NBER WORKING PAPER SERIES

INDUCED RETIREMENT, SOCIAL SECURITY, AND THE PYRAMID MIRAGE

Casey B. Mulligan

Working Paper 7679
http://www.nber.org/papers/w7679

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2000

I appreciate the comments of Chicago GSB Macro lunch eaters, Mark Duggan, Kevin Murphy, Tomas Philipson, Xavier Sala-I-Martin, Jaume Ventura, Chicago and Iowa State seminar participants, the research assistance of Shiqiang Zhan, and the financial support the National Science Foundation. The views expressed herein are those of the author and are not necessarily those of the National Bureau of Economic Research.

Induced Retirement, Social Security, and the Pyramid Mirage
Casey B. Mulligan
NBER Working Paper No. 7679
April 2000
JEL No. J26, D78, D91, H55, D72

ABSTRACT

Does Social Security redistribute across cohorts? Or is it a program for "purchasing the jobs" of the elderly? I formalize both models, showing how they have some predictions in common – the most important of which is that generational accounts have the appearance of a "pyramid scheme." I also derive important differences between the two interpretations, and compare those differences with data on the design and incidence of Social Security programs around the world. Since implicit and explicit tax rates on elderly labor income are so high, and so closely (and positively) related with the amount of Social Security spending, and because substitution effects of the program can be as large as its wealth effects, I conclude that Social Security's induced retirement motive is much more important for explaining differences among European countries than is the intergenerational redistribution motive. Furthermore, when policy at least in part designed to induce retirement, its generational incidence can be very different than the incidence of a pyramid scheme, even for those countries where the induced retirement motive is not the dominant one. The possibility of induced retirement also makes it difficult for perpetual intergenerational redistribution to be supported as a subgame perfect political equilibrium.

Casey B. Mulligan
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
and NBER
c-mulligan@uchicago.edu

Table of Contents

Social Security has for a long time been perceived as a vehicle for intergenerational redistribution (eg., Samuelson 1958, Barro 1974, Browning 1975, Kotlikoff 1992, Barro 1987, Becker and Murphy 1988, Murphy and Welch 1998, Cooley and Soares 1999). The perception has naturally led analysis to pose questions such as "What is the extent of intergenerational redistribution?", "For how long can cohorts enjoy fair returns from Social Security?", "Might intergenerational redistribution be understood as a political equilibrium?"

The intense focus on intergenerational redistribution by Social Security has distracted from what may be the main intent of Social Security – to induce the elderly to exit the labor market. Here I show how retirement-inducing policies are different, in terms of program incentives and policy incidence, from intergenerationally redistributive policies. I show how policy might achieve both objectives and, more importantly, how policy that is designed *both* to help the elderly *and* to induce retirement is very different from a policy designed only to help the elderly.

The predictions of the theory are compared with evidence, mainly from OECD countries, on the design and incidence of Social Security, and government policy more generally. On the whole, the evidence suggests that policy is designed to both redistribute and to induce retirement, and the policies are both qualitatively and quantitatively different from a (hypothetical) policy designed only to help the elderly. Within-European differences in program design and incidence can only be understood in terms of the retirement inducing motive. My estimates also suggest that Folk Theorem versions of the pyramid model (eg., Kotlikoff et al 1988, Bohn 1998, Cooley and Soares 1999 and, essentially, Browning 1975) cannot support as subgame perfect political equilibria the Social Security programs in Europe, or even programs elsewhere.

Since it is so important to understand, and so feasible to measure, the basic differences between policies designed to induce retirement and those designed to help the elderly, I do little in this paper to consider the important questions of *why* policy might have one or both motives.

Previous studies have offered some answers (eg., Cooley and Soares 1999, Mulligan and Sala-i-Martin 1999a,b,c, Sala-i-Martin 1996, Tabellini 1992), and have tried to derive differences *among* intergenerational-redistribution theories and differences *among* retirement-inducing theories (eg., Mulligan and Sala-i-Martin 1999b,c). Future work needs to improve on these answers, and can be usefully directed by the achievements of this paper: derivation of implications likely to be common to the intergenerational redistribution theories, derivation of implications likely to be common to the induced retirement theories, and comparison of those implications with government behavior around the world.

I. A Nested Model of Social Security

I.A. Setup of the Nested Model

Consider the familiar overlapping generations model, where each generation lives two periods and both young and old populate the economy in any given period. The generation born at time t has N_t members. The period t young and the period t old enjoy a time endowment of 1 and labor productivity of w_t^y and w_t^o, respectively. The time t government pays a lump sum transfer to one generation and finances that transfer with a labor income tax on one generation, or both generations.[1] I denote as $T_t \geq 0$ the net revenue paid by the period t the young to the period t old, as a proportion of elderly full private income w_t^o.[2] σ_t^i denotes the rate at which labor income of the age group i is taxed *at the margin* ($i = o,y$).

For simplicity, I assume that young individuals do not borrow or save. Old and young citizens choose consumption c and leisure l when young and old to maximize their utility functions $u(c^i,l^i) = \ln c^i + \gamma^i \ln l^i$ ($i = o,y$) where their consumption is constrained by their labor income net of the various taxes. Notice that, for simplicity, I assume utility functions are Cobb-Douglas.

[1]I presume that no individual can control his year of birth (or lie about it to the government) so that generation-specific transfers can indeed be lump sum. Generation-specific lump sum taxes are presumably more difficult to implement, since a taxpayer might refuse to work and thereby have no income that could be seized by the government.

[2]It will be the old who receive the net transfer in the versions of the model I consider, but later in the paper it will become obvious how to describe period t transfers to the young. For example, T_t would be negative, and the budget constraint for the old would have to be modified to rule out lump sum taxes on them.

The aggregate resource constraint at date t is:

$$N_t c_t^y + N_{t-1} c_t^o = N_t w_t^y (1 - l_t^y) + N_{t-1} w_t^o (1 - l_t^o)$$

with $l_t^i \in [0,1]$ and $c_t^i \geq 0$ ($i = o,y$). Policy at each date t, (T_t, σ_t^y, σ_t^o), is chosen to maximize a social welfare function:[3]

$$N_t\left(\ln c_t^y + \gamma^y \ln l_t^y\right) + \alpha_t N_{t-1}\left(\ln c_t^o + \gamma^o \ln l_t^o\right) + \beta_t (1 + \gamma^o) N_{t-1} \ln l_t^o \qquad \textbf{(1)}$$

where the terms in first and second parentheses are the utility of the young and old, respectively, and γ^o, γ^y are (positive) preference parameters. The policy objective need not equally weight the utility of young and old. The date t weight on old utility, $\alpha_t \geq 0$, can be interpreted as an index of the political power of the old, or an index of the young's altruism for the old, or some combination of these. The weight placed in the policy objective on *retirement* relative to the welfare of young and old varies according to the parameter $\beta_t \geq 0$.[4]

The policy objective (1) raises two important questions. First, what determines the magnitude of α and β? Second, to whom accrues the benefits indicated by the third term in the policy objective? As suggested above, α may be related to the political power of the old, or the degree of altruism felt by one generation for another. The literature has also suggested why governments may value retirement, and to whom the benefits of retirement accrue. One suggestion is that workers are more successful when retirement is common in the economy, and workers have a disproportionate influence

[3]The social welfare function could be defined over the remaining lifetime utility of everyone alive at the time without changing the results. One such case has intertemporally separable lifetime utility, no opportunity for aggregate saving, and future governments not otherwise reacting to policies of current and past governments. Other cases are analyzed in Appendix III.

[4]$\beta_t \geq 0$ is multiplied by the positive constant $(1+\gamma_o)$ in the third term merely to simplify the calculations. With this convention, the marginal social welfare of income for the old equals the marginal social welfare of retirement exactly when $\alpha_t = \beta_t$.

on policy decisions. Another may be that retirement promotes political stability, and policymakers value stability itself (over and above how it may benefit citizens).[5] Or retirement may otherwise enhance the political power of policymakers (this is a simplified version of Mulligan and Sala-i-Martin's 1999a argument). Or inclusion of retirement in the social welfare function may proxy for an external effect of retirement on the operation of labor markets accounted for by policymakers, as in the model of Sala-i-Martin (1996). Or β may reflect redistribution in a dimension other than age – from workers to retirees (perhaps this is another interpretation of Mulligan Sala-i-Martin 1999a). Nevertheless, much more research is needed to convincing explain why policy might weight heavily the utility of the old, or retirement.[6] But even in a model as abstract as (1), a number of different implications for the design and incidence of Social Security can be derived – implications shared with more detailed models of intergenerational transfers and induced retirement.[7]

The policy choice is subject to the aggregate resource constraint, and subject to the labor supplied by young and old citizens given the policy they face. In other words, policies solve a sequence of Ramsey-tax problems, which can be viewed in two stages for each period. In the second stage, young and old allocate their resources between consumption and leisure, taking as given the date t policy T_t, σ_t^y, σ_t^o. Second stage behavior is particularly interesting for the young

$$u_t^y = \max_{c_t^y, l_t^y} \ln c_t^y + \gamma^y \ln l_t^y$$

$$\text{s.t.} \quad c_t^y = w_t^y (1 - l_t^y)(1 - \sigma_t^y) \quad , \quad l_t^y \in [0,1]$$

The maximum attained in the second stage for the young is, up to a constant depending only on γ^y,

---

[5]A related argument is made by Piven and Cloward (1971) and Olson (1982).

[6]Diamond and Mirrlees (1978), Mulligan (1999), Mulligan and Sala-i-Martin (1999a,b,c), Sala-i-Martin (1996), and Tabellini (1992) have done some research on the reasons for publicly induced retirement.

[7]In order to derive some of the main results, it is important that my formulation rules out the possibility that old age leisure raises the marginal social welfare of consumption by the young (in this case, the *old* pay taxes in order to induce retirement). In this regard, my formulation is consistent with the induced retirement models from the literature.

$\ln w_t^y + \ln(1\text{-}\sigma_t^y)$. Because of the logarithmic functional form, we also have the simple result that young leisure time is $\gamma^y/(1+\gamma^y)$.

The old face a similar second stage problem, but it is less interesting because the government has two policy variables to influence their behavior, namely the marginal tax rate and the amount of the lump sum transfer. Therefore, without loss of generality, I have the government choose elderly consumption and leisure in the first stage. Optimal policy can be described as the solution to a sequence of "planner's" problems (2):

$$\max_{c_t^o, l_t^o, T_t, \sigma_t^y, u_t^y} \quad e^{n_t} u_t^y + \alpha_t \left(\ln c_t^o + \gamma^o \ln l_t^o \right) + \beta_t (1+\gamma^o) \ln l_t^o$$

$$\text{s.t.} \quad c_t^o = w_t^o (1 - l_t^o) + w_t^o T_t \quad , \quad l_t^o \in [0,1] \tag{2}$$

$$u_t^y = \ln(1-\sigma_t^y) + \ln w_t^y \quad , \quad T_t = e^{n_t + z_t} \frac{\sigma_t^y}{1+\gamma^y}$$

where $z_t \equiv \ln \frac{w_t^y}{w_t^o}$ and $n_t \equiv \ln \frac{N_t}{N_{t-1}}$ denote amounts (in log points) by which the period t wage and cohort size of the young, respectively, exceeds that of the old. The first line of (2) is the social welfare function (1). The constraint on u_t^y in the problem (2) reflects date t's second stage, namely that the government cannot control the behavior (and ultimate utility) of the young, except through the single policy instrument σ_t^y. The government has two instruments to control the behavior of the old so, for simplicity, c_t^o and l_t^o are entered as choice variables for the government. Policy values T_t, and σ_t^o are easily computed from (c_t^o, l_t^o) as $T_t = c_t^o/w_t^o - (1\text{-}l_t^o)$ and $\sigma_t^o = 1 - \gamma_t c_t^o/(w_t^o l_t^o)$, while c_t^o, l_t^o, σ_t^y are derived from the first order conditions of (2). The optimal policy is:[8]

$$T_t = \frac{\beta_t + (\alpha_t - \alpha_t^*)}{(\beta_t + \alpha_t)(1+\gamma^y)\, e^{-(n_t+z_t)} + \alpha_t^*} \quad , \quad \sigma_t^y = \frac{\beta_t + (\alpha_t - \alpha_t^*)}{\beta_t + \alpha_t + \frac{e^{n_t}}{1+\gamma}} \quad , \quad \sigma_t = \frac{\beta_t}{\beta_t + \frac{\gamma}{1+\gamma}\alpha_t} \tag{3}$$

[8]Results reported in the main text are for those parameter values for which the constraints l_t^o, $l_t^y \le 1$ do not bind. Appendix I studies corner solutions, and displays the parameter restrictions required for interiority.

where the age group superscript is suppressed when it refers to the old, a convention which I continue for the remainder of the paper, unless a superscript is necessary for clarity. $\alpha_t^* = \frac{1+\gamma^y}{1+\gamma^o} e^{-z_t}$ is a function of the model parameters.

