How Well Does the U.S. Social Insurance System Provide Social Insurance? *

Mark Huggett and Juan Carlos Parra
Georgetown University

Abstract

We analyze the insurance provided by the U.S. social security and income tax system within a model where agents receive idiosyncratic, wage-rate shocks that are privately observed. We consider two reforms: a piecemeal reform that optimally chooses the social security benefit function and a radical reform which eliminates the entire social insurance system and replaces it with an optimal tax on lifetime earnings. The radical reform outperforms the piecemeal reform and achieves nearly all of the maximum possible welfare gain when wages differ permanently over the lifetime. When wage shocks match properties in U.S. data, the piecemeal reform outperforms the radical reform.

JEL Classification: D80, D90, E21

Keywords: Social Insurance, Social Security, Idiosyncratic Shocks, Private Information

*We thank Robert Shimer, coeditor of this journal, two anonymous referees, John Bailey Jones, and seminar participants at different venues for their valuable suggestions.

"From the point of view of insurance, there seem to me to be two compelling theoretical arguments for having the State rather than the market provide a wide range of insurance, for old-age pensions, disability and sickness, unemployment and low income: the first is that the market handles adverse selection badly. The second is that, even if adverse selection were not important, people should take out insurance at an age when they are incapable of doing so rationally, namely zero." - Mirrlees (1995, p. 384)

I. Introduction

One rationale for a government-provided, insurance system is the provision of insurance for risks that are not easily insured in private markets. One can find this rationale in textbooks, in public policy documents and in the work of prominent economists.[1]

An important risk that is often discussed in the context of social insurance is labor income risk. Individual workers experience substantial variation in wage rates which are not related to systematic life-cycle variation or to aggregate fluctuations.[2] A common view is that labor income is not easily insured because it is partly under an individual's control by the choice of unobserved effort or unobserved labor hours and because a component of labor income risk is realized at a young age. It is often claimed that a progressive income tax system together with a progressive social security system may provide valuable insurance. The Economic Report of the President (2004, Ch. 6) claims that the progressive relationship between monthly social security benefit payments in the U.S. and a measure of lifetime labor income may be an important source of insurance.

We provide a benchmark analysis of how well a stylized version of the U.S. social insurance system provides social insurance. We do so by determining the maximum possible gain to superior insurance. We analyze only the retirement component of the social security system, treat social security together with income taxation as the entire social insurance system and focus only on a single but very important source of risk. The risk that is examined here is idiosyncratic, wage-rate risk.

Our methodology involves the analysis of two decision problems. One decision problem

[1] See Rosen (2002, Ch. 9), The Economic Report of the President (2004, Ch. 6) and Mirrlees (1995).
[2] See Heathcote, Storesletten, and Violante (2008) or Kaplan (2007).

is that of a cohort of ex-ante identical agents. Each agent maximizes expected utility in the presence of the model social insurance system. It is assumed that asset markets transfer resources over time and that the social insurance system (i.e. social security and income taxation) is the only way to transfer resources across different histories of wage shocks. We then contrast the ex-ante expected utility in the model insurance system with the maximum ex-ante expected utility that a planner could deliver to this cohort. The planner uses no more resources in present-value terms than are used by a cohort in a solution to the model insurance system. The planner is also restricted to choose allocations that are incentive compatible. The incentive problem arises from the fact that the planner observes each agent's earnings but not an agent's hours of work or an agent's wage.

The model we analyze is closely related to the work of Kaplan (2007). He first estimates a process for male wages that accounts for the variation in mean wages and the idiosyncratic component of wages over the life cycle. He then estimates preference parameters to best match moments characterizing the distribution of consumption, hours and wages over the life cycle. The main deviation from Kaplan's model is that we replace the proportional tax rates on labor and capital income in his model with the structure of the U.S. social security system and the U.S. federal income tax system.

We analyze two versions of this model. The full model captures the pattern of permanent, persistent and purely temporary idiosyncratic wage variation estimated from U.S. data, whereas the permanent-shock model shuts down the variance in the persistent and temporary shock components. The analysis of the permanent-shock model is motivated in part because we can solve the planner's problem for this model but not for the full model. Thus, we calculate maximum welfare gains to superior insurance only for the permanent-shock model. However, we calculate optimal parametric policy reforms in both models.

We find that the maximum welfare gain to improved insurance in the permanent-shock model is large. The maximum welfare gain is equivalent to a 4.09 percent increase in consumption each model period. Important differences in time spent working are behind this welfare gain. Specifically, high productivity agents work too little and low productivity agents work too much under the U.S. system as compared to the solution to the planning problem.

One reason for these differences in work time is that the pattern of intratemporal wedges in the planning problem differs markedly from the wedges under the U.S. system. In the planning problem, the wedge between the intratemporal marginal rate of substitution and the wage rate is zero for the highest wage agents at each age and increases as an agent's wage rate falls. Thus, the greatest wedge at each age is for the lowest productivity agent. In the U.S. system, the pattern of wedges is exactly the opposite because marginal income tax rates are progressive and because the social security benefit function is concave in a measure of lifetime earnings.[3]

We explore two main reforms. First, we conduct an optimal piecemeal reform by allowing the social security benefit function to be chosen optimally without changing the social security tax rate or the income tax system. This reform leads to almost no welfare gain in the permanent-shock model but a welfare gain equivalent to a 1.15 percent consumption increase each period in the full model.

The second reform is more radical. We eliminate the model social insurance system and replace it with an optimal tax on the present value of earnings. An optimal present-value tax achieves a welfare gain of 3.95 percent of consumption in the permanent-shock model - nearly all of the maximum possible welfare gain. The present-value tax performs so well because it approximates the wedges between marginal rates of substitution and transformation arising in a solution to the planning problem while allowing for a flexible relationship between lifetime earnings and lifetime consumption. In the full model this optimal reform leads to no welfare gain. Thus, while a present-value tax is well designed for models with only permanent labor productivity differences that remain over the entire lifetime it does not lead to a welfare gain in models with permanent, persistent and temporary sources of labor productivity variation that mimic properties in U.S. wage data.

Two literatures are most closely related to the analysis in this paper. First, there is the dynamic contract theory literature which analyzes optimal planning problems in which some key information is only privately observed.[4] Our work is similar in spirit to Hopenhayn and

[3]Average tax rates on lifetime earnings are substantially more progressive in a solution to the planning problem than in the model of the U.S. system. Thus, the large welfare gain originates both from too *little* progression in lifetime taxation and from the wrong pattern of marginal tax rates at each age.

[4]This work builds upon Mirrlees (1971). Golosov, Tsyvinski, and Werning (2006) review the recent

Nicolini (1997), Wang and Williamson (2002) and Golosov and Tsyvinski (2006). These papers analyze optimal planning problems and stylized social insurance systems. Second, there is the literature on social security systems with idiosyncratic risk (e.g. Imrohoroglu, Imrohoroglu, and Joines (1995), Huggett and Ventura (1999) and Storesletten, Telmer, and Yaron (1999)). Nishiyama and Smetters (2007) is one interesting paper from this literature. They consider various ways of partially privatizing the U.S. social security system. They find important efficiency gains when they abstract from idiosyncratic wage risk. When idiosyncratic risk is added, they find either no efficiency gains or very small gains for the reforms they analyze.

Our findings paint a different picture. We find that the maximum welfare gain to improved insurance substantially increases as the magnitude of idiosyncratic wage risk increases. Our work differs from Nishiyama and Smetters (2007) in at least two main ways. First, we focus on ex-ante welfare as is common in the contract theory literature rather than the ex-interim notion they use. This allows us to assess insurance provision over shocks realized early in life. Second, the methodology differs as we solve for allocations maximizing ex-ante welfare rather than trying particular reforms. This methodology allows one to determine if the maximum possible welfare gain is large or small and to determine which reforms are well focused. It also allows one to take steps towards designing superior insurance systems simply because properties of solutions to the planning problem are known in advance.

The paper is organized as follows. Section 2 presents the framework. Section 3 sets model parameters. Section 4 and 5 present the main results. Section 6 concludes.

theoretical literature.

