affect Finland?", *OECD Economics Department Working Papers*, No. 295, OECD Publishing.

Attanasio, O. and M. Browning (1996) "Consumption over the life cycle and over the business cycle", *American Economic Review* 85 (5), p. 1118-37.

Auerbach, A. and L. Kotlikoff (1987) *Dynamic Fiscal Policy*, Cambridge University Press.

Becker et al (XXX)

Bernheim, D., J. Skinner and J. Weinberg (2001) "What accounts for the variation in retirement wealth among U.S. households" *American Economic Review* 91, 4 p. 832-857.

Berstein, S., Reyes, G. and Pino, F. (2005). "Independent Workers: should they be mandated or subsidized to contribute? A third option", *En foco* Document Expansiva Nº 66 (in Spanish), www.expansiva.cl

Beveridge, W. (1942) *Social Insurance and Allied Services*, also *Beveridge Report*, HMSO reprint, London.

Beyer, H. and S. Valdés-Prieto (2004) "Proposal to increase the density of contribution", report for national Office for Women (SERNAM), March, Santiago; later version presented at international seminar organized by Centro de Estudios Públicos and Superintendency of AFP, November 11-12, Santiago (in Spanish).

Bucheli, M., A. Forteza and I. Rossi (2007) "Work history and the access to contributory pensions: The case of Uruguay", Working Paper, Departamento de Economía, Facultad de Ciencias Sociales, Universidad de la República, Uruguay.

Camerer, C. and D. Lovallo (1999) "Overconfidence and excess entry: an experimental approach", *American Economic Review* XXX

Cigno, A. and M. Werding (2007) *Children and Pensions*, The MIT Press.

Diamond, P.A. (1977) "A framework for social security analysis", *Journal of Public Economics* 8, p. 275-298.

Diamond, P. (1998) “Optimal income taxation: An example with a U-shaped pattern of optimal marginal tax rates”, *The American Economic Review* Vol. 88, No. 1, March, 83-95.

Diamond, P. and S. Valdes-Prieto (1994) “Social Security Reforms”, chapter 6 in Bosworth, B. R. Dornbusch and R. Labán (eds.), *The Chilean Economy: Policy Lessons and Challenges*, Brookings Institution, Washington, D. C. pp. 257-328.

Edwards, S. and A. Cox Edwards (2002) "Social Security Privatization Reform And Labor Market: The Case Of Chile", *Economic Development and Cultural Change*, v50 (3,Apr), 465-489.

Feldstein, M. (2005) “Rethinking Social Insurance”, NBER Working Paper 11250, March.

Godoy, Oscar and S. Valdés-Prieto (1997) “Democracy and pensions in Chile: Experience with two systems”, chapter 4 in S. Valdés-Prieto (ed.) *The Economics of Pensions*, Cambridge University Press, also in *Cuadernos de Economía*, Vol. 31, N° 93, pp. 135-160.

Gruber, J. (1997) “The Incidence of Payroll Taxation: Evidence for Chile”, *Journal of Labor Economics*, Vol. 15, No. 3, Part 2: Labor Market Flexibility in Developing Countries (Jul., 1997), pp. S72-S101.

Hubbard, G., J. Skinner and S. Zeldes (1995), “Precautionary saving and social insurance”, *Journal of Political Economy,* 103 (2), April.

Hurd, M. and S. Rohwedder (2008) “The Retirement Consumption Puzzle: Actual Spending Change in Panel Data” NBER Working Paper 13929, April.

Koufopoulos, K. (2005) “Asymmetric Information, Heterogeneity
in Risk Perceptions and Insurance: An Explanation to a Puzzle”, FMG Discussion Paper, Warwick Business School, September.

Krueger, D. and F. Kubler (2006) “Pareto improving social security reform when financial markets are incomplete!?” *American Economic Review* 96, 737-755.

Laibson, D. (1997) “Golden Eggs and Hyperbolic Discounting”, *Quarterly Journal of Economics*, Vol. 112, No. 2, May , Pages 443-477.

Madrian, B. and D. Shea (2001) “The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior”, *The Quarterly Journal of Economics*, Vol. 116, No. 4 (Nov.), pp. 1149-1187.

Malmendier, U. and G. Tate (2005) “CEO Overconfidence and Corporate Investment”, *The Journal of Finance*, XXX

Mirrlees, J. (1971), “An exploration in the theory of optimum income taxation”, *Review of Economic Studies*, 38, p. 175-208.