I.B. Intergenerational Redistribution and Induced Retirement as Special Cases

This model of social security includes three important models as special cases. The first has neither redistributive nor retirement-inducing motives:

<u>No Motives</u>: $\alpha_t = \alpha_t^*$, $\beta_t = 0$

The optimal policy abstains from both lump sum and distortionary taxes and transfers ($T_t = 0$, $\sigma_t^i = 0$, $i = o,y$). "No policy" is optimal because, roughly speaking, the policy objective weights each age group's utility according to its relative command of market resources e^{-zt} and its relative marginal utility of income $(1+\gamma^y)/(1+\gamma^o)$.

The second special case is the intergenerational redistribution (IGR) model:

<u>IGR model</u>: $\alpha_t \geq \alpha_t^*$, $\beta_t = 0$

Redistribution is the only policy motive in the IGR model. The optimal policy uses only a lump sum transfer (from young to old, $T_t \geq 0$)[9] – abstaining from distorting elderly labor supply ($\sigma_t^o = 0$) – and finances the transfer with a labor income tax on the young. This special case has a lot in common with many models in the literature viewing Social Security merely as a vehicle for intergenerational redistribution (eg., Samuelson 1958, Barro 1974, Browning 1975, Kotlikoff 1992, Barro 1987, Becker and Murphy 1988, Murphy and Welch 1998, Cooley and Soares 1999), not a policy designed to distort behavior.

The IGR special case can also "explain" the historical emergence of policies that redistribute

[9]With $\alpha_t < \alpha_t^*$ and $\beta_t = 0$ we have another intergenerational redistribution model, but perhaps one less interesting to the student of Social Security because redistribution is from old to young. In this case, (2) might be reformulated so that transfers to the young could be lump sum, and financed with a labor income tax on the old.

resources transferred across generations with a time sequence of α_t's. Namely, the IGR model with $\alpha_t = \alpha_t^*$ for $t < 0$ and $\alpha_t > \alpha_t^*$ for $t \geq 0$ has pay-as-you-go "social security" emerging at $t = 0$, because $T_t = 0$ for $t < 0$ and $T_t > 0$ for $t \geq 0$. Whether policy in the IGR model accurately reflects actual Social Security programs is another question, to which I return below.

The third special case is the induced retirement (INR) model:

$$\underline{\text{INR model}}:\ \alpha_t = \alpha_t^*,\ \beta_t \geq 0$$

Inducing retirement is the only policy motive in the INR model. The optimal policy distorts the labor supply of the old ($\sigma_t^o > 0$) in those periods when $\beta_t > 0$. The optimal policy also subsidizes the old ($T_t \geq 0$), in a sense, as compensation to the old for tolerating deviation from their preferred labor supply. The labor supply of the young is distorted whenever $\beta_t > 0$ as a necessary consequence of raising revenue, but the young marginal tax rate is still less than that for the old ($\sigma_t^y < \sigma_t^o$) in those periods.

The INR special case can also "explain" the historical emergence of policies that redistribute resources across generations with a time sequence of β_t's. Namely, the INR model with $\beta_t = 0$ for $t < 0$ and $\beta_t > 0$ for $t \geq 0$ has "social security" emerging at $t = 0$, because $T_t = 0$ for $t < 0$ and $T_t > 0$ for $t \geq 0$.

That Social Security is primarily designed to induce retirement has been discussed in policy circles (eg., the references cited by Sala-i-Martin 1996, Gruber and Wise 1999, p. 31). Inducing retirement is also sometimes cited as a motive for private pensions (eg., the instances cited by Clague et al 1971, p. 78) but, with rare exceptions (eg., Sala-i-Martin 1996, Mulligan and Sala-i-Martin 1999a,b,c), the induced retirement model of Social Security has neither been formalized nor systematically compared with the IGR model. My sections II and III do so.

II. Equivalent Implications of Redistributional and Retirement-Inducing Motives

Two basic observations have motivated the study of intergenerationally redistributive models of public policy. First, Social Security, and government policy as a whole, transfers revenues from young to old (eg., Auerbach et al 1992, Auerbach et al 1999, House Committee 1996 table 1-50). Second, the old appear to support the continuation and growth of Social Security (eg., see the AARP's web page www.aarp.org) and other related programs. I show how neither of these observations can distinguish the IGR model from the INR model or any of the hybrid cases with $\alpha_t > \alpha_t^*$ and $\beta_t > 0$.

II.A. Generational Accounting

Although IGR models are often motivated by the observation that governments transfer resources from young to old, such observations are also consistent with the INR model and all of the cases in between. This can be seen in the formula (3) for the optimal date *t* redistribution T_t in the nested model, a formula duplicated below for the reader's convenience:

$$T_t = \frac{\beta_t + (\alpha_t - \alpha_t^*)}{(\beta_t + \alpha_t)(1 + \gamma^y)e^{-(n_t + z_t)} + \alpha_t^*}$$

Notice how the redistributive motive (α_t - α_t^*) is a perfect substitute with the retirement-inducing motive β_t in terms of the amount of resources transfers from young to old, T_t. In other words, any sequence of intergenerational redistribution generated by the IGR model can be generated by the INR model (or any of the hybrid cases with $\alpha_t > \alpha_t^*$ and $\beta_t > 0$), and vice versa.

The two motives also have very similar implications for the so-called "generational accounts" studied by Kotlikoff (1992), Hagemann and John (1990), and many others. Date-of-birth generational accounts compute the lifetime present value of net transfers v_t for each generation, indexed by their date of birth *t*. In my notation, the date-of-birth generational accounts are:

$$v_t = -w_t^o T_t e^{-n_t} + e^{-r_{t+1}} w_{t+1}^o T_{t+1} = w_t^o T_t e^{-n_t}\left(e^{g_{t+1} - r_{t+1}} - 1\right) \quad \textbf{(4)}$$

where r_{t+1} is the generational interest rate and $g_{t+1} \equiv \ln \frac{w_{t+1}^o T_{t+1} N_t}{w_t^o T_t N_{t-1}}$ is the aggregate growth rate (in log points) of the Social Security program from date *t* to date *t*+1. Remaining lifetime generational accounts – namely, the lifetime net present value of all current and future taxes and transfers – are also easily computed in my nested model for those alive at date *t*. As of time *t*, v_t (as computed above) is the remaining lifetime generational account for the date *t* young and T_t is the remaining lifetime account for the date t old.[10]

Three conclusions can be derived from the expression (4) for generation t's date-of-birth generational account v_t. First, since the redistributive and retirement-inducing motives are perfect substitutes in terms of T_t, they are perfect substitutes in terms of v_t – any set of generational accounts generated by the IGR model can be generated by the INR model (or any of the hybrid cases), and vice versa.

A second conclusion is more familiar – a generation t a positive date-of-birth generational account if and only if the interest rate r_{t+1} is less than the generational growth rate g_{t+1}. In other words, positive date-of-birth accounts are enjoyed by transitional generations, who age at a time when Social Security emerges or grows sufficiently rapidly.[11] What is novel about this result is that the emergence of Social Security, and its early generations with positive date-of-birth generational accounts, might occur even without any redistributive motive (ie, with $\alpha_t = \alpha_t^*$ for all t).

Third, the remaining lifetime accounts for the old are positive in any period t in both the IGR and INR models, and in any hybrid case with $\alpha_t > \alpha_t^*$ and $\beta_t > 0$. The first six rows of Table 1 record this and other common implications of the IGR and INR models.

[10]Remaining lifetime accounts are typically computed for studies of the impact of policy reforms (as in the Auerbach et al 1999 studies), while date-of-birth accounts are computed for "money's worth" studies (eg., Leimer 1994, Geanakoplos et al 1998).

[11]In a "steady state" (ie, $\alpha_t = \alpha_{t+1}$ and $\beta_t = \beta_{t+1}$), positive accounts are enjoyed only if population and wage growth rates exceed the interest rate.

Table 1: Common and Uncommon Implications of Redistributive and Retirement-Inducing Motives

	IGR	INR
old age subsidies financed with contemporaneous taxes on the young	yes	yes
SS causes the elderly to work less	yes	yes
the elderly have relatively low private incomes	yes	yes
remaining lifetime generational accounts positive for the old	yes	yes
date-of-birth generational accounts positive only when SS growing	yes	yes
the old are better off with the continuation and/or expansion of SS	yes	yes*
only initial generations gain from SS	yes	no
elderly marginal tax rates are positive	no†	yes
elderly marg tax rates increase significantly with the amount of redistribution	no	yes
elderly both pay taxes and receive benefits (aka, "cross-hauling")	no	yes
policy motive index exceeds 0.5	no	yes
distortions exceed redistribution	no	yes
SS causes the elderly to consume less	no	yes

*assumes β not too large

†beneficiary behavior is distorted in some models of redistribution (eg., those in Appendix II)

II.B. "Social Security Alleviates Elderly Poverty"

Both models are consistent with the finding that the elderly have lower private incomes than do the young, and that the size of the Social Security program would increase with the age-income gap.[12] In the IGR model, low private incomes for the elderly might be modeled as a low α_t^* (remember that α_t^* is proportional to the relative private full income of the old), which motivates the redistribution across cohorts. In contrast, low private incomes for the elderly in the INR model are partly a consequence of policy (as long as labor supply is somewhat elastic), and redistribution across

[12]Whether in these findings are true is another matter. See, for example, Gratton (1996).

cohorts is a means of lowering elderly private incomes.

II.C. Generational Incidence I

Generational incidence, the distribution of policy-induced utility gains and losses across generations, is conceptually distinct from generational accounting, the distribution of taxes and subsidies across generations. But they are isomorphic in the IGR model – generations with positive date-of-birth accounts are better off than they are in the absence of policy and generations with negative date-of-birth accounts are worse off.[13] Ignoring the possibility that steady state growth might exceed the interest rate, date-of-birth generational accounts are positive only if Social Security is growing, so the generations gaining in the IGR model are only those living while the program emerges or is growing. The remaining lifetime generational accounts also make it clear that, even for a mature Social Security system (ie, $\alpha_t > 0$ and stable over time), the date t old are worse off in the IGR model if Social Security were immediately eliminated. The generational incidence in the IGR model suggests an explanation why the old would support the continuation and expansion of Social Security.

Generational accounts and generational incidence are not isomorphic in the INR model. Nevertheless, the INR model can explain why the old would support the continuation and expansion of Social Security because, for small β_t enough, the remaining lifetime *utility* of the date t old ($\ln c_t^o + \gamma^o \ln l_t^o$) increases with β_t. Roughly speaking, the old in the INR model willingly sell their jobs to the planner and an elimination of Social Security would mean that they would lose the opportunity to sell their jobs.

---

[13]At least for small policies. Since the interest rate relates to the *marginal* willingness to substitute consumption over the life cycle, generational accounting in the intergenerational redistribution model indicates the willingness to pay for a marginal increase in taxes and transfers rather than the willingness to pay to avoid the complete elimination of the program. The tax on the young is distortionary, with a dollar in taxes costing the young more than a dollar, and is thereby the source of another difference between generational accounts and generational incidence unless policies are sufficiently small.

My analysis does not focus on either of these conceptional difficulties with generational accounting, since they are not related to the differences between the "intergenerational redistribution" and "induced retirement" models. Fehr and Kotlikoff (1999, p. 49f) also suggest that the distinction between marginal and average willingness to pay for Social Security is not of practical significance.

The young may also favor Social Security in the INR model, and date-of-birth generational accounts tell little or nothing about the generational incidence of Social Security. I return to this important difference between the IGR and INR models below.

III. Unique Implications of Redistributional and Retirement-Inducing Motives

Both the IGR and INR models predict that the young will be taxed to subsidize the old, and both are consistent with the old's being better off under a large Social Security program rather than a smaller one. But there are a number of differences between the two models. One difference is the very weak correspondence between generational incidence and accounts in the INR model. The rest of the differences, roughly speaking, can be attributed to the importance of income effects in the IGR model and substitution effects in the INR model. These differences between the IGR and INR models are derived below and recorded in the bottom six rows of Table 1.