II. Framework

A. Preferences

An agent's preferences over consumption and labor allocations over the life cycle are given by a calculation of ex-ante, expected utility.

$$E\left[\sum_{j=1}^{J}\beta^{j-1}u(c_j,l_j)\right]=\sum_{j=1}^{J}\sum_{s^j\in S^j}\beta^{j-1}u(c_j(s^j),l_j(s^j))P(s^j)$$

Consumption and labor allocations are denoted $(c,l)=(c_1,...,c_J,l_1,...,l_J)$. Consumption and labor at age $j=1,...,J$ are functions $c_j:S^j\to R_+$ and $l_j:S^j\to[0,1]$ mapping j-period shock histories $s^j\in S^j$ into consumption and labor decisions. The set of possible j-period histories is denoted $S^j=\{s^j=(s_1,...,s_j):s_i\in S,i=1,...,j\}$, where S is a finite set of shocks. $P(s^j)$ is the probability of history s^j. An agent's labor productivity in period j, or equivalently at age j, is given by a function $\omega(s_j,j)$ mapping the period shock s_j and the agent's age j into labor productivity - effective units of labor input per unit of time worked.

B. Incentive Compatibility

Labor productivity is observed only by the agent. The principal observes the earnings of the agent which equals the product of a wage rate, labor productivity and work time. In this context, the Revelation Principle implies that the allocations (c,l) that can be achieved between a principal and an agent are precisely those that are incentive compatible.[5]

We now define incentive compatible allocations. For this purpose, consider the report function $\sigma\equiv(\sigma_1,...,\sigma_J)$, where σ_j maps shock histories $s^j\in S^j$ into S. The truthful report function σ^* has the property that $\sigma_j^*(s^j)=s_j$ in any period for any j-period history. An allocation (c,l) is *incentive compatible* (IC) provided that the truthful report function always gives at least as much expected utility to the agent as any other feasible report function.[6]

[5]See Mas-Colell, Whinston, and Green (1995, Prop. 23.C.1).

[6]A report function σ is feasible for (c,l) provided (1) $\omega(s_j,j)$ is always large enough to produce the output required by a report (i.e. $0\le l_j(\hat{s}^j)\omega(\sigma_j(s^j),j)\le\omega(s_j,j),\forall j,\forall s^j$, where $\hat{s}^j\equiv(\sigma_1(s^1),...,\sigma_j(s^j))$) and (2) σ maps true histories into reported histories that can occur with positive probability.

The expected utility of an allocation (c, l) under a report function σ is denoted $W(c, l; \sigma, s_1)$.[7] Using this notation, (c, l) is IC provided $W(c, l; \sigma^*, s_1) \geq W(c, l; \sigma, s_1), \forall s_1, \forall \sigma$.

$$W(c, l; \sigma, s_1) \equiv \sum_{j=1}^{J} \sum_{s^j \in S^j} \beta^{j-1} u\left(c_j(\hat{s}^j), \frac{l_j(\hat{s}^j)\omega(\sigma_j(s^j), j)}{\omega(s_j, j)}\right) P(s^j|s_1)$$

$$\hat{s}^j \equiv (\sigma_1(s^1), ..., \sigma_j(s^j))$$

C. Decision Problems

This paper focuses on two decision problems: the U.S. social insurance problem and the planning problem. These problems have the same objective but different constraint sets. V^{us} and V^{pp} denote the maximum ex-ante, expected utility achieved.

$$V^{us} \equiv \max_{(c,l)\in\Gamma^{us}} E\left[\sum_{j=1}^{J} \beta^{j-1} u(c_j, l_j)\right]$$

$$\Gamma^{us} = \{(c, l) : \sum_{j=1}^{J} \frac{c_j}{(1+r)^{j-1}} \leq \sum_{j=1}^{J} \frac{(w\omega(s_j,j)l_j - T_j(x_j, w\omega(s_j,j)l_j))}{(1+r)^{j-1}}$$
$$\text{and } x_{j+1} = F_j(x_j, w\omega(s_j, j)l_j, c_j), x_1 \equiv 0\}$$

$$V^{pp} \equiv \max_{(c,l)\in\Gamma^{pp}} E\left[\sum_{j=1}^{J} \beta^{j-1} u(c_j, l_j)\right]$$

$$\Gamma^{pp} = \{(c, l) : E\left[\sum_{j=1}^{J} \frac{(c_j - w\omega(s_j,j)l_j)}{(1+r)^{j-1}}\right] \leq Cost \text{ and } (c, l) \text{ is IC } \}$$

The constraint set Γ^{us} is specified by a tax function T_j and a law of motion F_j for a vector of state variables x_j. The tax function states the agent's tax payment at age j as a function of period earnings $w\omega(s_j, j)l_j$ and the state variables x_j. Earnings equal the product of a wage rate w per efficiency unit of labor, labor productivity $\omega(s_j, j)$ and work time l_j. Allocations in Γ^{us} have the property that the present value of consumption is no more than the present

[7] $W(c, l; \sigma, s_1)$ is defined only for $\omega(s_j, j) > 0$. Later in the paper, we will set labor productivity to zero beyond a retirement age. It is then understood that labor supply is set to zero at those ages.

value of labor earnings less net taxes for any history of labor-productivity shocks.[8] The next section demonstrates that this abstract formulation can capture important features of the U.S. social security and income tax system.

The constraint set Γ^{pp} for the planning problem has two restrictions. First, the expected present value of consumption less labor income cannot exceed some specified value, denoted $Cost$. We set $Cost$ to the present value of resources extracted from a cohort in a solution to the U.S. social insurance problem: $Cost \equiv E[\sum_{j=1}^{J} -\frac{T_j(x_j, w\omega(s_j,j)l_j^{us})}{(1+r)^{j-1}}]$. As all shocks are idiosyncratic, a known fraction of agents $P(s^j)$ in a cohort receives any shock history $s^j \in S^j$. Thus, while the resources extracted from a single agent over the lifetime is potentially random, the resources extracted from a large cohort is not random. Second, allocations (c, l) must be incentive compatible (IC).

Ex-ante expected utility can be ordered in these problems so that $V^{pp} \geq V^{us}$. The argument is based on showing that if the allocation (c^{us}, l^{us}) achieves the maximum, then (c^{us}, l^{us}) is also in Γ^{pp}. Since (c^{us}, l^{us}) satisfies the present value condition in Γ^{us}, then it also satisfies the expected present value condition in Γ^{pp} by the choice of $Cost$. It remains to argue that (c^{us}, l^{us}) is incentive compatible. However, the fact that (c^{us}, l^{us}) is an optimal choice implies that it is incentive compatible.

D. Model Tax-Transfer System

The tax function and law of motion (T_j, F_j) are now specified to capture features of U.S. social security and federal income taxation. The tax function T_j is the sum of social security taxes T_j^{ss} and income taxes T_j^{inc}. The state variable $x_j = (x_j^1, x_j^2)$ in T_j has two components: x_j^1 is an agent's average earnings up to period j and x_j^2 is an agent's asset holdings.

$$T_j(x_j, w\omega(s_j, j)l_j) = T_j^{ss}(x_j^1, w\omega(s_j, j)l_j) + T_j^{inc}(x_j^1, x_j^2, w\omega(s_j, j)l_j)$$

[8]The constraint set can equivalently be formulated as a sequence of budget restrictions where the agent has access to a risk-free asset, starts life with zero units of this asset and must end life with non-negative asset holding.

1. Social Security

The model social security system taxes an agent's labor income before a retirement age R and pays a social security transfer at and after the retirement age. Specifically, taxes are proportional to labor earnings $(w\omega(s_j, j)l_j)$ for earnings up to a maximum taxable level e_{max}. The social security tax rate is denoted by τ. Earnings beyond the maximum taxable level are not taxed. At and after the retirement age, a transfer $b(x^1)$ is given that is a fixed function of an accounting variable x^1. The accounting variable is an equally-weighted average of earnings before the retirement age R (i.e. $x^1_{j+1} = [\min\{w\omega(s_j, j)l_j, e_{max}\} + (j-1)x^1_j]/j$). The earnings that enter into the calculation of x^1_j are capped at a maximum level e_{max}. After retirement, the accounting variable remains constant at its value at retirement.

$$T^{ss}_j(x^1_j, w\omega(s_j, j)l_j) = \begin{cases} \tau \min\{w\omega(s_j, j)l_j, e_{max}\} & : \quad j < R \\ -b(x^1_j) & : \quad j \geq R \end{cases}$$

The relationship between average past earnings x^1 and social security benefits $b(x^1)$ in the model is shown in Figure 1. Benefits are a piecewise-linear function of average past earnings. Both average past earnings and benefits are normalized in Figure 1 so that they are measured as multiples of average earnings in the economy. The first segment of the benefit function in Figure 1 has a slope of .90, whereas the second and third segments have slopes equal to .32 and .15. The bend points in Figure 1 occur at 0.21 and 1.29 times average earnings in the economy. The variable e_{max} is set equal to 2.42 times average earnings.