Munnell,A., A. Webb and F. Golub-Saas (2007) “Is There Really A Retirement Savings Crisis? An Nrri Analysis”, Center for Retirement Research, Boston College, Brief 7-11.

Ted O'Donoghue, T. and M. Rabin (1999) “Doing it now or later”, *The American Economic Review*, Vol. 89, No. 1 (March), pp. 103-124.

Palme, J. (2005), “Features of the Swedish pension reform”, *The Japanese Journal of Social Security Policy*, Vol. 4, No.1 (June), p. 42-53.

Pedersen, A.W. (2005) “Two technical choices with critical implications: Issues in Scandinavian pension reform”. *NOVA-Norwegian Social Research* Working paper 6/05, Oslo.

Poblete, Dante (2005) “Optimal degree of targeting of noncontributory pensions in the Third World”, Thesis in Economics, Instituto de Economía, Universidad Católica de Chile, www.economia.puc.cl

Puri, M. and D.T. Robinson (2007) “Optimism and economic choice” *Journal of Financial Economics*, XXX

Rocha, R. and C. Thorburn (2007) *Developing Annuity Markets: The Experience of Chile*, The World Bank, Washington D.C., 251 pages.

Salditt, F., P. Whiteford and W. Adema (2007) "Pension Reform in China: Progress and Prospects", OECD Social, Employment and Migration Working Papers No. 53, OECD, Paris.

Scholz, J.K., A. Seshadri and S. Khitatrakun (2006) "Are Americans saving optimally for retirement?" *Journal of Political Economy,* Vol 114, nº 4 (August). p. 607-643.

Slemrod, J., S. Yitzhaki, J. Mayshar and M. Lundholm (1994) "The optimal two-bracket linear income tax", *Journal of Public Economics* 53, 2, p. 269-290.

Titelman, D. and Uthoff, A. (2005), "The Role of Insurance in Social Protection in Latin America", *International Social Security Review*, Vol. 58, 2-3/2005, Oxford (UK) and Malden (USA), p. 43-68.

Torche, A. and G. Wagner (1997) "Social Security: individual valuation of a mandated benefit ", *Cuadernos de Economía*, No 103, p. 363-390, (in Spanish) available in www.economia.puc.cl.

Valdés-Prieto, S. (2002) *Pension policies and pensions markets: a university textbook for Latin America*, Ediciones Universidad Católica, 930 pages, Santiago (in Spanish). Available in www.edicionesuc.cl

Valdés-Prieto, S. (2005) "Pay-as-you-go Securities", *Economic Policy*, 42, April, p. 215-251, CEPR.

Valdés-Prieto, S. (2007) "Solidarity pension: advantages, defects and proposals", *Temas de la Agenda Pública* series Nº 13, December, Santiago de Chile, (in Spanish), www.puc.cl/agendapublica/

Valdés-Prieto, S. (2008a) "A theory of non-contributory pension design", Working Paper Nº 335, Instituto de Economía, U. Católica de Chile, Santiago, in www.economia.puc.cl

Valdes-Prieto, S. (2008b) "The 2008 Chilean reform to first-pillar pensions" (companion paper), mimeo Instituto de Economia, Catholic University of Chile.

van Ginneken, W. (2007) "Extending Social Security Coverage: Concepts, Global Trends and Policy Issues" . *International Social Security Review*, Vol. 60, No. 2-3, April-September, pp. 39-57.

van Ginneken, W. (editor) (1999) *Social Security for the Excluded Majority*, ILO, Geneva.

Weinstein, N.D. (1980) "Unrealistic optimism about future life events", *Journal of Personality and Social Psychology*, Vo. 39, Nº 5, p. 806-820.

Weinstein, N.D. and W.M. Klein (1996) "Unrealistic Optimism: Present and Future", *Journal Of Social And Clinical Psychology*, Guilford Press.

World Bank (1994) Averting the old age crisis: Policies to protect the old and promote growth, Oxford University Press.

World Bank (2000) *The Korean Pension System at a Crossroads,* Report Nº2004-KO, Washington D.C.