III.A. Generational Incidence II

Ignoring the possibility that the steady state growth might exceed the interest rate, date-of-birth generational accounts in both models are positive only if Social Security is growing. Hence policy has the *appearance* of a "chain letter" or "pyramid scheme" because, like a pyramid scheme, policy has only those at the peak of the pyramid gaining *revenue* at the expense of those at the broader base.

Policy is indeed a chain letter in the IGR model, because coming out ahead in terms of generational accounts is, with the caveats mentioned above, the same as coming out ahead in terms of utility. But the appearance of a chain letter is deceiving in the INR model because, for two reasons, the relationship between generational accounting and incidence is weak. First, the amount of the subsidy T_t paid to the old is not monotonically related with the welfare of the old. To see this consider increasing β_t, starting from zero and holding constant α at all dates and β at all dates other than t. The amount of the date t subsidy from young to old, T_t, unambiguously increases with β_t as shown by my formula (3). The utility of the old u_t^o (= ln $c_t^o + \gamma^o$ ln l_t^o) increases with β_t at first but, for some parameter values, later decreases with β_t and can be less than old utility when $\beta_t = 0$![14] The

[14]Since governments often do not force people to retire, leaving the retirement decision up to employers and employees, perhaps continued work (even without benefits) is a feasible choice

reason for the nonmonotonic relationship between T_t and u_t^o shown as a dashed blue curve in Figure 1 is that the old are simultaneously subsidized and forced to change their behavior in the direction of less consumption and more leisure, and the latter effect becomes relatively large as the old's marginal utility of consumption grows with β_t.[15] Even if the old are better off with Social Security, their welfare improvement can be a small fraction of what it would be if their subsidy T_t were lump sum. It is also possible and, as the evidence in Section IV shows, likely that the old are not as well off under the largest Social Security programs as they are under medium size programs because the larger programs do so much to discourage work. This possibility is seen in Figure 1 for the INR model where the utility of the elderly is maximized under the program transferring T_t^{max}; programs with $T_t > T_t^{max}$ are worse for the elderly even if they are better than no Social Security at all.

for the elderly so they are no worse off than they would if Social Security were eliminated? This logic is not correct (although the conclusion may be) because, even when it is feasible to continue work and forego Social Security benefits, most governments still require elderly workers to pay Social Security taxes.

Downward sloping schedules in the $[T_t, u_t]$ plane are more likely when population and wage growth are low, and when the elderly preference for leisure is relatively small.

[15]Holding fixed α_t, there is a maximum β_t beyond which additional increases in β_t have no effect on policy because the constraint $l_t \leq 1$ binds. The amount of redistribution for this case is denoted T_t^{corner} in the Figure. T_t^{corner} is computed by evaluating the Appendix I formula for T_t at $\alpha_t = \alpha_t^*$.

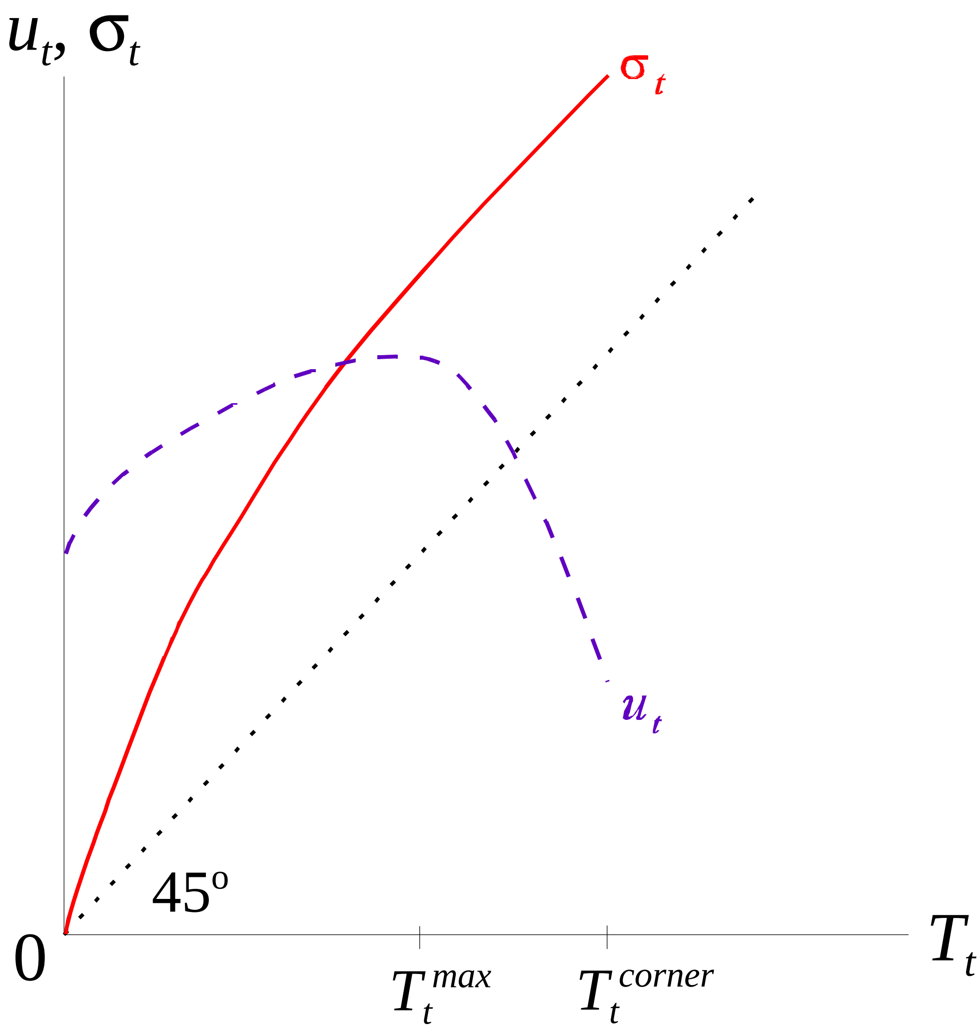


Figure 1 INR Model: Elderly Utility and Marginal Tax Rate, as Functions of the Size of the Social Security Program

Second, the young in the INR model may be better off under Social Security, if they enjoy at enough of the benefits modeled by the β_t term in the policy objective (1). For example, Sala-i-Martin (1996) suggests that retirement improves productivity for the young, and all generations in his model are better off from retirement-inducing Social Security. Others have suggested that retirement creates jobs for the young.

Generational accounts can be revised so that they more accurately reflect generational incidence. Consider a set of revised date-of-birth accounts v_t^*:

$$v_t^* = -w_t^o E_t^y e^{-n_t} + e^{-r_{t+1}} w_{t+1}^o E_{t+1}^o \equiv w_t^o E_t^y e^{-n_t}\left(e^{g_{t+1} - r_{t+1} - \varepsilon_{t+1} + \lambda_t} - 1\right)$$

There are two differences between my revised generational accounts and those computed in the literature. The first difference regards the term E^o_{t+1}, which is a "compensating variation" – the amount by which the date (t+1) old value the Social Security program at date (t+1), taking into account that their benefit is not a lump sum transfer.[16] $\varepsilon_{t+1} \equiv \ln (T_{t+1}/E^o_{t+1})$ is a convenient transformation of the gap between the budgetary cost of old age subsidies and their valuation of them, and is nonnegative because the (t+1) old value Social Security in the amount T_{t+1} the most when it is a lump sum. $\varepsilon_{t+1} = 0$ when $\sigma_{t+1} = 0$, and otherwise depends on σ_{t+1}, T_{t+1}, and the shape of elderly utility functions. ε_{t+1} is easy to compute when the utility of the old is $\ln c_t^o + \gamma \ln l_t^o$:

$$e^{-\varepsilon_t} = \frac{(1+T_t)\left(1 - \frac{\sigma_t}{1+\gamma}\right)^{-1}(1-\sigma_t)^{1/(1+\gamma)} - 1}{T_t} \qquad \textbf{(5)}$$

E_t^y is the second departure of my revised generational accounts and those computed in the literature. It is also "compensating variation" – the cost to the young of the Social Security program at date t, taking into account that taxes may change their behavior and they may benefit from the retirement of the date t old.[17] $\lambda_t \equiv \ln (T_t/E_t^y)$ is a convenient transformation of the gap between taxes paid by the young and the cost to them of the program. λ_t is positive as long the young derive enough benefit from retirement by the old. If the young derive no benefit from retirement by the old, then λ_t is negative, with its magnitude determined by σ_t^y, because the young would prefer to pay a lump sum tax rather than a distortionary one. The young's benefit from retirement depends on the third term in the policy objective (1), and the fraction of that term that accrue to the young, rather than to the old or someone else influencing policy. Hence there are three reasons why λ_t is more difficult to compute than ε_{t+1}. First, the magnitude of third term depends not only on σ_{t+1}, T_{t+1}, and

[16]In other words, E^o_{t+1} is the size of the lump sum benefit that would give the (t+1) old the same utility as they have under the program, expressed as a fraction of their full private income.

[17]More precisely, E_t^y is the lump sum amount the date t young would forego in the absence of the program in order to achieve the same utility as they do under the program.

the shape of elderly utility functions, but also on the magnitude of β_t, z_t, and n_t. Second, neither theory nor evidence gathered to date tells us much about the share of those benefits accruing to the young. Third, even if these first two items were known, their value to the young depends on the shape of their utility function.

Comparing the revised accounts with those (4) computed in the literature, we see that old age benefits are effectively discounted at rate ($r_{t+1} + \varepsilon_{t+1} - \lambda_t$) rather than r_{t+1}. Except in the IGR model where transfers are lump sum and retirement is not valued *per se*, this rate is higher than r_{t+1} for those retiring under a new or growing Social Security system because induced retirement was relatively unimportant (or nonexistent) when that generation was young (hence, $\varepsilon_{t+1} > \lambda_t$). In other words, my revised accounts suggest that benefits are smaller, or even negative, for the early participants in a pay-as-you-go Social Security system. For those participating in a mature system, it could be that $\varepsilon_{t+1} < \lambda_t$; old age benefits are effectively discounted at less than the interest rate. This means that generations participating in a mature system could benefit, in the lifetime sense, from the system even though when interest r_{t+1} exceeds the rate of growth of the program. I show in Section IV how the revised generational accounts are quite different in magnitude from those computed in the literature.

Except when the program growth rate exceeds the interest rate, even the usual generational remaining-lifetime accounts (4) show why the young would oppose, and the old favor, the continuation of Social Security. Generational accounts cannot be computed in my model for the "middle aged" because it has only two periods, but it is well known (eg., Browning 1975, Kotlikoff et al 1988) that, according to the usual generational accounts, the middle aged strongly favor a *temporary* reduction in Social Security but may weakly oppose a *permanent* reduction. Browning (1975), and students of the Folk Theorem for repeated games (eg., Kotlikoff et al 1988, Bohn 1998, Cooley and Soares 1999) have exploited the middle-aged opposition to permanent reductions to argue that perpetual intergenerational redistribution can be a subgame perfect equilibrium of an infinitely repeated political game determining the sequence of coefficients $\{\alpha_t\}$. However, all of this presumes that benefits are valued at their budgetary cost (as, for example, in Bohn 1998) so that the usual generational accounts (4) are closely related to generational incidence. The fact that the transfer to the elderly is not lump sum means that they do not value benefits at their budgetary cost; future benefits are effectively discounted at rate ($r_{t+1} + \varepsilon_{t+1}$), rather than r_{t+1}. Even a small revision can mean that the middle aged favor a permanent reduction is Social Security. More importantly, ε_{t+1}

needs to be large enough for *only one generation*, even if that generation is very far in the future, for the young and middle aged to form a winning coalition eliminating Social Security and eliminating as subgame perfect equilibria intergenerational redistribution in prior periods – even periods when the middle aged favor the continuation of Social Security.

III.B. Marginal Tax Rates

While the different predictions of the models for generational incidence are important, they are difficult to verify. A more easily verified implication is for marginal tax rates, and their relationship with the size of the Social Security program. The IGR predicts that marginal tax rates are zero, while the INR model predicts a positive marginal tax rate for the old, $\sigma_t = \frac{\beta_t}{\beta_t + \frac{\gamma}{1+\gamma}\alpha_t^*}$.[18]

For the hybrid models, the magnitude of elderly marginal tax rates, σ_t, is an indicator of the absolute importance of the retirement inducing motive and, to some degree, an indicator of its importance relative to the redistributive motive. Recall from the formula (3) that σ_t depends only on γ and the ratio β_t/α_t. Hence, large elderly marginal tax rates indicate a strong retirement-inducing motive and, holding constant α_t^*, a relatively weak redistributive motive. Large elderly marginal tax rates are consistent with strong retirement inducing motives *and* strong redistributive motives if α_t^* is small, say, because the elderly would have relatively low private incomes in the absence of Social Security.