We set the bend points and the maximum earnings e_{max} equal to the actual multiples of mean earnings used in the U.S. social security system. We also set the slopes of the benefit function equal to actual values.[9] Figure 1 says that the social security retirement benefit payment is about 45 percent of mean earnings in the economy for a person whose average

[9]In the U.S. social security system, a person's monthly retirement benefit is based on a person's averaged indexed monthly earnings (AIME). For a person retiring in 2002, this benefit equals 90% of the first $592 of AIME, plus 32% of AIME between $592 and $3567, plus 15% of AIME over $3567. Dividing these "bend points" by average earnings in 2002 and multiplying by 12 gives the bend points in Figure 1. Bend points change each year based on changes in average earnings. The maximum taxable earnings from 1998-2002 averaged 2.42 times average earnings. All these facts, as well as average earnings data, come from the Social Security Handbook (2003). The retirement benefit above is for a single-person household. We abstract from spousal benefits.

earnings over the lifetime equals mean earnings in the economy.

Two differences between the model system and the old-age component of the U.S. system are the following:[10]

(i) The accounting variable in the U.S. system is an average of the 35 highest earnings years, where the yearly earnings measure which is used to calculate the average is capped at a maximum earnings level. [11] In the model, earnings are capped at a maximum level just as in the actual system, but earnings in all pre-retirement years are used to calculate average earnings.

(ii) In the U.S. system the age at which benefits begin can be selected within some limits with corresponding actuarial adjustments to benefits. In the model the age R at which retirement benefits are first received is fixed.

2. Income Taxation

Income taxes in the model economy are determined by applying an income tax function to a measure of an agent's income. The empirical tax literature has calculated effective tax functions (i.e. the empirical relationship between taxes actually paid and income).[12] We use tabulations from the Congressional Budget Office (2004, Table 3A and Table 4A) for the 2001 tax year to specify the relation between average effective federal income tax rates and income. Figure 2 plots average effective tax rates for two types of households: head of household is 65 or older and head of household is younger than 65. The horizontal axis in Figure 2 measures income in 2001 dollars. Figure 2 shows that average federal income tax rates increase strongly in income.

In the model economy, we choose income taxes $T_j^{inc}(x_j^1, x_j^2, w\omega(s_j, j)l_j)$ before and after the retirement age R to approximate the average tax rates in Figure 2. We proceed in

[10] We do not try to capture the degree to which the progressivity of the old-age component of social security is mitigated by a positive correlation between survival rates and earnings.

[11] The 35 highest years are calculated on an indexed basis in that indexed earnings in a given year equal actual nominal earnings multiplied by an index. The index equals the ratio of mean earnings in the economy when the individual turns 60 to mean earnings in the economy in the given year. In effect, this adjusts nominal earnings for inflation and real earnings growth.

[12] See, for example, Gouveia and Strauss (1994).

three steps. First, we approximate the data in 2001 dollars with a continuous function. Specifically, we use the quadratic function passing through the origin that minimizes the squared deviations of the tax function from data. This gives average tax functions before and after the retirement age. Second, we express model income in 2001 dollars.[13] Third, the average tax rates on model income are given by the function estimated in the first step after expressing model income in 2001 dollars. Model income equals the sum of labor income $w\omega(s_j, j)l_j$, asset income $x_j^2 r$ and social security transfer income $b_j(x_j^1)$, where initial assets are zero (i.e. $x_1^2 = 0$).

[13]This is done using the ratio between the average U.S. earnings and average model earnings. The figure for average U.S. earnings is $\$32,921$. This comes from the benefit calculation section of the Social Security Handbook (2003).

III. Parameter Values

The results of the paper are based upon the parameter values in Table 1. Model parameters are principally set equal to the values estimated by Kaplan (2007). The goal of Kaplan's work is to understand many dimensions of cross-sectional inequality from the perspective of a standard, incomplete-markets model with endogenous labor supply. Model parameters are estimated to account for the cross-sectional, variance-covariance patterns of hours, consumption and wages at different ages over the life cycle.[14]

One key departure from Kaplan's model is that our tax-transfer system differs. We consider a tax-transfer system that captures features of social security and federal income taxation. Thus, net marginal tax rates will vary with an agent's age and state. Capital and labor taxes in Kaplan's work are proportional taxes that are age and state invariant.[15]

Table 1: Parameter Values

There are $J = 56$ model periods in an agent's life. Retirement occurs at model period $R = 41$. At the retirement age labor productivity is zero and an agent starts collecting social security benefits. One model period corresponds to one year. Thus, we view the agent as starting the working life at a real-life age of 25, retiring at age 65 and dying after age 80.

An agent's labor productivity is $\omega(s_j, j) = \mu_j \; exp(s_j^1 + s_j^2 + s_j^3)$. The wage at age j is determined by a fixed wage rate w per efficiency unit of labor and by labor productivity $\omega(s_j, j)$. Labor productivity is given by a deterministic component μ_j and by an idiosyncratic shock component $s_j = (s_j^1, s_j^2, s_j^3)$ which captures permanent, persistent and temporary sources of productivity differences. The permanent component s^1 stays fixed for an agent over the life cycle and is distributed $N(-\sigma_1^2/2, \sigma_1^2)$. The persistent component follows an autoregressive process $s_j^2 = \rho s_{j-1}^2 + \eta_j$, $\eta_j \sim N(0, \sigma_2^2)$. The temporary component s_j^3 is independent across periods and is distributed $N(-\sigma_3^2/2, \sigma_3^2)$.

[14] Heathcote et al. (2008) analyze a related model with time-varying variances of different components of wages to account for the change in cross-sectional hours, wage, earnings and consumption inequality in the U.S. over time.

[15] There are two other departures. First, we do not allow for heterogeneity in the preference parameters. Second, the working lifetime is 40 years rather than the 38 years in Kaplan (2007). We thank Greg Kaplan for providing his estimates of the mean productivity profile based upon 40 working years.

We consider a benchmark model with only permanent shocks as well as a full model with all three stochastic components. The parameters are set to estimates from Kaplan (2007). A one standard deviation permanent shock leads to about a 24 percent permanent change in wages, whereas a one standard deviation innovation to the persistent component changes wages by about 14 percent. The persistent shock is set to zero for each agent at the beginning of the working life cycle. The deterministic wage component μ_j is given in Figure 3. This component implies that wages approximately double over the life cycle. We approximate each productivity process with a discrete number of shocks.[16]

The period utility function in the model is additively separable $u(c,l) = \frac{c^{(1-\nu)}}{(1-\nu)} + \phi\frac{(1-l)^{(1-\gamma)}}{(1-\gamma)}$. Utility function parameters are set equal to Kaplan's estimates. The coefficient of relative risk aversion is $\nu = 1.66$. The coefficient $\gamma = 5.55$ governs the Frisch elasticity of labor (i.e. $\epsilon_{labor} = \frac{1}{\gamma}\frac{(1-l)}{l}$ so that the Frisch elasticity is 0.27 evaluated at $l = .4$). These values lie well within a range of values estimated in the literature based upon micro-level consumption and labor data - see Browning, Hansen, and Heckman (1999). The value $\phi = 0.13$ is the mean value estimated by Kaplan.

One important restriction on the utility function $u(c,l)$ is the assumption of additive separability. Much of the literature on dynamic contract theory with a labor decision employs this assumption. We make use of this assumption when we design a procedure to compute solutions to the planning problem.[17]

The parameters of the model tax-transfer system are set to capture features of social security and federal income taxation in the U.S. Thus, the social security tax rate τ is set to equal 10.6 percent of earnings. This is the combined employee-employer tax for the old-age and survivor's insurance component of social security. The social security benefit function $b(x)$ and the income tax function T_j^{inc} are given by Figure 1 and Figure 2, which were

[16]We approximate the permanent component with 5 equally-spaced points in logs on the interval $[-\sigma_1^2/2 - 3\sigma_1, -\sigma_1^2/2 + 3\sigma_1]$. Following Tauchen (1986), probabilities are set to the area under the normal distribution, where midpoints between the approximating points define the limits of integration. The persistent component is approximated with 3 equally-spaced points on the interval $[-4\sigma_2, 4\sigma_2]$. Transition probabilities are calculated following Tauchen (1986). The temporary component is approximated with 2 values.