Appendix: Parametrization for simulation exercises

To facilitate comparability, the parametrization used in this paper follows Auerbach and Kotlikoff (1987, p. 27):

(A.1) $u(c_a, \hat{l}_a) = (1/(1-1/\gamma)) \cdot \left[c_a^{\ 1-1/\varphi} + \alpha \cdot \hat{l}_a^{\ 1-1/\varphi} \right]^{(1-1/\gamma)/(1-1/\varphi)}$ $\gamma \neq 1$

(A.2) $v(c_p, \hat{l}_p) = (1/(1+\delta)) \cdot (1/(1-1/\gamma)) \cdot \left[c_p^{\ 1-1/\varphi} + \alpha \cdot \hat{l}_p^{\ 1-1/\varphi} \right]^{(1-1/\gamma)/(1-1/\varphi)}$ $\gamma \neq 1$

where γ, δ, α and φ are taste parameters.

The parameter values suggested by Auerbach and Kotlikoff (1987, p. 50-52) are used here as well, so γ= 0.25, φ =0.80 and α=1.5. This last parameter is consistent with a proportion of time devoted to leisure in active life of $\hat{l}_a$ = 0.60, which given a labor endowment of 5,000 hours per year, implies 2,000 hours of work per year.

The next assumption is that δ is 1.5% per annum, from the same source. This implies that $\delta = (1+0.015)^{30} - 1 = 0.563$ or 56.3% per 30-year period. Since average labor supply during old age lasts only up to age 65, the hours of work during "old age" are half of those worked while active. Thus, the proportion of time devoted to leisure in "old age" is $\hat{l}_p$ = 0.80.

Consider the parameters for the pension system. To evaluate Bismarckian social insurance on a stand-alone basis, fiscal incentives for old-age saving are assumed to be zero. The net internal rate of return paid by the contributory system to each generation, ρ^c, depends on the financing method and on the tax treatment of the returns of pension funds. To represent a range of experiences, two values for ρ^c are considered: $\rho^c = g_{CR}$, for pay-as-you-go finance, and $\rho^c = r$ for fully funded finance with tax exemptions for pension funds.

The net tax rate on earnings from covered jobs in the active phase is set at t_a = 0.15. This figure is based on common values for the sum of the contribution rates for other branches of social insurance different from old age, across countries, and on the assumption that in those branches the marginal increase in benefits for individuals that contribute more is zero.

The tax rate levied on the return from voluntary saving, which combines the corporate tax rate and the personal income tax rate, is set at τ_S = 0.20. This implies that the after-tax return on independent voluntary saving is reduced to 2.4% per year and to $(1+0.024)^{30} - 1$ = 103.70% per 30-year period. This tax rate is levied only when the individual chooses a positive amount of saving. If the individual chooses negative net saving (consumer credit), there is no tax deduction proportional to the interest paid. Thus, the tax rate τ_S does not reduce the spread s.

Following studies on the tax burden and the tradition of regulatory economics, the proportion of the adjustment tax that is dissipated in inefficiencies is set at $\lambda = 0.30$.

Pre-tax factor prices are given by the international economy. The assumed values are $w = 1$ and $r = (1+0.03)^{30} - 1 = 1.427$ (142.7% per 30-year period). The spread in consumer credit is set at 10 percentage points per annum, according to the empirical evidence for the U.S. This implies that $r+s = (1+0.03+0.10)^{30} - 1 = 38.116$ (3,811.6% per 30-year period).

In a small open economy, the growth rate of the economy and of tax bases is given by worldwide technological progress, which is exogenous for this country. It is assumed that labor productivity rises at 1% per year, and that the reward to effective labor remains constant over time. Thus $g_{CR} = 0 + d \ln x_a^c / dt = (1+0.01)^{30} - 1 = 0.3478$ (34.8% per 30-year period).

Since all individuals have the same labor productivity, x^c_a = 1 for the generation that is young in the present period (x^c_a does grow over time, see below). Ageing is also assumed to affect all individuals to the same degree. Given the absence of empirical data on ageing, three values for $\psi^p{}_p$ are considered: 0.70, 0.60 and 0.50.

Individuals differ in the size of the optimistic bias ($b \equiv \psi^{biased}{}_p - \psi^p{}_p$). Since $\psi^p{}_p < \psi^{biased}{}_p < 1$, it follows that $0 < b < 1- \psi^p{}_p$. Given the absence of empirical data on the distribution of parameter b, it is assumed that it is uniformly distributed, within the bounds given by the assumption on ageing: $b \; \varepsilon \; [0 \, , \, 1- \psi^p{}_p]$. The sensitivity analysis performed on the value of $\psi^p{}_p$, which takes the values 0.70, 0.60 and 0.50, implies sensitivity analysis on the distribution of the bias b.