The INR model not only predicts a positive marginal tax rate on the old, but also a positive correlation between the magnitude of the marginal tax rate and the size of the Social Security program. This can be seen by increasing β_t, starting from zero and holding constant α at all dates and β at all dates other than t, and comparing T_t with σ_t. σ_t varies monotonically with T_t according to:

$$\sigma_t = \frac{(1+\gamma)T_t\left[1+(1+\gamma^y)e^{-(n_t+z_t)}\right]}{T_t\left[1+\gamma+(1+\gamma^y)e^{-(n_t+z_t)}\right]+\gamma} > T_t \tag{6}$$

[18]Notice that the INR model and the IGR model have the same number of free parameters, because the INR model sets $\alpha_t = \alpha_t^*$ for all t while the IGR model sets $\beta_t = 0$ for all t. Hence there are facts that would be consistent with the IGR model but not the INR model, and vice versa.

This function is graphed as a solid red curve in Figures 1 and 2. Even in the hybrid models, σ_t varies monotonically with T_t. The main difference in this regard between the hybrid models and the INR model is that σ_t does not exceed T_t unless β_t is sufficiently large (roughly speaking, the solid red curve shifts down and to the right as α_t exceeds α_t^*, as shown by the solid blue curve in Figure 2).

Also notice from equation (6) that the slope of the relationship between σ_t and T_t can exceed one. As shown below, such a steep slope cannot be derived from the IGR model, or from hybrid models with a dominant redistributive motive.

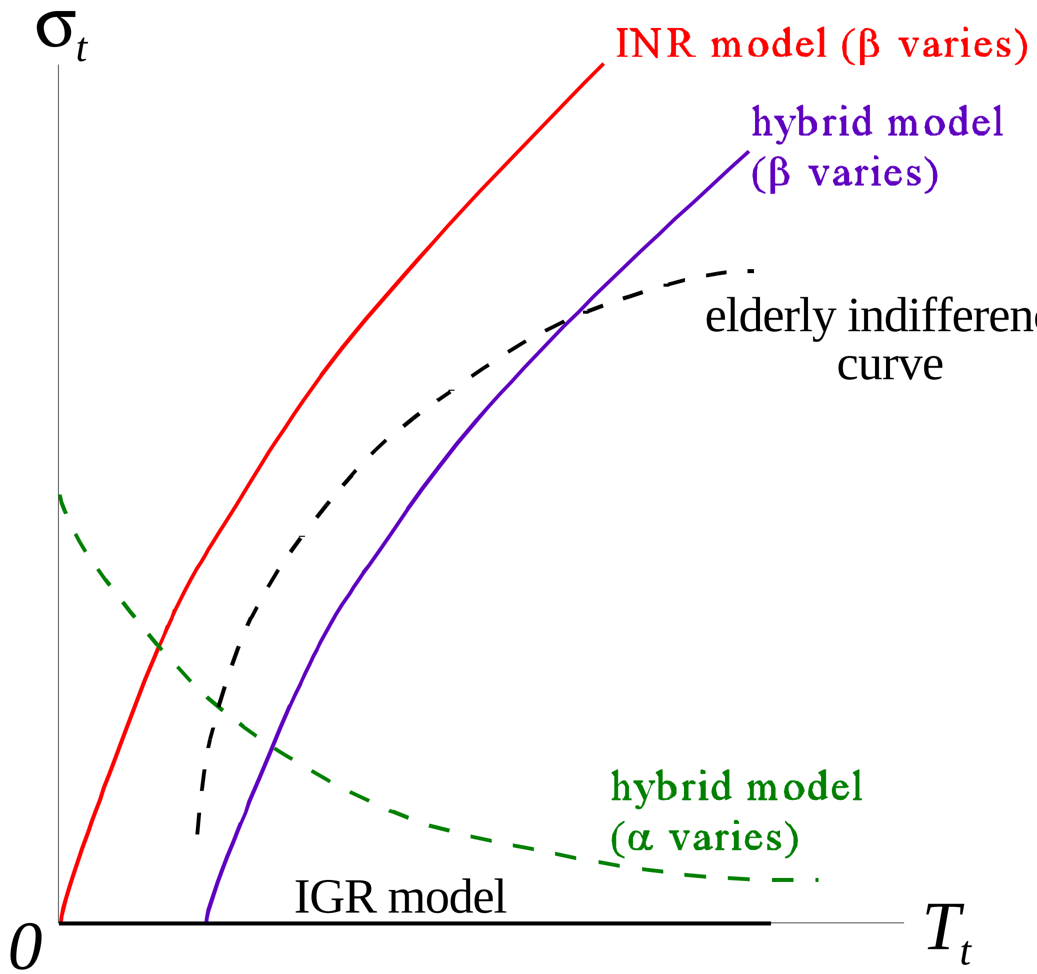


Figure 2 Elderly Marginal Tax Rate vs Program Size in Three Models

Equation (6) shows that σ_t and T_t are positively correlated in the INR model. (6) also suggests that σ_t is a concave function of T_t, but the concavity of the relationship is due only to the logarithmic functional forms for the utility and social welfare function. If cost of taxing the young increased sufficiently rapidly – perhaps because their marginal utility increases more rapidly than it does with log utility – T_t would increase more and more slowly relative to σ_t as the induced retirement motive became stronger.

Variation in the size of the Social Security program T_t is, in the INR model, attributed to differences in the strength of the retirement-inducing motive β_t. There is no such motive in the IGR model, so the same variation must be explained by differences in the redistributive motive α_t. Since the marginal tax rate is zero in all cases of the IGR model (and Figure 2 is labeled accordingly), regardless of α_t, it trivially follows that the relationship between σ_t and T_t is not increasing. Perhaps the hybrid models are more interesting in this regard, where we can study the relationship between σ_t and T_t induced by differences in the redistributive motive α_t holding constant the strength of the retirement-inducing motive β_t. In other words, consider increasing α_t, starting from α_t^* and holding constant β at all dates and α at all dates other than t. The induced relationship between σ_t and T_t is:

$$\sigma_t = \frac{(1+\gamma)\beta_t\left[1 - T_t(1+\gamma^y)e^{-(n_t+z_t)}\right]}{\beta_t + \gamma\alpha_t^* + T_t\left[\gamma\alpha_t^* - \beta_t(1+\gamma^y)e^{-(n_t+z_t)}\right]}$$

which is decreasing in T_t as shown by the downward sloping dashed green line in Figure 2.

Perhaps the IGR model is oversimplified because it does not allow any distortion of the old's behavior as a response to the redistributive motive. My Appendix II, and other work (Mulligan 2000, Mulligan and Sala-i-Martin 1999c), considers ways of embellishing the IGR model to allow for some distortion, but two conclusions are robust to those and other likely modifications of the IGR model. First, the marginal tax rate will not increase too rapidly with the size of the program because, holding constant the amount of redistribution T_t, the utility of the old falls with the marginal tax rate σ_t. In other words, there is a tradeoff between T_t and σ_t from the date t old's point of view; an elderly indifference curve such as the dashed curve in Figure 2 slopes up in the $[T_t,\sigma_t]$ plane. If marginal tax rates increase more rapidly with size of the program than indicated by indifference curves for the elderly, then the bigger programs make the old worse off and cannot be explained by the redistributive motive.[19] As shown above, a relationship between σ_t and T_t with slope greater than one indicates that the retirement-inducing motive dominates.

Second, redistributive motives cannot explain the prevalence of "cross-hauling." Third, income effects must be more important than substitution effects in the IGR model, and less important in the INR model. We see the third conclusion in an heuristic way in the formula (6) where the "substitution effect" σ_t exceeds the "income effect" T_t. The next two subsections derive the second and third conclusions more rigorously.

Fourth, it would seem that, if transfers and distortions could be targeted at a subset of the old, then the targeting would be very different in the IGR and INR model. In particular, an IGR planner would target transfers, and the associated unavoidable marginal tax rates, toward the least responsive old because, in effect, it is cheaper to raise their utility while optimal INR marginal tax rates would not vary among the old according to their responsiveness. This fourth point is just a conjecture, and cannot be proven in my model since the old are homogeneous. But it is testable, so I leave it to future research to explore the different implications of redistributive and retirement-inducing motives for the targeting of pensions.

III.C. Cross-Hauling

Those subsidized by the government are often taxpayers too, a phenomenon known as "cross-hauling." Cross-hauling occurs with old age subsidy programs too, since the old pay some taxes and, when they work, they pay Social Security taxes. Of course, when taxes and subsidies are lump sum, most economic approaches suggest that what matters is the *net* tax or subsidy, and not how any net tax is composed of taxes and subsidies. But, when their incentive structures are different, the composition of taxes and subsidies paid and received by an individual or group matter for behavior

---

[19]With $u^o(c,l) = \ln c + \gamma \ln l$, the old's indirect utility as a function is T and σ is:

$$u^o = \ln(1-\sigma) - (1+\gamma)\ln(1+\gamma-\sigma) + (1+\gamma)\ln(1+T) + \text{constants}$$

which implies that redistributive motives cannot cause σ to increase with T more rapidly than $(1-\sigma)(1+\gamma-\sigma)(1+\gamma)/[\sigma\gamma(1+T)]$, an upper bound which is smallest for large T and large σ.

and welfare.

Since policy is designed to improve the welfare of the elderly in the IGR model, the elderly are subsidized and their optimal marginal tax rate is zero ($\sigma_t = 0$). Hence the IGR model cannot explain why the old would also pay the same (distortionary) labor income tax as the young, or any labor income tax for that matter, unless the old were also paid a subsidy with a marginal tax rate *negative* enough to render the net transfer lump sum. Nor can the hybrid models with relatively small β_t explain why the old would also pay the same labor income tax as the young, because the optimal marginal tax rate is smaller for the old in those models. The INR model, and hybrid models with relatively large β_t, have larger optimal marginal tax rates for the old, so they are perfectly consistent with the old both paying a labor income tax at rate σ_t^y and receiving a subsidy that implicitly taxes their labor income, bringing the combined marginal tax rate to ($\sigma_t > \sigma_t^y$).

It might be argued that exempting old people from Social Security taxes would be administratively complex, but there are reasons to be skeptical of such a claim. First, age is easily verifiable, and is already used to determine personal income tax liabilities (eg., IRS *1999 Form 1040*, line 35a). Indeed, the American *Form 1040* already includes a line (62) for refunding excessive Social Security Tax withholding, and it would be trivial to modify the rules so that all some or all of the Social Security taxes paid by those over age 65 were defined to be "excessive." Nor has administrative complexity stopped state and local governments from exempting the elderly from some of their property taxes.

III.D. Importance of Income and Substitution Effects

Policy in the IGR model is motivated to improve the utility of the old, and hence optimal transfers to the old have only an "income effect" on them. Policy in the INR model is motivated to change the behavior of the old, and hence has a substantial "substitution effect" on their behavior. "Income" and "substitution" effect can be defined more precisely, but the basic idea that substitution effects are dominant in the INR model and income effects are dominant in the IGR model is the essence of my analysis and should be robust to modifications of those models.

The relative importance of income and substitution effects can also be used to classify the hybrid models where there are both redistributive and retirement-inducing motives. Such a classification is likely to be very useful in applications, since it is reasonable to believe that there is some role for both redistributive and retirement-induce motives in understanding Social Security. The question then is the relative quantitative importance of the two motives. Consider first a policy motive index ρ_t which takes values on the unit interval [0,1] and, heuristically, is computed as:

$$\rho_t \equiv \frac{\text{substitution effect}_t}{\text{substitution effect}_t + \text{income effect}_t}$$

More precisely, consider an index calculated as:

$$\rho_t = \frac{\hat{c}_t - c_t}{(\hat{c}_t - c_t) + (\hat{c}_t - c_t^*)}$$

$$\hat{c}_t \equiv \underset{c}{argmax}\ u^o\left(c, 1 + T_t - \frac{c}{w_t^o}\right)$$

$$c_t^* \equiv \underset{c}{argmax}\ u^o\left(c, 1 - \frac{c}{w_t^o}\right)$$

In other words, $\hat{c}_t$ - c_t is the amount by which old consumption is reduced by the optimal policy as compared to what would be consumed if the old were given their subsidy as a lump sum, and can be thought of as a substitution effect.[20] The income effect on consumption is the difference $\hat{c}_t$ - c_t^* between what the old would consume if given their subsidy as a lump sum and what they consume if they received no subsidy and faced a zero marginal tax rate.