[17]It is used in Theorem A1 in the Appendix to establish which incentive constraints bind and to reduce dimensionality when we compute solutions to the permanent-shock problem.

discussed in the previous section.

The model is explicitly a partial equilibrium model in that wage w per efficiency unit of labor and the real interest rate r are exogenous. They do not vary as we consider alternative social insurance arrangements. Nevertheless, we choose the value of the agent's discount factor β so that a steady state of a general equilibrium version of the full model produces the interest rate $r = .042$ in Table 1. This interest rate is the average of the real return to stock and to long-term bonds over the period 1946-2001 (see Siegel (2002, Tables 1-1 and 1-2)). The value of the wage w in the model is then set to the value consistent with the factor inputs that produce this real return as explained in the Appendix.[18]

Figure 4 displays the evolution of the variance of (log) wages, earnings, work hours and consumption within the full model. The dispersion in wages early in life reflects the sum of the permanent and temporary components of productivity. The rise in wage dispersion with age reflects the role of persistent shocks. The dispersion in earnings over the life cycle closely mimics the pattern for wages. One reason for this is that, absent preference heterogeneity, the model produces little dispersion in work hours. The rise in consumption dispersion over the life cycle reflects mainly the role of persistent shocks. The levels of consumption, earnings and wage dispersion are lower at all ages within the full model compared to the U.S. facts documented in Heathcote, Storesletten, and Violante (2005). This is because Kaplan (2007) analyzes residual dispersion - dispersion after controlling for observable sources of variation such as those related to differences in education - rather than total dispersion. Although the estimate of the permanent wage shock variance is reduced compared to the estimates in Heathcote et al. (2008), the parameters related to persistent and temporary shocks are not greatly affected.

[18]The notion of a steady state and how to compute it is standard and follows Huggett (1996). This involves choosing an aggregate production function and setting factor prices to marginal products. The Appendix describes in detail how this is carried out.

IV. Analyzing Welfare Gains

This section analyzes welfare gains within the permanent-shock model.

A. Maximum Welfare Gains

The maximum welfare gain to improved insurance is measured by the percentage increase α in consumption in the allocation (c^{us}, l^{us}) solving the U.S. social insurance problem so that ex-ante expected utility is the same as in an allocation (c^{pp}, l^{pp}) solving the planning problem.[19] These allocations use the same expected present value of resources. This calculation is shown below. The results of this section are based on computing solutions to each problem. Computational methods are described in the Appendix.

$$E\left[\sum_{j=1}^{J}\beta^{j-1}u(c_j^{us}(1+\alpha), l_j^{us})\right] = E\left[\sum_{j=1}^{J}\beta^{j-1}u(c_j^{pp}, l_j^{pp})\right] \equiv V^{pp}$$

Figure 5 highlights the maximum welfare gains attainable for a range of values of the variance of the permanent component of wage shocks. Figure 5 shows that the welfare gain is increasing in this variance. This is true both when the model social insurance system only includes social security and when the model social insurance system includes both social security and income taxation.

To quantify the size of the maximum welfare gain, we need an estimate of this variance. Kaplan (2007) estimates that $\sigma_1^2 = .056$ for permanent shocks. Thus, a one standard deviation shock increases wages permanently over the lifetime by about 24 percent. Heathcote et al. (2008) estimate a wage process with a similar structure to Kaplan (2007) but find that $\sigma_1^2 = .109$. One reason for this difference is that in a first stage regression Kaplan controls for permanent differences in wages related to education whereas Heathcote et al. do not. It is valuable to keep both estimates in mind in viewing Figure 5a. Using Kaplan's estimate,

[19] When the range of the period utility function of consumption is not bounded from above, then there is always a value α solving this equation. The utility to consumption is bounded above by zero for the period utility function in Table 1. Nevertheless, as Figure 4 highlights, α is well defined for all the examples analyzed.

Figure 5a shows that the maximum welfare gain in the model of the combined social security and income tax system is equivalent to a 4.1 percent increase in consumption each period.

The analysis in Figure 5a is based upon the idea that while earnings are publicly observed both individual hours of work and individual wage rates are only privately observed. This implies that any mechanism determining consumption and labor over the lifetime must respect the incentive compatibility constraints. Figure 5b describes how important private information is for limiting the size of the gains to superior insurance. Figure 5b plots the maximum welfare gain in the economy with social security and income taxation when wage rates are private information and when they are public information. At the value $\sigma_1^2 = .056$, the maximum welfare gain under public information is equivalent to a 6.1 percent change in consumption at each age. This gain is achieved by having all agents of a given age consume the same amount despite large differences in earnings across agents with different productivities.

The remainder of section 4 develops an understanding of what lies behind the patterns in Figure 5. In doing so, the following questions are addressed: (1) How do patterns of lifetime taxation differ in the two problems?, (2) To what degree can welfare be improved by reallocating consumption, fixing the labor allocation?, (3) How do marginal rates of substitution in the model insurance system differ from those in the planning problem? and (4) Why does the welfare gain increase as the shock variance increases?

B. Patterns of Lifetime Taxation

To get a preliminary idea of the economics behind the maximum welfare gains, it is useful to examine patterns of lifetime taxation. Figure 6 graphs the present value of earnings and consumption for agents at each of the five values of the permanent shock. This is done both in the model social insurance system and in the planning problem for the benchmark variance of $\sigma_1^2 = .056$. Figure 6 shows that lifetime taxation is progressive in both allocations in that the ratio of the present value of consumption to the present value of earnings falls as lifetime earnings increase. Furthermore, there is much more progression in lifetime average tax rates in the planning allocation than in the allocation under the model social insurance system. One additional feature of Figure 6 is that both allocations involve extracting resources in

present-value terms from a cohort. This last point is clear as the lifetime tax patterns under the model social insurance system is below the 45 degree line for agents at all permanent shock levels.[20]

A quick look at Figure 6 reveals that the labor allocation must be quite different across these two allocations as the present value of earnings differs sharply. To highlight this, we plot work time over the life cycle. Figure 7 shows that in the planning problem the highest productivity shock agents work the greatest fraction of time and the lowest productivity shock agents work the least. In the model social insurance system this pattern of work time is exactly reversed.

One issue raised by Figures 6 and 7 is the extent to which the maximum welfare gains arise from simply reallocating consumption across agents with different permanent shocks, holding the labor allocation fixed. The remaining gains are related to changing work time. Thus, if it were possible to raise the consumption of low shock agents and lower that of high shock agents, how far would such a reallocation go to improving welfare? While such a reallocation would improve ex-ante utility because the utility function is concave in consumption, this reallocation can only be pushed up to the point where the incentive constraints bind.

To answer this question, we calculate the new allocation (c^*, l^{us}) which maximizes ex-ante utility, holding labor fixed at l^{us}, while imposing incentive compatibility and the present value resource constraint. We find that at the benchmark value $\sigma^2 = .056$ the new allocation (c^*, l^{us}) increases welfare over (c^{us}, l^{us}) by 2.9 percent, compared to a maximum 4.09 percent achieved in the planning problem. Thus, important parts of the maximum welfare gain are due both to reallocating consumption and changing the labor allocation.

C. Analyzing Wedges

We now try to better understand the sources of the welfare gains documented in Figure 5. To do so, we focus on the wedges between marginal rates of substitution and transformation. One wedge is the intratemporal wedge between the consumption-leisure marginal rate of

[20]Intuitively, a pay-as-you-go social security system alone should extract resources from current and future birth cohorts to pay for "free" benefits to previous cohorts. Fullerton and Rogers (1993, Table 4-14) calculate that lifetime average tax rates in the U.S. are roughly progressive in lifetime income and that resources are extracted in present-value terms from the cohorts they analyze.

substitution and the agent's wage. The other wedge is the intertemporal wedge between the marginal rate of substitution of consumption intertemporally and the gross interest rate. We will see shortly that the differences in work hours across the two problems turn out to be related to the differences in the intratemporal wedge.

Consider first the social insurance problem. The income tax system causes the marginal rate of substitution of consumption intertemporally to be below the gross interest rate. In fact, the progressivity of the income tax system, previously documented in Figure 2, implies that within the model the intertemporal wedge is greatest for high productivity agents. These are the agents who end up receiving high incomes.