Individuals also differ in the productivity of uncovered job options, z^{ex}. The absence of empirical data on the distribution of parameter z^{ex} also recommends sensitivity analysis. It is assumed that z^{ex} is uniformly distributed: $z^{ex} \; \varepsilon \; [k - 0.5 \, , \, k + 0.5]$, where k takes three values: k = 0.9, 0.8 and 0.7 ; This specification keeps constant the degree of inequality in z^{ex}. The higher value of k may represent an economy where earnings in the covered sector are not much better than in uncovered jobs, while entrepreneurs with high z^{ex} are relatively well rewarded. The lower value of k represents an economy where uncovered jobs are much less productive and where entrepreneurs with high z^{ex} are not so well rewarded. Finally, it is assumed that b and z^{ex} are independent. Thus $f(z^{ex},b) = 1/(1- \psi^p{}_p -0) \cdot (0.5+k-(k-0.5))$ for all z^{ex} and b.

Expected results for the optimal uniform social insurance rate

A uniform contribution rate creates additional inefficiencies originated in heterogeneity across individuals. The main question to be answered in this section is whether the benefits of a Bismarckian mandate are sufficient to compensate those inefficiencies, in the view of individuals that assess lifetime utility when old, free from unrealistic optimism.

Problem P4 is solved numerically with MATLAB. Table 1 shows the results.

Table 1: Optimal uniform contribution rate to social insurance (to be completed)
(The other parameter assumptions are indicated in this Appendix)

$\theta^*_{uniform}$	$\rho^c = g_{CR}$ = 1% per annum (real)			$\rho^c = r$ = 3% per annum (real)		
	$\psi^p{}_p = 0.70$	$\psi^p{}_p = 0.60$	$\psi^p{}_p = 0.50$	$\psi^p{}_p = 0.70$	$\psi^p{}_p = 0.60$	$\psi^p{}_p = 0.50$
$k = 0.9$						
$k = 0.8$						
$k = 0.7$						

If the optimal contribution rate is positive, there is a role for benevolent social insurance.

Social insurance distorts the labor market, because unrealistic optimism induces some young individuals to choose uncovered jobs and independent saving. The average distortion in the labor market can be captured by the average density of contribution. The coverage rate and the labor distortion LD are defined as:

(A.3) $Cov(\theta) \equiv \iint D^{biased}(\theta) \cdot f(z^{ex}, b) dz^{ex} db \quad ; \quad LD \equiv -[Cov(\theta^*) - Cov(0)]$

Density and coverage are linked in a simple way when lives have only two periods.

Table 2: Equilibrium coverage and labor distortion with optimal social insurance (to be completed)

(The other parameter assumptions are indicated in this Appendix)

Cov, LD	$\rho^c = g_{CR}$ = 1% per annum (real)			$\rho^c = r$ = 3% per annum (real)		
	$\psi^p{}_p = 0.70$	$\psi^p{}_p = 0.60$	$\psi^p{}_p = 0.50$	$\psi^p{}_p = 0.70$	$\psi^p{}_p = 0.60$	$\psi^p{}_p = 0.50$
$k = 0.9$	XX, ZZ					
$k = 0.8$						
$k = 0.7$						

Social insurance also distorts the savings market, because unrealistic optimism induces some young individuals to substitute the forced saving bundled to coverage jobs for a reduced amount of independent saving. The average saving rate adds independent and covered saving. This rate and the saving distortion SD are defined as:

(A.4)

$$Sav(\theta) \equiv \iint \left[F^{biased} + \theta \cdot w \cdot x_a^c \cdot (1 - \hat{l}_a) \cdot D^{biased}\right] \cdot f(z^{ex}, b) dz^{ex} db \quad ; \quad SD \equiv -[Sav(\theta^*) - Sav(0)]$$

Table 3: Equilibrium saving rate and savings distortion with optimal social insurance (to be completed)

(The other parameter assumptions are indicated in this Appendix)

Sav, SD	$\rho^c = g_{CR}$ = 1% per annum (real)			$\rho^c = r$ = 3% per annum (real)		
	$\psi^p{}_p = 0.70$	$\psi^p{}_p = 0.60$	$\psi^p{}_p = 0.50$	$\psi^p{}_p = 0.70$	$\psi^p{}_p = 0.60$	$\psi^p{}_p = 0.50$
$k = 0.9$	XX, ZZ					
$k = 0.8$						
$k = 0.7$						