ρ_t measures the importance of the retirement-inducing motive relative to the redistributive motive. Two critical values of the index are of particular interest. The first is $\rho_t = 0$, which occurs if and only if $c_t = \hat{c}_t$. In other words, policy has only an income effect and not a substitution effect – as it does in the IGR model where there is no retirement-inducing motive. $\rho_t = 0.5$ is another critical value which occurs when the substitution effect exactly equals the wealth effect (ie, $c_t = c_t^*$). $\rho_t = 0.5$ is perhaps the most interesting one because it can be used to partition the hybrid cases: $\rho_t >$ 0.5 ($\rho_t < 0.5$) indicates that the substitution effect is larger (smaller), so that the retirement-inducing

[20]By definition, $\hat{c}_t - c_t = w_t(l_t - \hat{l}_t)$, so the "substitution effect" on consumption is the same as the substitution effect on earnings.

motive dominates (is dominated by) the redistributive motive. ρ_t exceeds 0.5 in the INR model.[21]

The test of whether ρ_t exceeds 0.5 has several other interesting, and intuitive, interpretations. The following "tests" are equivalent:

(i)	ρ_t	< (>)	0.5
(ii)	c_t^*	< (>)	c_t
(iii)	$w_t(l_t - l_t^*)$	< (>)	$w_t T_t$
(iv)	β_t	< (>)	$(\alpha_t - \alpha_t^*)(1+\gamma^y)e^{nt+zt}$

The equivalence of (i) and (ii) follow directly from the definition of ρ_t. The budget constraints $c_t^* = w_t(1 - l_t^*)$ and $c_t = w_t(1 - l_t) + w_t T_t$ imply the equivalence of (ii) and (iii). The assumed functional forms imply that (i) and (iv) are equivalent. In words, the retirement inducing motive dominates the redistributive motive when ρ_t exceeds 0.5, β_t exceeds $(\alpha_t - \alpha_t^*)(1+\gamma^y)e^{nt+zt}$, the dollar value of distortions $w_t(l_t - l_t^*)$ exceeds the dollar amount of redistribution $w_t T_t$, and when the old are forced "to pay" (ie, reduce c_t) for at least some of their increased leisure.[22]

While budget constraints and the definition of ρ_t imply the equivalence of (i)-(iii), (iv) is equivalent only because the functional forms. In other words, policy *outcomes* determine whether (i)-(iii) point to the dominance of the redistributive or retirement-inducing motives. (iv), in contrast, measures the dominant motive according the policy objective – whether β_t exceeds $(\alpha_t - \alpha_t^*)(1+\gamma^y)e^{nt+zt}$. Which is more the more relevant test depends on the purpose for conducting a test,

---

[21]Given the assumed functional forms for social welfare and elderly utility functions,

$$\rho_t = \left[1 + \frac{1 + \frac{\alpha_t - \alpha_t^*}{\beta_t}}{1 + (1+\gamma^y)e^{-(n_t + z_t)}}\right]^{-1}$$

which approaches one for INR models with negative $(n_t + z_t)$.

[22]Since there are a number of studies that compute $w_t T_t$ program-by-program and for all government programs taken together, Mulligan and Philipson (2000) suggest that (iii) may be the computationally most economical version of the test (for the purposes of showing the redistributive motives are less important for all government programs taken together), because $w_t(l_t - l_t^*)$ might be computed just for the largest program and still exceed $w_t T_t$ for all government programs taken together.

and the believed accuracy of the logarithmic functional forms. If retirement were very inelastic to taxes, then β_t could far exceed $(\alpha_t - \alpha_t^*)(1+\gamma^y)e^{nt+zt}$ while the measured income effects of the program exceeded the substitution effects.

IV. The Design and Generational Incidence of Social Security – Stylized Facts

There are six readily available facts that are relevant for determining the importance of Social Security's retirement inducing motive: (1) elderly marginal tax rates are positive and large, (2) Social Security taxes young and subsidies old, (3) elderly marginal tax rates increase significantly with the size of the Social Security program, (4) the elderly are liable for the same payroll tax as the young, (5) Social Security reduces the earnings of the elderly by about as much as it pays them, and (6) the usual generational accounts are only weakly related to generational incidence. I present these empirical findings, and compare them with the predictions of the nested model.

IV.A. Social Security Induces Retirement

As of 1995, over 100 countries had public pension programs.[23] Among the 88 of those countries reporting to the U.S. Social Security Administration sufficient detail of their public pension benefit formulas, 75% pay pension benefits in such a way as to discourage work by its elderly citizens. The most typical means by which benefit formulas induced retirement is remarkably transparent: retirement is a necessary condition for receiving public pension benefits, and no credit is given to those who decide to retire later and collect benefits for fewer years. Other countries had more complicated benefit formulas extending some less-than-actuarially fair credits to those who delay retirement, or allowing employed elderly to collect partial benefits, or both (the case for U.S. Social Security for elderly aged 65-69). But the more complicated formulas have much the same effect as the simple one: elderly labor income is implicitly taxed.

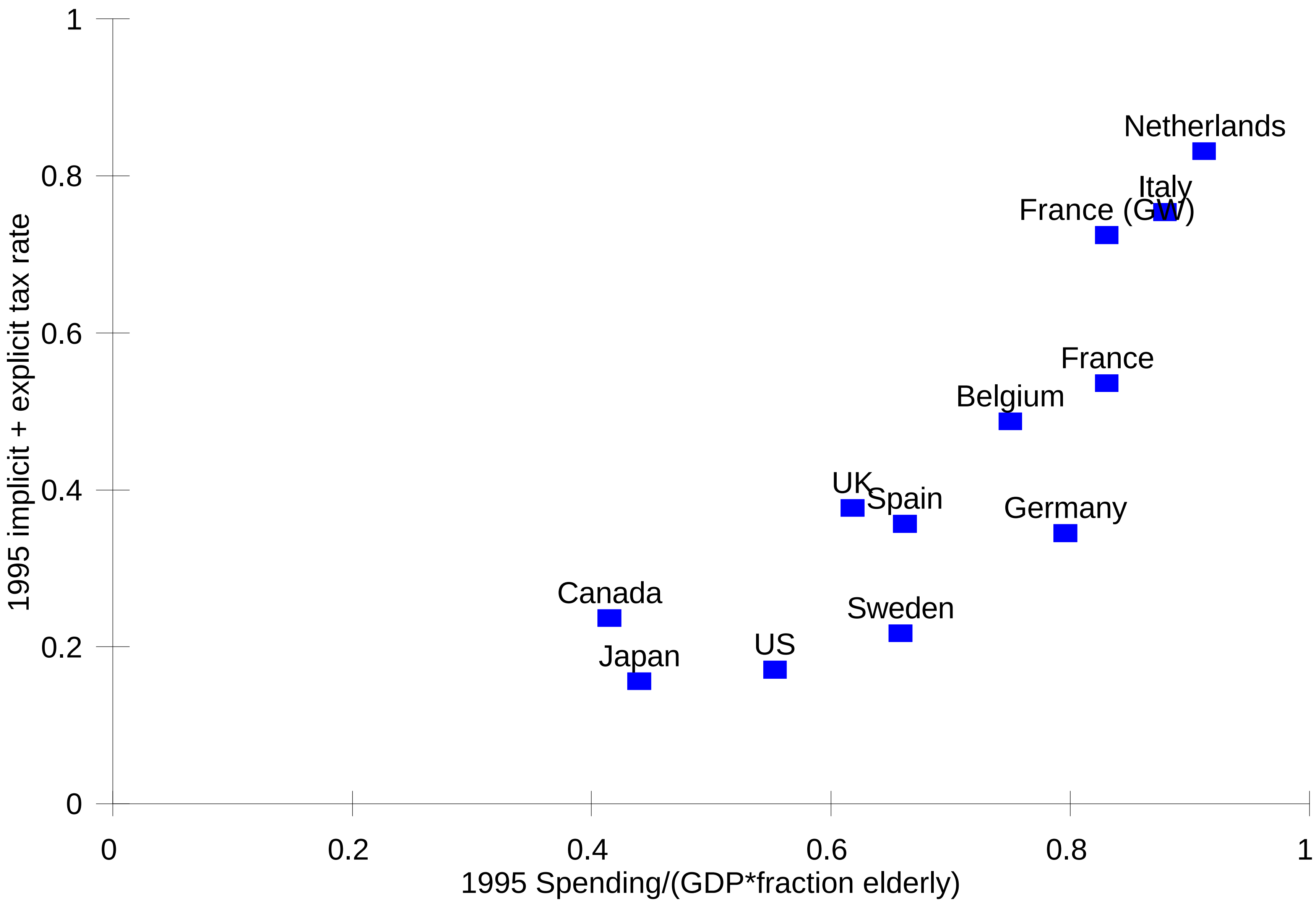
At least in higher income countries, the rates of implicit and explicit taxation are enormous. Contributors to the Gruber and Wise (1999) volume attempt to quantify the effective marginal rates for 11 countries. According to their calculations for the early 1990's, the "typical" marginal tax rate for "someone of retirement age" ranges from roughly 20% for Japan, U.S., and Canada, to more than

---

[23]Data in this paragraph are reported and described in more detail by Mulligan and Sala-i-Martin (1999a,b) and Sala-i-Martin (1996).

75% for Belgium, the Netherlands, and maybe France.[24]

[24]In any one country, marginal implicit rates vary with earnings, age, calendar year, and other variables. For a person of age *t* in the early 1990's, where *t* is between 60 and 69, the 1999 volume computes for a worker of median earnings the present value of public pension benefits foregone (net of Social Security taxes) by delaying retirement one year, and express it as a fraction τ_t of earnings (after income and payroll taxes) for that year. I average τ_t, which can be interpreted as an implicit tax rate, from $t = 60$ to 69 to arrive at the "typical" implicit tax rate for each country's elderly. This is basically the same exercise done by Gruber and Wise (1999) in their Table 1, although they compute a sum (aka, "tax factor") somewhat differently for different countries. Also, I am unable to reconcile the French tax factors reported in their Tables 1 and 3.5, so I graph France twice in the Figure – once based on their Table 1's data ("France (GW)") and once based on their Table 3.5's ("France").


1995 implicit + explicit tax rate
1995 Spending/(GDP*fraction elderly)
1
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1
Netherlands
Italy
France (GW)
France
Belgium
UK
Spain
Germany
Canada
Sweden
US
Japan

Gruber and Wise make some reasonable judgements as to when programs not called "old age", such as disability insurance and unemployment benefits, are in practice earlier retirement programs paying most of their benefits to relatively old citizens. However, Mulligan and Sala-i-Martin (2000) point out that old age, disability, and unemployment programs are not the only government programs that strongly induce retirement. Other important examples include national labor union rules requiring retirement, tax-favored retirement savings, and mandatory private defined-benefit pensions. Their results suggest that the true marginal tax rate is probably higher than shown in Figure 3, although too little is known about these programs to say whether the slope of the relationship shown in Figure 3 is accurate.

It is interesting, but difficult, to compute marginal Social Security tax rates for the young and compare them with those for the old shown in Figure 3. The payroll tax rate is one proxy for σ_t^y although, as Feldstein and Samwick (1992) explain, it overstates the marginal rate to the extent that young workers anticipate their marginal contribution to increase their retirement benefits. For example, their Table I reports gaps between the payroll tax rate and the marginal tax rate ranging from one to nine percentage points. With this in mind, we can compare the difference between the tax rates reported in Figure 3 and the payroll tax rate (combined for employer and employee). Those two rates are basically identical for Japan and Sweden, and differ by roughly 50 percentage points for France (GW), Italy, and the Netherlands. Hence, we can conclude that elderly marginal tax rates significantly exceed those for the young.

Gruber and Wise also suggest that the large elderly marginal tax rates have a significant impact on retirement behavior in the countries they study. But it is important to note that the basic insights of the IGR and INR model obtain even if retirement is not very responsive to marginal tax rates. In particular, optimal policy in the INR model involves a positive marginal tax rate regardless of the responsiveness of labor supply because the marginal tax rate serves the purpose of "internalizing" into an elderly person's decision to supply labor the social effect represented by equation (1)'s third term.

IV.B. Generational Accounts

Public retirement funds are almost always paid for by the young. It is rare for a country to have a fully-funded program (Mulligan and Sala-i-Martin 1999), so that most Social Security programs redistribute from younger generations to older ones. In fact, the cross-cohort redistribution is much more important than redistribution in any other dimension by these programs (e.g., Auerbach et al 1992, Auerbach et al 1999, Jensen and Raffelhuschen 1997, Hagemann and John 1997, House Committee 1996 table 1-50).