Consider next the intratemporal wedge. Figure 8 graphs the ratio of the intratemporal marginal rate of substitution to the agents wage for each value of the permanent shock.[21] Any deviation of this ratio from unity will be labeled a wedge.

Within an age group, Figure 8 shows that this wedge increases as an agent's wage and productivity increases. The wedge is smallest for low productivity agents for two reasons. First, these agents have relatively low incomes and marginal income tax rates are relatively low at low income levels. Second, the nature of the social security system implies that at any age the marginal tax rate on additional earnings arising from social security increases as an agent's productivity shock increases.

This second point merits some discussion. The marginal tax rate mentioned above equals the social security tax rate τ less the present value of marginal social security benefits incurred from an extra unit of earnings. This applies to agents who are below the maximum taxable earnings level. This second component differs across agents within the same age group. The reason is that agents in the model will anticipate ending up on different sections of the social security benefit function. High productivity agents will end up on the flat part of the social security benefit function and thus will incur a low marginal benefit in present value. The situation is reversed for low productivity agents as they will end up on the steep part of the benefit function. This reasoning implies that marginal tax rates arising from social security increase with productivity within the model.[22]

[21]Recall from section 3 that the wage rate in the permanent-shock model is $w\omega(s,j) = w\mu_j exp(s^1)$ and that there are five equally-spaced shock values $s_1^1 < s_2^1 < ... < s_5^1$.

[22]A previous version of this paper calculated how the marginal tax rate arising from the model social

We now analyze the nature of wedges that arise in a solution to the planning problem. Solutions to the planning problem will involve some incentive compatibility constraint binding. As a consequence, at a solution it will not be true that all marginal rates of substitution are equated to marginal rates of transformation.

While there is an intertemporal wedge in the model social insurance problem arising from the income tax there is no intertemporal wedge in a solution to the planning problem. This difference accounts for some of the welfare gains. To see why there is no intertemporal wedge in the planning problem, assume that there is a solution with a wedge. If so, then it is possible to deliver both the same expected utility and the same ex-post utilities at lower expected present value cost, without changing the labor allocation. This can be done by eliminating the intertemporal wedge. The extra resources saved can then be used to make a uniform increase in utility to agents receiving all shocks while preserving incentive compatibility.[23]

Now consider the intratemporal wedge. The intratemporal marginal rate of substitution will differ from an agent's wage rate in a solution to the planning problem depending on which incentive constraints bind. It turns out that only the local downward incentive constraints hold with equality in a solution. These constraints require that an agent with a given permanent shock weakly prefers his/her own allocation to the allocation received by pretending to have the next lowest shock. An important consequence of this (see Theorem A1 in the Appendix) is that the marginal rate of substitution between consumption and labor is then strictly below the wage rate $w\omega(s,j)$ in all periods for all agents except the agent receiving the highest shock.[24] For the agent with the highest shock, there is no gain to distorting the consumption-labor margin at any age. The reason is that no other agent envies the consumption and output allocation of this agent. All other agents get strictly lower lifetime utility by pretending to be the high shock agent and allocating enough labor

security system varied with age for a median productivity agent. Early in life the marginal tax rate is slightly below $\tau = .106$. It decreases with age but remains positive at all ages. Broadly, our results are similar to the marginal social security tax rates calculated by Feldstein and Samwick (1992, Table 1).

[23]Rogerson (1985) and Golosov, Kocherlakota, and Tsyvinski (2003) present a necessary condition on this margin in planning problems with a more general structure of shocks. Their main result is the "inverse" Euler equation. The result stated in the text is a special case of their result as the inverse Euler equation reduces to the claim made above, absent period-by-period shocks. With period-by-period shocks, a solution to the planning problem will have an intertemporal wedge.

[24]A similar result holds in the one-period model studied by Mirrlees (1971).

time to produce the higher output required.

Next, we examine the size of the intratemporal wedge. Figure 9 graphs the ratio of the marginal rate of substitution to the agent's wage rate at each age for each of the five possible values of the permanent shock. Figure 9 shows that the intratemporal wedge is positive for all agents with the exception of the agent with the highest permanent shock. Furthermore, within an age group the magnitude of this wedge decreases as an agent's wage increases.

In the context of the permanent-shock model, we are not aware of any existing theoretical result which describes how the wedge at each age moves as productivity increases. However, for the static Mirrlees model there are theoretical and computational results (see, for example, Tuomala (1990), Saez (2001) and the references cited in these papers). In the Mirrlees model, the lognormal distribution of productivity is important for wedges to decline as productivity increases. We have computed the nature of wedges in the permanent-shock model when we replace the lognormal distribution with a Pareto distribution. The literature has argued that the upper tail of the earnings distribution has fat tails which are more in line with a Pareto distribution. For the Pareto distribution with the same mean and variance, we find that wedges do not decrease as productivity increases.[25]

We conjecture that the differences in wedges and the differences in lifetime taxation are the key reasons why the maximum welfare gains increase as labor productivity risk increases. There is too little progression in lifetime taxation in the model social insurance system compared to the planners problem as risk increases. Furthermore, the intratemporal wedge on high productivity agents typically increases as risk increases in the model social insurance system whereas the wedge on the highest productivity agents within an age group is always zero in the planning problem.

[25]Following Tauchen (1986), we approximate a Pareto distribution with five equally-spaced points one standard deviation apart. The resulting wedge is positive and displays little variation across ages. The wedge for the lowest four shock levels averages approximately $.12, .10, .16$ and $.20$, in order of increasing productivity. The wedge for the highest productivity level is approximately zero in computations.

V. Reforming the Social Insurance System

We examine two ways to reform the model social insurance system. Reform 1 is a piecemeal reform in which a component of the social insurance system is changed while maintaining the remainder of the system. In Reform 1 we change the social security benefit function without changing income taxation or the social security tax rate. Reform 2 is a radical reform as social security and income taxation are eliminated and are replaced with a tax on the present value of earnings.

Reform 1 and 2 are optimal parametric reforms. In each case we search over the parameters of the respective tax functions to find the parameter vector which maximizes ex-ante expected utility of the cohort of agents.[26] In each reform the same present value of resources is extracted from the cohort as in the original social insurance system. The Appendix describes computational methods. The Appendix is also useful for understanding how to achieve a tax on the present value of earnings using a period-by-period tax system. We note that a present-value tax is compatible with the provision of retirement benefits as such a tax can be achieved with very different timings of taxes and transfers over the lifetime.

A. Motivation

The policy literature is full of discussions of piecemeal reforms. In the social security literature, it is common to find the suggestion that the value of marginal social security benefits incurred by extra earnings should be more closely linked with marginal taxes paid in order to improve efficiency or a welfare measure. These considerations motivate the analysis of Reform 1 which is an optimal piecemeal reform that flexibly changes the benefit function.

The motivation for Reform 2 is that it is simple and that there are reasons to think that it might work well within the permanent-shock model. Within the permanent-shock model, a present-value tax has two important properties. First, it imposes no intertemporal wedge. Second, it imposes an age-invariant wedge on the intratemporal margin that can be made to

[26] Our analysis of optimal parametric reforms is similar in some respects to the work of Conesa, Kitao, and Krueger (2009). They choose the parameters of a labor income tax function and a linear capital income tax to maximize ex-ante lifetime utility in steady state.

flexibly differ across agents.[27] The previous section argued that the first property holds in a solution to the planning problem and that the second property is approximately supported in computations.

B. Analysis

The welfare gain to each reform is given in Table 2. Welfare gains are stated in terms of the permanent percentage increase in consumption in the allocation in the model without the reform which is equivalent to the expected utility delivered under the optimal reform. Welfare gains are calculated for both the full model (i.e the model with permanent, persistent and temporary shocks) and the permanent-shock model.

Table 2: Welfare Gains to Optimal Parametric Reforms

We first discuss the results for the permanent-shock model. For Reform 1, we calculate the best constant benefit, the best linear benefit and the best quadratic benefit as a function of average lifetime earnings. The best constant benefit function in the permanent-shock model leads to a welfare gain of 0.14 percent. A constant social security benefit increases the progressivity of lifetime earnings taxation but also increases marginal earnings taxes across earnings levels. The best linear benefit function has a positive intercept and a negative slope and leads to a welfare gain of 0.18 percent. The best quadratic benefit function that we find does not improve welfare over the best linear function. This class of reforms achieves only a small fraction of the maximum possible welfare gain. This occurs because these reforms are poorly focused: greater progression in lifetime taxation is achieved by imposing an even larger intratemporal wedge on high productivity agents and the change in the benefit function does not eliminate the wedge on the intertemporal consumption margin.

In contrast, an optimal present-value tax leads to a large welfare gain worth a 3.95 percent increase in consumption. We obtain this result when the class of tax functions are increasing

[27]Werning (2007) shows that a present-value tax system is optimal in some contexts. Specifically, he shows that such a tax implements a solution to a planning problem in the context of an infinitely-lived agent model where labor productivity takes on two possible values, labor productivity is private information and preferences are of the constant Frisch elasticity of labor form.

step functions. This reform achieves nearly all of the maximum possible welfare gain in the permanent-shock model.

We highlight two reasons why the optimal present-value earnings tax works well in the permanent-shock model. First, it allows for a flexible choice of lifetime taxation. Indeed, the graph of the present value of consumption as a function of the present value of earnings which turns out to be optimal is essentially the pattern in the planning problem - previously displayed in Figure 6. Second, the present-value tax is able to closely approximate the pattern of intratemporal and intertemporal wedges found in a solution to the planning problem.[28]

We now discuss results for the full model. For Reform 1, the best constant, linear, and quadratic benefit functions lead to gains worth a $0.56, 1.07$ and 1.15 percent increase in consumption, respectively. The best quadratic benefit function has a positive intercept, but negative values for the coefficients on the slope and quadratic terms. Thus, the piecemeal reform that maximizes ex-ante welfare does not involve more closely linking the value of marginal benefits received to marginal taxes paid. Greater progression in lifetime taxation is achieved within this reform by increasing intratemporal wedges. For Reform 2 we find that in the full model the best present-value tax that is within the piecewise-linear class leads to a small welfare loss equivalent to a 0.07 percent decrease in consumption. Thus, even though a present-value tax is both a simple and well-focused reform within the permanent-shock model, this class of reforms does not lead to welfare gains within the richer idiosyncratic shock structure of the full model.

To get some insight into what is behind these results, we first examine the pattern of lifetime taxation. Figure 10 shows that the progression in lifetime taxation is greater in Reform 1 and Reform 2 than in the benchmark model.[29] Moreover, the pattern of lifetime

[28]At a deeper level, a present-value tax may work well within these economies for two quite different reasons. First, one might conjecture that interior solutions to the planning problem with (i) constant Frisch elasticity of labor preferences (i.e. $u(c_j, l_j) = u(c_j) + \phi\frac{l_j^{1+\gamma}}{1+\gamma}$) and (ii) permanent proportional productivity differences have the property that only local downward incentive constraints bind. If so, such allocations can always be implemented by a present-value tax system. A key property of such a solution, given assumptions (i)-(ii), is that the intratemporal wedge is age invariant - see the proof of Theorem A1(iii) in the Appendix. Second, the preferences used in Table 1 may effectively be close to those with constant Frisch elasticity of labor.

[29]The 10th, 50th and 90th percentile of the present value of earnings distribution in the benchmark model occur at values 10.7, 17.4 and 26.1 in Figure 10.

taxation is broadly similar in both reforms over much of the domain. So the difference in welfare gain between Reform 1 and Reform 2 does not seem to come from differences in this measure of tax progression. The optimal present-value tax function in Reform 2 is roughly linear over most of the domain but is eventually flat well past the 99th percentile of the distribution - this occurs at a present value of earnings equal to 45.

We now describe how the reforms impact consumption. Both reforms produce a downward shift in the distribution of the present value of earnings compared to the benchmark model. The result is that mean consumption at almost all ages is lower in both reforms than in the benchmark model but, perhaps surprisingly, only Reform 1 substantially reduces measures of the dispersion in consumption at all ages compared to the benchmark model. This implies that the component of expected utility due to consumption is slightly lower in Reform 1 compared to the benchmark model but is even lower in Reform 2 compared to Reform 1 or to the benchmark model.

Next we describe how the reforms impact work hours. Reform 1 reduces the mean hours of work at all ages compared to the benchmark model and it produces about the same coefficient of variation in hours at all ages. Thus, the ex-ante utility from leisure is greater in Reform 1 than in the benchmark model. Reform 2 reduces mean hours of work at all ages below that in the benchmark model and below that in Reform 1. However, Reform 2 nearly doubles the coefficient of variation of hours early in the life cycle compared to the benchmark model. The overall effect of Reform 2 is to increase the ex-ante utility from leisure compared to the benchmark model. Both reforms increase the correlation between work hours and labor productivity at all ages compared to the benchmark model. Figure 10 suggests that different income effects on high and low lifetime earnings agents is partly behind the increase in correlation. This increase in correlation is a key part of the mechanism within the permanent-shock model for achieving the maximum possible welfare gain.

We now consider Reform 3 to determine if an important part of the welfare gain obtained by Reform 2 in the permanent-shock model comes simply from eliminating capital income taxation and the associated intertemporal wedge. Reform 3 is a piecemeal reform that maintains social security and income taxation but exempts capital income from entering into taxable income. An additional proportional labor income tax is added to satisfy the

present-value resource constraint. Eliminating capital income taxation in this way produces a welfare gain of 0.22 percent in the permanent-shock model and a welfare loss of -0.22 percent in the full model. Thus, simply eliminating intertemporal wedges in this crude way, without substantially increasing the progressivity of lifetime taxation or altering the pattern of intratemporal wedges, does not go very far towards producing the maximum welfare gain in the permanent-shock model.

All of the analysis in the paper is based upon the assumption that factor prices are fixed and do not change as the social insurance system is changed. We now take a step towards determining how a closed-economy analysis might differ by simply calculating how the aggregate capital and labor evolve over time at fixed factor prices within the full model. We assume that each reform applies only to each successive cohort of newborn agents and that all other agents who are alive at the start of the reform face the original social insurance system. The original social insurance system was calibrated to be consistent with a steady state in the full model with no government debt. We view any change in the capital-labor ratio over time as reflecting a need for factor prices to adjust in a closed-economy analysis. An increase in the ratio is viewed as a force which depresses the interest rate and raises the wage rate.

In Reform 1 the capital-labor ratio changes by well under one percent over the first 40 periods. In contrast, Reform 2 and 3 show much larger movements. After 40 periods this ratio falls by 10 percent in Reform 2 and increases by 18 percent in Reform 3. This is due almost entirely to the movement in the numerator - total asset holdings less government debt. We conjecture that little of the welfare gains we find for Reform 1 would vanish in a closed-economy analysis simply because the large effects at the individual level wash out almost entirely for factor inputs both within age group and at the aggregate level. It is less clear whether or not the results for Reform 2 and 3 would continue to hold.

In closing this section, we think that finding parametric tax systems that work well within the full model is a useful problem. This problem connects the policy literature to the literature on optimal taxation. We acknowledge that the tax systems that we have explored can be improved upon as both reforms violate the inverse Euler equation which is a necessary

condition on the intertemporal margin for a solution to the planning problem.[30] Further theoretical and computational work that give insight into wedges arising in planning problems would be useful for finding parametric tax systems that produce larger welfare gains.

[30]Rogerson (1985) and Golosov et al. (2003) present the inverse Euler equation result. Kocherlakota (2005) provides an implementation theorem for solutions to planning problems using this result.

VI. Conclusion

The question of whether to or how to fundamentally redesign social security systems has been and continues to be a major policy issue in the U.S. and in many other countries. One's position on this issue is likely to depend upon one's view of the rationale for social security and for social insurance more broadly. One standard rationale is the provision of insurance for risks that are not easily insured in private markets.

We provide a quantitative analysis of the U.S. social insurance system within a framework with important idiosyncratic, labor-market risks. We find that large welfare gains to changing the social insurance system are possible. Systems that can achieve such welfare gains need not be more complicated than the current U.S. system. Specifically, we find that an optimal tax on the present value of earnings does this within the model with only permanent shocks and that changing only the social security benefit function does this within the model with permanent, persistent and purely temporary productivity shocks of the nature found in U.S. wage rate data. These results are based upon maximizing ex-ante utility for a cohort. Thus, the objective reflects an insurance role both for productivity differences present at the start of the working lifetime as well as for productivity shocks occuring throughout the working lifetime.