IV.C. Elderly Marginal Tax Rates Increase Significantly with the Size of the Program

Figure 3 measures Social Security spending for 1995 on the horizontal axis[25] for the 11 countries studied by Gruber and Wise, normalizing by GDP and the fraction of the population over age 65. Social Security is a very large program in these countries, with governments spending a fraction of GDP on elderly that is nearly as large as their fraction of the population. More relevant for comparing the IGR and INR models, however, is that the largest Social Security programs also have the largest elderly marginal tax rates (σ). σ increases rapidly with the size of the program (T), especially among the countries with bigger programs. This finding is difficult to reconcile with a redistributive motive that operates at the margin, because it is likely that the elderly in France, Italy, or the Netherlands are worse better off (or at least not much better off) with their Social Security program than with a program like that in Germany where the government does less to distort elderly labor even while it spends a lesser fraction of its GDP on each elderly person.[26] The difference between Germany and these other countries is more easily explained by the retirement-inducing

[25]"Social Security spending" is the sum of public old age cash, disability cash, and survivor's benefits from OECD (1997, lines 1, 2, and 6). These include civil servant and military pensions and in kind benefits for survivors, but exclude private pensions mandated by the government, worker's and unemployment compensation, sickness and family benefits, and in kind benefits for the elderly and disabled.

[26]Within Europe, the slope of marginal tax rate with respect to (spending per elderly)/(per capita GDP) is at least 2 in Figure 3. If we think each country's elderly person's private full income roughly equal to 4 times that country's per capita GNP (see my computations in section IV.F), then Figure 3 displays a European slope of σ_t with respect to T_t of at least 8 (remember that T_t is expressed as a fraction of elderly private full income). A logarithmic elderly utility function with $\gamma = 2$ implies a marginal rate of substitution of σ_t with respect to T_t of much less than 8 for a typical European country. In other words, elderly indifference curves in the $[T_t,\sigma_t]$ plane are flatter than the empirical relationship shown in Figure 3. Relatively steeper indifference curves are implied by the theory if the Frisch labor supply elasticity were less than one, $\gamma > 2$, or the ratio of elderly private full income to per capita GNP were much less than 4.

motive.

It might be said that implicit marginal tax rates and program size are correlated because both derive from a social "taste" for government activity. "Social taste" is one way to interpret the parameters of the social welfare function, but the question posed in this paper is whether that the "taste for government" is mainly a taste for redistribution, a taste for altering private-sector decisions, or some combination of these.

IV.D. Cross-hauling

Most countries around the world have a Social Security payroll tax (SSA 1995, as tabulated by Mulligan and Sala-i-Martin 1999a), whose rules apply equally to young and old. [27] Furthermore, the rates of payroll taxation are quite significant and, including the old age, disability, survivor, health, and maternity portions of the tax, are near 50% in a number countries, such as Netherlands, Egypt, and Italy. In other words, even if Social Security *benefits* were lump sum, the old would face large marginal tax rates. Since cross-hauling is so prevalent and important in Social Security programs, how can it be that the programs are designed to help the elderly at the expense of the young when deciding to work results in an old person's being treated (by the program) like a young one?

IV.E. Distortion Exceeds Redistribution?

Test (iii) of the dominance of the retirement-inducing motive is perhaps the most intuitive and straightforward to implement, although its interpretation relies most heavily on the assumed logarithmic functional form. The test requires three data items: the amount w_tT_t transferred by the government from young to old, the marginal elderly labor product w_t, and the change in elderly work induced by the program l_t - l_t^* (which includes both income and substitution effects). I make estimates for the three items, for the U.S. (1970 and 1995), and for three of the countries shown in Figure 3: Spain, Netherlands, and Belgium (1995 only).

The first data item, w_tT_t, is relatively easy to obtain. w_tT_t is reported as a fraction of GDP in Table 3, from the National Income and Product Accounts for the U.S. and from OECD (1997, lines 1, 2, 6) for the others, and includes public Old Age, Disability, and Survivors pensions paid to

[27]Sweden is one exception, where part of the "employer's share" of the payroll tax is forgiven when the worker reaches age 65 (Palme and Svensson 1999, p. 370).

(former) government and civilian employees.

I obtain the second item for men, w_t, from the gross annual pay used for the "base case" in various chapters of Gruber and Wise (1999). w_t is computed for women as 80% of its male value. Mine might be an underestimate because it excludes the value of fringes provided by employers, but might be an overestimate because the marginal product of those age 60+ may be less than those younger.

The third item – the change in elderly work induced by Social Security l_t - l_t^* – is much more difficult to estimate. Constructing trustworthy estimates is well beyond the scope of this paper,[28] especially when it is recognized that most pensioners are women, but a few rough calculations can be informative. I offer three, each of which measures l_t - l_t^* as the labor force participation rate for four elderly groups: men age 60-64, men aged 65+, women age 60-64, and women aged 65+. My three calculations differ according to the source of their estimate for l_t^* for each of the four groups. The first sets (1-l_t^*) = equal to the 1995 gender-specific aged 25-54 labor force participation rate which, even though it presumes that Social Security has no effect on hours worked, is obviously and upper bound on a more accurate estimate. My second calculation, "Method 1," sets $l_t^* = l_{1950}$ for men and 1-(1-l_{1950})(1-$l^y{}_{1995}$)/(1-$l^y{}_{1950}$) for women, where (1-l_t^y) is the year t aged 25-54 female labor force participation rate and (1-l_t) is the gender-specific elderly labor force participation rate.[29] "Method 2" is the same as "Method 1," except that it assumes Social Security does not affect labor force participation for those aged 65+. The three calculations are displayed in the second, third, and fourth columns of Table 2, with changes in the labor force participation rates for the four groups are displayed as small numbers and the results of the calculation displayed as large numbers.

---

[28]There are many studies attempting estimates. Kapteyn and do Vos' (1999, p. 302) summary of studies for the Netherlands is equally applicable for the other countries: "[There is] ample evidence for the dominant role of financial incentives and eligibility rules in the explanation of the low labor force participation rate among the elderly in the Netherlands. However, no study has yet fully quantified the part of the decrease that can be ascribed to the changes in incentives and eligibility rules that have occurred over the last three decades."

[29]In other words, "Method 1" assumes that, in the absence of Social Security, elderly labor force participation 1950-95 would have been constant for men, and would have grown for old women in the same proportion it did for young women.

Table 2: "Distortion" Exceeds Redistribution
(large numbers are % of GDP)

country	year	redistribution	"distortion"								
			Upper bound			Method 1			Method 2		
US	1995	5.9	11.7	37	73	3.7	28	29	0.9	28	0
				38	67		13	11		13	0
Spain	1995	10.6	10.2	48	90	7.3	44	63	1.8	44	0
				39	53		33	33		33	0
Nether	1995	11.9	14.5	72	91	5.8	62	29	2.6	62	0
				54	62		31	16		31	0
Belgium	1995	12.0	8.6	75	91	2.7	59	17	1.4	59	0
				61	65		30	13		30	0
US	1970	4.7	10.1	16	65	2.3	7	20	0.3	7	0
				10	40		0	4		0	0

Notes: (1) large numbers are redistribution and "distortion" as a percentage of GDP
(2) small numbers are the estimated Social-Security-induced changes in labor force participation rates, used to calculate distortion, in percentage points (each country-method cell's top row is male aged 60-64, 65+; bottom row is female aged 60-64, 65+)

My "upper bound" estimates of the distortion $w_t(l_t - l_t^*)$ are close to, or larger than, the size of the Social Security program w_tT_t. For example, the 1995 US estimate of the upper bound of the reduction of elderly earnings by Social Security is 11.7% of GDP, while the program paid "only" 5.9% of GDP to the elderly. However, the Method 1 and 2 estimates of the distortion are smaller than the size of the program for each of the countries. Hence, my preliminary calculations suggest that the value of distortions is similar to the cost of the program, perhaps with the former a bit smaller; both redistributive and retirement-inducing motives are important for explaining the average dollar of Social Security.

The amount of redistribution grows by 26% in the U.S. between 1970 and 1995, while the amount of distortion grew by 16%, 61%, and 313% according to the upper bound, Method 1, and Method 2, respectively. Since the amount of redistribution did not grow rapidly 1970-95, it seems that, as long as it is true that Social Security substantially increased retirement since 1970 as suggested by all three methods, the retirement-inducing motive has been particularly important in the

U.S. since 1970.

Table 2 makes it clear how an estimate of "amount of distortion" is sensitive to estimates of the effect of Social Security on the elderly labor market. If, for example, one believed that retirement were insensitive to Social Security retirement incentives (ie, both its income and substitution effects were small), then distortion estimates would be smaller. However, also remember that my interpretation of the gap between distortion and redistribution in terms of the parameters α_t and β_t depends on the logarithmic functional form; distortions that are less than redistribution are consistent with β_t larger than α_t when retirement is insensitive to policy.

IV.F. Revised Generational Accounts

At least among the developed countries, it appears that the INR model explains more of the differences over time and across countries in the design and incidence of Social Security. It is less clear whether the INR or IGR model better explains the average dollar of Social Security spending. But even if intergenerational redistributive motives were dominant on average and at the margin, the presence of *any* retirement-inducing motive means that the generational incidence of Social Security can be very different than suggested by the generational accounts (4) as computed in the literature.

IV.F.1. Revision 1: Benefits are Valued at Less than Budgetary Cost

Remember from section III that the presence of an induced retirement motive implies two revisions to the date-of-birth account for generation *t*. The first revision I denote ε_{t+1}, when measures the logarithmic gap between the budgetary cost of old age subsidies T_t and their value to the elderly. With a logarithmic elderly utility function ($\ln c_t^o + \gamma \ln l_t^o$), ε_{t+1} is easily computed using the formula (5), which is repeated below for convenience:

$$e^{-\varepsilon_t} = \frac{(1+T_t)\left(1 - \frac{\sigma_t}{1+\gamma}\right)^{-1}(1-\sigma_t)^{1/(1+\gamma)} - 1}{T_t}$$

This formula depends on three things: (i) the marginal elderly tax rate σ_t, (ii) the size of the program as a fraction T_t of elderly private full income, and (iii) the demand for leisure in the absence of the program (dictated by γ).[30] σ_t can be read from the data in Figure 3. Figure 3 also reports government spending per elderly person, but as a fraction of per capita GDP rather than as a fraction of elderly full private income. If we suppose that elderly full private income is proportional to per capita GDP, then what we need to transform the Figure 3 data into a series for T_t is the proportion. In the U.S., for example, 1995 GDP per capita was $28,000 while GDP per worker was $58,000 (Council of Economic Advisers 1998, Tables B-1, B-34, and B-35). A worker's full income is, of course, greater than actual income, so perhaps elderly annual full private income is between $75,000 and $150,000. At $100,000, elderly full private income is nearly 4 times larger than per capita GDP, so I divide Figure 3's elderly spending data by 4 to compute a series for T_t. Finally, I conservatively choose a fairly large value for γ, 2 (the revision ε_{t+1} decreases with γ), although the formula (5) is not particularly sensitive to γ. I report the results in the first two columns of Table 3 for each country shown in Figure 3. Since my policy data is for the 1990's, these computations of ε_{t+1} are for those generations *t* elderly in the 1990's.

[30] $1/(1+\gamma)$ is the fraction of time the elderly would spend working if there were no Social Security program and they had no financial assets.