We mention three directions to pursue in future work. First, it would be valuable to know quantitative properties of the solution to the planning problem within the full model. This would require important theoretical and/or computational advances.[31] Second, this paper treats labor productivity as being unaffected by the social insurance system. We expect that human capital models (e.g. Huggett, Ventura, and Yaron (2007)) will be central both as positive models of inequality and as models for the analysis of social insurance issues. Because skill acquisition responds to policy in human capital models, labor productivity will not be policy invariant. Whether the gains to adopting superior systems are even larger within such models is an open question. Third, future work might expand the analysis of the social insurance system to go beyond income taxation and social security as well as provide a closed-economy analysis to complement the open-economy analysis pursued in this work.

[31]Fernandes and Phelan (2000) provide a recursive formulation of a planning problem with persistent shocks. Such a formulation is not computationally viable for the full model described in Table 1.

Appendix A

The Appendix contains two sections. Section A.1 describes our methods for computing solutions to the planning problem, the social insurance problem and the parametric planning problems. Section A.2 proves Theorem A1. In the Appendix the labor-productivity function is sometimes set to $\omega(s_j, j) = s_j$ solely to shorten and simplify expressions. FORTRAN programs that compute solutions to all the problems analyzed in the paper are available upon request.

A.1. Computation

A.1.1. Social Insurance Problem

The social insurance problem is stated below as a dynamic programming problem. This involves reformulating the present value budget constraint as a sequence of budget constraints where resources are transferred across periods with a risk-free asset. Risk-free asset holding must then always lie above period and shock specific borrowing limits $\underline{a}_j(s)$ consistent with solvency at the terminal age. The state variable is $x = (a, s, z)$, where a is asset holdings, s is the period shock vector determining productivity and z is average past earnings. The functions T_j and F_j describe the tax system and the law of motion for average past earnings. The shock is Markovian with transition probability $\pi(s'|s)$.

$V_j(a, s, z) = \max_{(c,l,a')} u(c, l) + \beta \sum_{s'} V_{j+1}(a', s', z')\pi(s'|s)$
(1) $c + a' \leq a(1 + r) + w\omega(s, j)l - T_j(x, w\omega(s, j)l)$
(2) $c \geq 0, a' \geq \underline{a}_j(s); l \in [0, 1]$
(3) $z' = F_j(z, w\omega(s, j)l)$

This problem is solved computationally by backwards induction. The value function V_j is computed at selected grid points (a, s, z) by solving the right-hand-side of Bellman's equation. We use the simplex method (see Press et al (1994)). Evaluating the right-hand-side of Bellman's equation involves a bi-linear interpolation of the function $V_{j+1}(a', s', z')$ over the asset and average past earnings dimensions: (a', z'). We set the borrowing limit to a fixed value $\underline{a}$ in each period. We then relax this value so that it is not binding. This is a device for imposing period and state specific limits $\underline{a}_j(s)$. To use this device, penalties are imposed for states and decisions implying negative consumption.[32]

We compute ex-ante, expected utility V^{us} and the expected cost, denoted $Cost$, of running the social insurance system by simulation, under the assumption that an agent starts out with no assets. Specifically, we draw a large number (100,000) of lifetime labor-productivity profiles, compute realized utility and realized cost for each profile, using the computed optimal decision rules, and then compute averages. The same 100,000 histories are used in the calculation of expected utility and expected cost in the analysis of reforms.

A.1.2. Steady State Calibration

We calibrate the discount factor β using the algorithm below. This algorithm is based on computing a stationary equilibrium. To set up this framework, we assume that (i) there is an aggregate production

[32]The backward induction procedure takes as given a value for average earnings in the economy. This value is used to determine the tax function T_j. Thus, an additional loop is needed so that guessed and implied values of average earnings coincide.

function $Y = F(K, L) = K^{\alpha}L^{1-\alpha}$ stated in terms of aggregate capital K and labor L, (ii) physical capital depreciates at rate δ and (iii) population growth is n.

We define an equilibrium using the recursive language - see Huggett (1996). To keep track of agent heterogeneity, we use probability measures ψ_j to describe the fraction of age j agents that have a state vector $x = (a, s, z)$ lying in particular subsets of the state space X. The relative size of different age cohorts is given by ϕ_j, where $\phi_{j+1} = \phi_j/(1+n)$ and $\sum_j \phi_j = 1$. Denote aggregate capital, labor and government spending and consumption (K, L, G, C): $K \equiv \sum_j \phi_j \int a d\psi_j$, $L \equiv \sum_j \phi_j \int \omega(s, j) l(x, j) d\psi_j$ and $C \equiv \sum_j \phi_j \int c(x, j) d\psi_j$. The probability measures must be consistent with one another. This is captured by the recursion $\psi_{j+1} = \Gamma_j(\psi_j)$, where $\Gamma_j(\psi_j)(\cdot) \equiv \int P(x, j, \cdot) d\psi_j$, and P is a transition function induced by the transition probabilities on shocks and by the period j decision rules. We do not write down all the details associated with the construction of this transition function partly because the algorithm below calculates the relevant integrals by simulating a large number of histories rather than by calculating probability measures on a rich collection of subsets of the state space and then integrating. However, details of how to do so are in Huggett (1996).

Definition: A stationary equilibrium is $(c(x, j), l(x, j), a(x, j), w, r, G)$*, tax-transfer functions* $(T_1, ..., T_J)$ *and probability measures* $(\psi_1, ..., \psi_J)$ *such that*

1. *(c,l,a) solve Bellman's equation (Appendix A.1.1), given* (w, r) *and* T_j*.*
2. $w = F_2(K, L)$ *and* $r = F_1(K, L) - \delta$
3. $\psi_{j+1} = \Gamma_j(\psi_j), \forall j$
4. $G = \sum_j \phi_j \int T_j(x, w\omega(s, j) l(x, j)) d\psi_j$
5. $C + K(n + \delta) + G = F(K, L)$

Algorithm:

1. Fix $(\alpha, \delta, n) = (.33, .06, .01)$.
2. Set $r = .042$ and $w = 1.19461$. Given (r, α, δ), equilibrium condition 2 pins down the wage w at the value stated and pins down the capital-labor ratio K/L.
3. Guess the discount factor and average earnings $(\beta, \bar{e})$.
4. Compute decision rules (c, l, a) solving Bellman's equation, given the information in steps 1-3 using the procedures described in Appendix A.1.1.
5. Calculate implied values of aggregates $(K', L', \bar{e}', \sum_j \phi_j \int T_j(x, w\omega(s, j) l(x, j)) d\psi_j)$ via simulation using the decision rules.
6. If $K'/L' = K/L$, $\bar{e}' = \bar{e}$ and $\sum_j \phi_j \int T_j(x, w\omega(s, j) l(x, j)) d\psi_j > 0$, then stop. Otherwise, update $(\beta, \bar{e})$ and repeat steps 4-5.

Comments:

1. We compute β for the full model at the parameters listed in Table 1 and fix this value for all subsequent analysis.

2. The initial value of β in step 3 is set to $\beta = 1/(1 + r)$. In carrying out this algorithm we first adjust average earnings $\bar{e}$ in steps 3-6 until $\bar{e}' = \bar{e}$. The value of β is increased until step 6 approximately holds. We choose $\bar{e}$ in step 3 because the tax-transfer function is only specified once $\bar{e}$ is known - see section 2.4.1.

A.1.3. Planning Problem

We show how to compute V^{pp} for the case of permanent shocks, given the value of $Cost$. The strategy is to analyze the Relaxed Problem. The Relaxed Problem is the same as the planning problem with permanent shocks except that only the local downward incentive constraints are imposed rather than all the incentive constraints. The local downward incentive constraints are the constraints stating that truth telling from shock s dominates claiming to be one shock lower, denoted s^-. There are N shock values that are ordered $s_1 < s_2 < ... < s_N$. Below, we let $\omega(s_j, j) = s_j$ solely to shorten and simplify expressions.