Table 3: Revised Generational Accounts
(for those elderly in 1995)

	revision ε_{t+1} for 1990's elderly		revision λ_{t+1} for 1990's workers	
Country	avg value of benefits	ROR (%/yr, -) adjustment	avg cost of taxes	ROR (%/yr, +) adjustment
Netherlands	-0.27	∞	-0.17	∞
Italy	0.10	7.8	-0.11	∞
France	0.67	1.4	-0.09	∞
Belgium	0.72	1.1	-0.62	∞
U.K.	0.83	0.6	0.59	1.7
Spain	0.86	0.5	0.45	2.6
Germany	0.89	0.4	0.30	4.0
Canada	0.92	0.3	0.18	5.7
Sweden	0.95	0.2	0.89	0.4
U.S.	0.97	0.1	0.76	0.9
Japan	0.97	0.1	0.89	0.4

Notes: (1) "avg value of benefits" is the program's lump sum equivalent (E^o) divided by size of the program (T) = $e^{-\varepsilon t+1}$
(2) Elderly Rate of Return adjustment is computed according to equation (5), and then annualized = $\varepsilon_{t+1}/30$
(3) both elderly calculations assume log utility with $\gamma = 2$
(4) "avg cost of taxes" is the program's lump sum equivalent (E^y) divided by size of the program (T) = $e^{-\lambda t+1}$
(5) Workers' Rate of Return adjustment is annualized = $\lambda_{t+1}/30$
(6) Canada pays for much of its Social Security from general revenues (SSA 1995). If a payroll tax rate of 12% (rather than 6%) were used in the E^y calculation, then the avg cost of taxes would be 0.60
Source: Figure 3 and author's calculations

The first column is the average value of benefits, which is the Social Security program's lump sum equivalent divided by the size of the program. For example, the average value of benefits is 0.97 for the U.S. – the representative elderly person with log utility ($\gamma = 2$), would accept a 3% reduction in benefits if those benefits were paid as a lump sum (ie, paid independent of earnings after age 61). For the typical European country studied by Gruber and Wise, the average value of benefits is about 0.8. Elderly in two countries – Italy and the Netherlands – value benefits at well less than half of their budgetary cost. Indeed, I calculate a negative value of benefits for the Netherlands, suggesting that their elderly might be better off with no subsidy and incentive to retire.

The second number reported in Table 3 for each country is the additional discounting by the revised generational accounts so that, ignoring the intergenerational incidence of the benefits of induced retirement indicated by the third term in the policy objective (1), the accounts are indicative of generational incidence. To put the return adjustments in perspective, compare them with the annual "Social Security" rates of return (in my notation, $g_{t+1}/30$) of 3 to 5% computed by Geanakoplos et al (1998) and Leimer (1994) for the U.S. Social Security system for cohorts retiring in the 1990's. Hence, the 0.1% adjustment indicated by Table for the U.S. and Japan is quantitatively unimportant, while adjustments for Belgium, Netherlands, France, and Italy are quite substantial.

Both the Table and equation (5) show that estimates of the average value of benefits are sensitive to the marginal tax rate measures, especially for high tax rates. More research is needed to accurately estimate those rates, but a few things are clear from the computations reported in the Gruber-Wise (1999) volume. First, estimated marginal tax rates would be higher if discount rates were higher than those on government bonds, which might be the case if the elderly have some difficulty borrowing against future Social Security benefits. Second, there is variability in marginal rates across persons within a country which, since the value of benefits is a concave function of the marginal tax rate, means that value of benefits averaged across persons is lower than the value of benefits for the average person reported in the Table. Both of these considerations suggest that I may substantially overestimate the average value of benefits for the European countries.

Remember from Section III that the revised remaining-lifetime generational accounts must *always* be sufficiently close to the usual accounts (4) in order for perpetual intergenerational redistribution to be supported as subgame perfect political equilibrium. The first two columns of Table 3 show that the revised accounts are quite different for current generations in several European countries. More importantly, given the current prevalence of high elderly marginal tax rates, is it reasonable to suppose that voters in countries with small revisions like the U.S. or Japan expect that their country will *never* have such high marginal tax rates? This supposition is required if perpetual intergenerational redistribution is to be supported as a subgame perfect political equilibrium.

IV.F.2. Revision 2: Budgetary Cost May Overstate the Young's Cost

In the case that half of the benefits of the third term in the policy objective (1) accrue to the young, λ_t can be computed as:

$$e^{-\lambda_t} = \frac{1 - (1 - \sigma_t^y)^{(1 - l_t^y)} \left(\frac{l_t}{l_t^*} \right)^{\frac{\sigma_t}{2} \frac{\sigma_t^y (1 - l_t^y) + e^{-z_t}}{1 - \sigma_t^y} \frac{l_t^*}{1 - \sigma_t (1 - l_t^*)}}}{(1 - l_t^y)\, \sigma_t^y}$$

Data on σ_t comes from Figure 3, and σ_t^y is measured as the payroll tax rate reported by SSA (1995). In addition, computing λ_t requires data on n_t, z_t, l_t^y and the change in elderly leisure time induced by the program l_t/l_t^*. As an example, I set $z_t = 0$, $l_t = 2/3$, and estimate n_t and l_t/l_t^* for each country using demographic and labor force data. Since my policy data is for the 1990's, λ_t can be computed for cohorts of working age in the 1990's, and hence a younger cohort than those for whom the average value of benefits is computed in Table 3.

Table's 3 third column shows that the average cost of taxes to workers can be quite small, which implies the often large rate of return adjustments to their date of birth generational accounts shown in the fourth column. To put the return adjustments in perspective, compare them with the expected annual "Social Security" rates of return of 1.5-2% computed by Geanakoplos et al (1998) and Leimer (1994) for the U.S. Social Security system for cohorts retiring in the 1990's, and with the 1-3% gap between the lifetime rates for retired and working cohorts. For nearly every country, the fourth column's revisions meet or exceed the level of, and cohort gaps in, returns computed in the literature.

The average cost can even be negative, which means that workers gain more from induced retirement than they pay in taxes. In this case, their lifetime rate of return is "infinite" because they gain both when young and old! Of course, the estimates in the third and fourth column are much rougher than those in the first two columns because additional data – including the fraction of the benefits of induced retirement that accrue to the young – is required to make the estimate. And a complete calculation requires estimates of λ_t for the cohorts elderly in the 1990's and ε_{t+2} for the 1990's working cohorts. But there is one important, and less easily refuted, reason why the revisions reported in the second two columns of the Table are so large – all 11 countries have substantial marginal tax rates which, in my nested model of Social Security, indicates that induced retirement is of substantial value to somebody other than the retiree.

In summary, two lessons are learned from my revision of date-of-birth generational accounts. First, early generations gain less from Social Security than suggested by the generational accounts computed in the literature, and may not gain at all. This adjustment is of quantitative importance mainly in countries like Belgium, Netherlands, France, and Italy where marginal elderly tax rates well exceed 50%. Second, later generations lose less than suggested by the generational computed in the literature, and may actually gain. This second revision is uncertain, since little is known about exactly why governments encourage retirement, but my calculations suggest that the second revision can be very substantial.

V. Conclusions

I model redistributive and retirement-inducing motives for paying social security benefits. I do not say much about why governments would want to induce retirement, or to improve the welfare of the old, but basic similarities and differences can be derived even without saying much about these motives. Both motives are consistent with programs redistributing funds from young to old, and can give rise to a sequence of policies that look like a "pyramid scheme." Both are consistent with programs that improve the welfare of the elderly.

However, the two approaches have a number of different implications. Intergenerational redistributive motives are difficult to reconcile with programs that strongly encourage retirement and, when they operate at the margin, difficult to reconcile with a strong positive relationship between retirement incentives and Social Security spending. Generational accounts and generational incidence are only weakly related, if at all, under retirement-inducing motives. Income effects do the most to change behavior under redistributive motives while, assuming the substitution elasticity of retirement is large enough, substitution effects dominate under retirement-inducing motives.

These differences between the redistributive and retirement-inducing motives mean that it is important to determine the relative importance of each motive for the creation of, and marginal changes in, actual Social Security programs. I show how most Social Security programs around the world induce retirement, and strongly so in more than a few countries. I show how cross-country or time-series relationships between elderly marginal tax rates and Social Security spending can be used to gauge the marginal importance of redistributive and retirement-inducing motives. For the eleven OECD countries studied by Gruber and Wise (1999), elderly marginal tax rates are closely (and positively) related with the amount of Social Security spending. Rates increase so rapidly with program size in Europe that inducing retirement seems to be the only model for understanding European differences in Social Security design and incidence. It appears that cross-hauling is a prevalent and important feature of Social Security program; it is hard to justify with a cohort-redistributive motive why, to the extent that an old person works, he is treated like a young one by the system.

Much more work is needed to accurately measure the elderly substitution effects of Social Security. They appear to be large, but not quite as large as the income effects. The importance of income effects suggest that the redistributive motive may be important for understanding the creation of Social Security, even if it is relatively unimportant for understanding its recent growth and differences across countries.

I introduce two revisions to generational accounts, so that they more accurately measure the generational incidence of government policy. The first revision accounts for the gap between the budgetary cost of benefits and their value to the elderly, a gap created by the positive elderly marginal tax rates. The second revision values the benefits to the young of induced retirement. Even though Social Security's income effects may exceed its substitution effects on the elderly, the evidence shows the first revision is quite substantial for some countries while the second revision may be substantial for all of the eleven countries studied. Policy is at least in part designed to induce retirement, and its generational incidence is probably very different than the incidence of a pyramid scheme.

VI. Appendix I: Corner Solutions in the Nested Model

The analysis in the text ignores the constraints $l_t^y \leq 1$ and $l_t^o \leq 1$ in the planner's problem (2). For some values of α_t and β_t, the constraints bind. Here I derive results for the case when $l_t^o \leq 1$ binds, but $l_t^y \leq 1$ does not. From the first order conditions of the planner's problem (2), it is easy to show that $l_t^o \leq 1$ binds if and only if

$$\beta_t > \beta_t^{corner} \equiv (\alpha_t^* - \alpha_t) + \frac{1 + (1+\gamma^y)e^{-(n_t + z_t)}}{1+\gamma}\alpha_t$$

In this case, the optimal solution to (2) is:

$$T_t = \frac{\alpha_t}{\alpha_t(1+\gamma^y)e^{-(n_t+z_t)} + (1+\gamma)\alpha_t^*}, \quad \sigma_t^y = \frac{\alpha_t}{\alpha_t + e^{n_t}},$$

$$\sigma_t = \frac{\alpha_t(1+\gamma^y)e^{-(n_t+z_t)} + (1+\gamma)\alpha_t^* - \gamma\alpha_t}{\alpha_t(1+\gamma^y)e^{-(n_t+z_t)} + (1+\gamma)\alpha_t^*}$$

As long as $\beta_t > \beta_t^{corner}$, neither T_t nor σ_t does not vary with β_t. Hence, changes in β_t do not induce a positive relationship between T_t and σ_t, as they do when $\beta_t < \beta_t^{corner}$. Changes in α_t induce a negative relationship between T_t and σ_t, as they do when $\beta_t < \beta_t^{corner}$.

VII. Appendix II: Is Distortion an Accidental Byproduct of Redistribution?

Perhaps the IGR model is oversimplified because it does not allow any distortion of the old's behavior as a response to the redistributive motive. In particular, might distortion of the elderly labor supply decision be an accidental byproduct of redistribution? There is not much research attempting to answer this question, and a thorough answer is beyond the scope of this paper, but this appendix briefly considers some embellishments of the IGR model. I suggest that two results derived in the main text for the IGR model are robust to these and other likely modifications: (1) the marginal tax rate will not increase too rapidly with the size of the program and (2) income effects must be more important than substitution effects in the IGR model.

VII.A. Means Test

Perhaps redistribution in general, and intergenerational redistribution in particular, is not politically feasible unless it favors the poor. And one of the realities of public finance is that helping the poor involves distorting labor supply. Tabellini (1992) has built one such model of Social Security, where the elderly invite the poor to form a winning coalition in a government for which policies are chosen by majority vote.

The main problem with such an explanation is that, in practice, Social Security does not favor the poor, nor do other government policies do nearly as much redistribute across income classes as they do across cohorts. First, in 98% of the 88 countries studied by Mulligan and Sala-i-Martin (1999a), Social Security benefit formulas have no link to the beneficiary's non-labor income. An elderly person can make millions of dollars in the stock market without sacrificing public pension benefits, but will make a sacrifice with only a small amount of income from work.

Second, it does not appear that Social Security, tax, and other government policies for the elderly are progressive. Many studies of American SS (Burkhauser and Warlick 1981, Garrett 1995), Medicare (McClellan and Skinner 1997), and elderly tax policy (Nelson 1983) suggest that government policy toward the elderly is neither progressive nor regressive.[31] Third World Social Security Programs appear to be regressive (Pampel and Williamson 1989, page 10; Midgley 1984). Several European programs have far more generous benefits at higher salary levels (apRoberts 1996, pp. 109, 112) and may thereby be more regressive than American SS. A comparison of empirical studies of the generational incidence (eg., Auerbach et al 1999 study 14 countries) of government policy with studies of its income incidence (eg., Pechman 1985) makes it clear that generational redistribution is vastly more important than redistribution across income classes.

The worldwide scarcity of large mean-tested programs is easy to understand in the INR model where inducing retirement, not helping the poor, is the primary policy motive. A means-tested program discourages retirement savings which, because retirement savings facilitate *retirement*, implies that a means-tested program has a weaker impact on retirement than does a program that implicitly taxes the earnings of the elderly without taxing their assets or unearned income.

VII.B. Allocative Efficiency

If the IGR model were modified to rule out the use of lump sum taxes on the young, then clearly the optimal policy would involve $\sigma_t^y > 0$. But, since year of birth is both observable and difficult to change, it seems that lump sum subsidies for the old are still feasible. Would the optimal policy in the IGR model still involve $\sigma_t^o = 0$? Because a policy setting $\sigma_t^y > \sigma_t^o = 0$ violates the law

[31]Boskin et al (1987) is one study showing a little progressivity in the OASI system.

of one price, it might not be optimal, depending on the ability of the young to evade their payroll taxes. With $\sigma_t^o = 0$, an old person has an incentive to hire an "unemployed" young person to do his work for him, pay "under the table" the young worker his marginal product, and report the young man's earnings as his own. Such arrangements might occur in families and firms, or even in the black market, and would reduce the ability of the government to redistribute across cohorts.[32] Setting $\sigma_t^o = \sigma_t^y$ would eliminate incentives for such arrangements, and would clearly induce a positive correlation between the amount of intergenerational redistribution (which varies directly with σ_t^y) and the magnitude of elderly marginal tax rates. However, this version of the IGR model rules out $\sigma_t^o > \sigma_t^y$, and requires that σ_t^o not increase too rapidly with σ_t^y (for the reasons shown in Figure 2). Also, as long as the elderly population is small relative to the young population, the substitutions effects on young and old are small relative to the income effects on old.

VII.D. Screening

Many have suggested that redistribution is optimally distortionary in order to prevent excessive entry into the subsidized group. For example, Stigler (1975, pp. 115f) suggests that this is why special interests do not lobby for cash subsidies, and Besley and Coate (1991) suggest that this may be a reason why some antipoverty programs do not pay in cash. If subsidies were in cash, they argue, too many people would become farmers or claim to be poor, or whatever it takes to join the subsidized group. Many policy distortions may in fact serve to screen recipients, but is this a good theory of Social Security? Are benefit distortions needed to prevent too many people from becoming old, or claiming to be old? While it may not be possible for government to determine who would be a farmer, or who would be poor, in the absence of a subsidy, governments easily observe a citizen's year of birth, and already use its observations to determine benefits.

VIII. Appendix III: A Nested Model of Social Security and Aggregate Capital Accumulation

Here I show similar implications might be derived for an infinite horizon economy with capital. For brevity, the Appendix assumes that taxes on the young can be lump sum, so that the only budget constraint is an intertemporal budget constraint for the planner:

---

[32] See Mulligan (1999) for additional analysis of the allocative efficiency of policies affecting young and old labor markets.

$$\sum_{t=0}^{\infty} R_t \Big[N_t c_t^y + N_{t-1} c_t^o - N_t w_t^y (1 - l_t^y) - N_{t-1} w_t^o (1 - l_t^o) \Big] \le k_0$$

where R_t is the interest rate factor for period t, and depends on the return to capital between periods 0 and t. r is the life cycle rate of time preference, which is the same for citizens and government (ie, government relatively weights a person's young and old utility the same way the person does). k_0 is the initial capital stock, and N_{-1} is the initial population of old people. The time zero government cares about the discounted sum of remaining lifetime utilities of those alive and unborn as of time 0:

$$N_{-1}\Big\{a_{-1} u^o(c_0^o, l_0^o) + b_{-1} v(l_0^o)\Big\} + \sum_{t=0}^{\infty} N_t \Big\{a_t \Big[u^y(c_t^y, l_t^y) + e^{-r} u^o(c_{t+1}^o, l_{t+1}^o)\Big] + b_t e^{-r} v(l_{t+1}^o)\Big\}$$

where $\{a_t, b_t\}_{-1}^{\infty}$ are sequences of constants. The first term in braces is the sum of the per capita utility of the initial old, and the extra social benefit (per initial old) of their retirement. The terms in square brackets are the per capita discounted lifetime utilities of those born at date 0 or later. The term under the sum, but outside the square brackets, is the extra social benefit retirement (per old person). Using the definitions $\alpha_t = e^{-r} a_{t-1}/a_t$, $\beta_t = e^{-r} b_{t-1}/a_t$, the time zero government's objective can be written as:

$$\sum_{t=0}^{\infty} N_{t-1}\, a_t \Big[e^{n_t} u^y(c_t^y, l_t^y) + \alpha_t u^o(c_t^o, l_t^o) + \beta_t v(l_t^o)\Big]$$

Notice the similarity with the objective of (2) in the text. Let the time zero government choose a sequences consumption and leisure for young and old subject to the intertemporal budget constraint above, and derive the optimal policies $\{T_t, \sigma_t^y, \sigma_t^o\}$ as in the text (T_t be the difference between earnings and consumption of the time t young). Using the log functional forms from the text, the optimal policies are:

$$T_t = \frac{(\beta_t + \alpha_t)(1 - I_t) - \alpha_t^*}{(\beta_t + \alpha_t)\, e^{-(n_t + z_t)} + \alpha_t^*} \quad , \quad \sigma_t^y = 0 \quad , \quad \sigma_t = \frac{\beta_t}{\beta_t + \frac{\gamma}{1+\gamma}\alpha_t}$$

where I_t is optimal date t national savings per young person. The optimal marginal tax rates are the same as in a static economy. T_t is different than reported in the text, but mainly because there is an ambiguity between savings and taxes paid by the young.

IX. References

Auerbach, Alan J., Jagadeesh Gokhale, and Laurence J. Kotlikoff. "Social Security and Medicare Policy from the Perspective of Generational Accounting." in Poterba, James M., ed. *Tax Policy and the Economy*. Volume 6. Cambridge: MIT Press, 1992: 129-45.

Auerbach, Alan J., Laurence J. Kotlikoff, and Willi Leibfritz, eds. *Generational Accounting Around the World*. Chicago: University of Chicago Press, 1999.

Barro, Robert J. "Are Government Bonds Net Wealth?" *Journal of Political Economy*. 82(6), December 1974: 1095-1117.

Becker, Gary S. and Kevin M. Murphy. "The Family and the State." *Journal of Law and Economics*. 31(1), April 1988: 1-18.

Besley, Timothy and Stephen Coate. "Public Provision of Private Goods and the Redistribution of Income." *American Economic Review*. 81(4), September 1991: 979-84.

Bohn, Henning. "Will Social Security and Medicare Remain Viable as the U.S. Population is Aging?" Working paper, Department of Economics, UCSB, June 1998.

Boldrin, Michele, Sergi Jimenez-Martni, and Franco Peracchi. "Social Security and Retirement in Spain." NBER Working Paper No. 6136, August 1997.

Browning, Edgar K. "Why the Social Insurance Budget Is Too Large in a Democracy." *Economic Inquiry*. 13(3), Sept. 1975: 373-88.

Clague, Ewan, Balraj Palli, and Leo Kramer. *The Aging Worker and the Union*. New York: Praeger Publishers, 1971.

Cooley, Thomas F. and Jorge Soares. "A Positive Theory of Social Security Based on Reputation." *Journal of Political Economy*. 107(1), February 1999: 135-60.

Costa, Dora L. *The Evolution of Retirement*. Chicago: University of Chicago Press, 1998.

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

Fehr, Hans and Laurence J. Kotlikoff. "Generational Accounting in General Equilibrium." in Alan J. Auerbach, Laurence J. Kotlikoff, and Willi Leibfritz, eds. *Generational Accounting Around the World*. Chicago: University of Chicago Press, 1999: 43-72.

Feldstein, Martin and Andrew Samwick. "Social Security Rules and Marginal Tax Rates." *National Tax Journal*. 45(1), March 1992: 1-22.

Geanakoplos, John, Olivia S. Mitchell, and Stephen P. Zeldes. "Social Security Money's Worth." NBER Working paper series no. 6722, September 1998.

Gratton, Brian. "The Poverty of Impoverishment Theory: The Economic Well-Being of the Elderly, 1890-1950." *Journal of Economic History*. 56(1), March 1996: 39-61.

Gruber, Jonathan and David Wise, eds. *Social Security Programs and Retirement Around the World*. Chicago: University of Chicago Press for NBER, 1999.

Hagemann, Robert P., and Christoph John, "Fiscal Reform in Sweden: What Generational Accounting Tells Us." *Contemporary Economic Policy*. 15(3), July 1997: 1-12.

Kapteyn, Arie and Klass de Vos. "Social Security and Retirement in the Netherlands." in Gruber, Jonathan and David Wise, eds. *Social Security Programs and Retirement Around the World*. Chicago: University of Chicago Press for NBER, 1999: 269-303.

Kotlikoff, Laurence J. *Generational Accounting*. New York: Free Press, 1992.

Kotlikoff, Laurence J., Torsten Persson, and Lars Svensson. "Social Contracts and Assets: A Possible Solution to the Time Inconsistency Problem." *American Economic Review*. September 1988: 662-77.

Leimer Dean R. "Cohort-Specific Measures of Lifetime Net Social Security Transfers." Office of Research and Statistics Working Paper No. 59. Washington, DC: Social Security Administration, February 1994.

Miron, Jeffrey A. and David N. Weil. "The Genesis and Evolution of Social Security." NBER Working paper #5949, March 1997.

Mulligan, Casey B. “Can Monopoly Unionism Explain Publicly Induced Retirement?” NBER Working paper, May 2000.

Mulligan, Casey B. and Tomas J. Philipson. “Merit Motives and Government Intervention: Public Finance in Reverse.” NBER Working paper, May 2000.

Mulligan, Casey B. and Xavier Sala-i-Martin. “Gerontocracy, Retirement, and Social Security.” NBER Working Paper #7117, May 1999a.

Mulligan, Casey B. and Xavier Sala-i-Martin. “Social Security in Theory and Practice I: Facts and Political Theories.” NBER Working Paper #7118, May 1999b.

Mulligan, Casey B. and Xavier Sala-i-Martin. “Social Security in Theory and Practice II: Efficiency Theories and Implications for Reform.” NBER Working Paper #7119, May 1999c.

Mulligan, Casey B. and Xavier Sala-i-Martin. “Retirement-Inducing Policy Wears Many Disguises.” Working Paper, Columbia University, March 2000.

Murphy, Kevin M. and Finis Welch. “Perspectives on the Social Security Crisis and Proposed Solutions.” *American Economic Review*. 88(2), May 1998: 142-50.

Myers, Robert J. *Social Security*. 4th edition. Philadelphia: University of Pennsylvania Press, 1993.

OECD. *Social Expenditure Database*. 1997.

Olson, Laura Katz. *The Political Economy of Aging: The State, Private Power, and Social Welfare*. New York: Columbia University Press, 1982.

Palme, Martin and Ingemar Svensson. “Social Security and Retirement in Sweden.” in Gruber, Jonathan and David Wise, eds. *Social Security Programs and Retirement Around the World*. Chicago: University of Chicago Press for NBER, 1999: 355-402.

Pechman, Joseph A. *Who Paid the Taxes 1965-85*. Washington, DC: Brookings Institution, 1985.

Pestieau, Pierre and Jean-Phillippe Stijns. “Social Security and Retirement in Belgium.” NBER working paper #6169, September 1997.

Piven, Francis Fox and Richard A. Cloward. *Regulating the Poor: The Functions of Social Welfare*. New York: Vantage Books, 1971.

Sala-i-Martin, Xavier X. “A Positive Theory of Social Security.” *Journal of Economic Growth*. (1)2, June 1996: 277-304.

Samuelson, Paul A. “An Exact Consumption Loan Model of Interest With or Without the Social Contrivance of Money.” *Journal of Political Economy*. 66(6), December 1958: 467-82.

Stigler, George J. *The Citizen and the State: Essays on Regulation*. Chicago: University of Chicago Press, 1975.

Tabellini, Guido. "A Positive Theory of Social Security." mimeo IGIER, Milano, Italy, February 1992.

United States Census Bureau. Population Division, Population Projections Branch. *Annual Population Estimates by Age Group and Sex, Selected Years from 1990 to 1999*. 1999.

United States Census Bureau. *Historical Statistics of the United States, Colonial Times to 1970.* Bicentennial edition. Washington: U.S. Dept. of Commerce, Bureau of the Census, 1975.

United States Congress, House Committee on Ways and Means. *Overview of Entitlement Programs, 1996 Green Book*. Washington, D.C.: U.S. Government Printing Office, 1996.

United States Council of Economic Advisers. *Economic Report of the President*. Washington, DC: U.S. Government Printing Office, February 1998.

United States Social Security Administration. *Social Security Programs Throughout the World - 1995*. Washington, DC: Government Printing Office, 1995.