Relaxed Problem: $\max_{(l_j(s),c_j(s))} \sum_s [\sum_j \beta^{j-1}(u(c_j(s)) + v(l_j(s)))]P(s)$
subject to

(i) $\sum_s [\sum_j (c_j(s) - wl_j(s)s)/(1+r)^{j-1}]P(s) \leq Cost$
(ii) $\sum_j \beta^{j-1}(u(c_j(s)) + v(l_j(s))) \geq \sum_j \beta^{j-1}(u(c_j(s^-)) + v(l_j(s^-)s^-/s))), \forall s > s_1$

The strategy is to compute solutions to the Relaxed Problem and to verify that at the computed solution all incentive constraints hold. We compute solutions to the Relaxed Problem by solving the Equivalent Problem below. The Equivalent Problem is useful as it reduces the dimension of the control variables. The claimed equivalence follows from several facts about solutions to the Relaxed Problem. Specifically, at a solution (i) the cost constraint must hold with equality, (ii) consumption is chosen without intertemporal distortion (i.e. $u'(c_j(s)) = \beta(1+r)u'(c_{j+1}(s)), \forall j, s$) and (iii) all local downward incentive constraints bind. As the first result is straightforward, we only formally state the last two in Theorem A1. Theorem A1 also provides an additional theoretical insight. Specifically, since the Lagrange multipliers on the incentive constraints are strictly positive, the Kuhn-Tucker conditions imply that at a solution the intratemporal marginal rate of substitution is strictly below labor productivity for all agents at any age except for the agent with the highest productivity shock. This is a generalization of a standard result for the one-period Mirrlees problem.

Theorem A1: Assume $u(c,l) = u(c) + v(l)$, u and v are concave and differentiable, u and v are strictly increasing and decreasing respectively. At an interior solution to the Relaxed Problem the following hold:

(i) all local downward incentive constraints bind,

(ii) $\dfrac{u'(c_j(s))}{\beta u'(c_{j+1}(s))} = 1 + r, \forall j, \forall s$

(iii) $-\dfrac{v'(l_j(s))}{u'(c_j(s))} < ws, \forall j, \forall s < s_N$ *and* $-\dfrac{v'(l_j(s))}{u'(c_j(s))} = ws, \forall j$ *and for* $s = s_N$.

Proof: See Appendix A.2.

In the Equivalent Problem the choice variables are labor and the lifetime utility of consumption $u(s)$. The cost constraint makes use of the function $COST$. $COST(u)$ describes the resource cost of obtaining lifetime utility u from consumption, given that $u'(c_j(s)) = \beta(1+r)u'(c_{j+1}(s))$.[33] As all constraints are equality

[33] When $u(c) = c^{1-\rho}/(1-\rho)$ and $\rho \neq 1$, then $COST(u) = (\sum_j a^{j-1})[\frac{(1-\rho)u}{\sum_j b^{j-1}}]^{1/(1-\rho)}$, where $a = \frac{[\beta(1+r)]^{1/\rho}}{1+r}$ and $b = \beta[\beta(1+r)]^{(1-\rho)/\rho}$.

constraints, it is also possible to reduce dimensionality further by solving these constraints to express lifetime utility of consumption $u(s)$ as a function of all labor profiles l and $Cost$ as follows: $u(s) = g(l, s, Cost)$.

We use the simplex method from Press et al (1994) to solve the Equivalent Problem. This involves maximizing over $(l_1(s), ..., l_{R-1}(s))$, where R is the retirement period. These choices lie in an $(R-1) \times N$ dimensional space as there are $R-1$ labor periods and N possible permanent shocks.

Equivalent Problem: $\max_{(u(s), l_j(s))} \sum_s [u(s) + \sum_j \beta^{j-1} v(l_j(s))] P(s)$

subject to

(i) $\sum_s [COST(u(s)) - \sum_j wsl_j(s)/(1+r)^{j-1}] P(s) = Cost$

(ii) $u(s) + \sum_j \beta^{j-1} v(l_j(s)) = u(s^-) + \sum_j \beta^{j-1} v(l_j(s^-)s^-/s), \forall s > s_1$

A.1.4. Optimal Parametric Planning Problems

We examine a number of parametric tax systems. For any parametric tax system we choose the parameters of these tax systems to maximize ex-ante utility, given that agents behave optimally and that the present value resource constraint cannot be violated. We describe how we compute the optimal parametric tax system for the case of a tax on the present value of earnings. The computation of other optimal parametric tax systems is similar.

The agent's problem and the planner's problem are described below. The agent's state variable is $x = (a, s, pv)$, where pv is the present value of earnings earned from previous periods. The tax function T_j maps the present value of earnings from previous periods and earnings in period j into the tax paid or transfer received in period j. T_j depends upon a parameter vector α. Solutions to the agent's problem are computed using the methods from section A.1.1.

$V_j(a, s, pv; \alpha) = \max_{(c,l,a')} u(c, l) + \beta \sum_{s'} V_{j+1}(a', s', pv'; \alpha) \pi(s'|s)$
(1) $c + a' \leq a(1+r) + w\omega(s, j)l - T_j(pv, w\omega(s, j)l; \alpha)$
(2) $c \geq 0, a' \geq \underline{a}_j(s); l \in [0, 1]$
(3) $pv' = pv + \frac{w\omega(s,j)l}{(1+r)^{j-1}}$

Parametric Planning Problem: $\max_\alpha E[V_1(0, s, 0; \alpha)]$
subject to

$$E[\sum_j \frac{c_j(s^j;\alpha) - w\omega(s_j, j) l_j(s^j;\alpha)}{(1+r)^{j-1}}] \leq Cost$$

In the planner's problem the only constraint facing the planner, given agent's choices to any tax system are optimal, is the cost constraint. This is because the allocation induced by a solution to the agent's problem is incentive compatible. We compute solutions to the planner's problem by (i) drawing α, (ii) computing optimal decision rules solving the agent's problem, given α and (iii) simulating these decision rules to determine whether or not the resource constraint is violated at the allocation induced by α. We use the simplex method to search over the space of parameters describing the tax function to maximize the objective function. The objective function is ex-ante utility less a penalty term when the cost constraint is violated.

We now describe how we choose the tax function T_j in the agent's problem. Start with a tax function $T(pv; \alpha)$ mapping the present value of realized earnings over the lifetime into the present value of taxes paid

over the lifetime. Define the period tax function $T_j(pv, w\omega(s,j)l;\alpha)$ as indicated below. The tax paid in period j is based on the increment added to the present value of earnings. By the end of the working lifetime, the present value of taxes paid is simply $T(pv;\alpha)$, where pv is the realized present value of earnings over the working lifetime. This is one way to carry out a present-value tax T with a period-by-period tax system T_j for $j = 1, ..., J$.[34]

$$T_j(pv, w\omega(s,j)l;\alpha) = \begin{cases} [T(pv + \frac{w\omega(s,j)l}{(1+r)^{j-1}};\alpha) - T(pv;\alpha)](1+r)^{j-1} & : \quad j \geq 2 \\ T(w\omega(s,j)l;\alpha) & : \quad j = 1 \end{cases}$$

In our numerical implementation, we focus on two classes of parametric functions T. We use the class of piecewise-linear functions for the full model and the class of increasing step functions T for the permanent-shock model. We choose as many steps as there are permanent shocks.

A.2. Theorem A1

Theorem A1: Assume $u(c,l) = u(c) + v(l)$, u is strictly concave and differentiable, v is concave and differentiable, u and v are strictly increasing and decreasing respectively. At an interior solution to the Relaxed Problem the following hold:

(i) all local downward incentive constraints bind,

(ii) $\dfrac{u'(c_j(s))}{\beta u'(c_{j+1}(s))} = 1 + r, \forall j, \forall s$

(iii) $-\dfrac{v'(l_j(s))}{u'(c_j(s))} < ws, \forall j, \forall s < s_N$ *and* $-\dfrac{v'(l_j(s))}{u'(c_j(s))} = ws, \forall j$ *and for* $s = s_N$.

Proof:

(i) We study the Lagrange function below. Let $\gamma(s)$ denote multipliers on the local downward incentive constraints and λ denote the multiplier on the resource constraint. A superscript $+$ or $-$ denotes one higher or lower shock, respectively.

$$L = \textstyle\sum_s [\sum_j \beta^{j-1}(u(c_j(s)) + v(l_j(s)))]P(s) + \lambda[Cost - \sum_s [\sum_j (c_j(s) - wl_j(s)s)/(1+r)^{j-1}]P(s)]$$
$$+ \textstyle\sum_{s>s_1} \gamma(s) \sum_j \beta^{j-1}[u(c_j(s)) + v(l_j(s)) - u(c_j(s^-) - v(l_j(s^-)s^-/s)]$$

At an interior solution the Kuhn-Tucker conditions $dL/dc_j(s) = 0$ and $dL/dl_j(s) = 0$ hold:

$$\frac{dL}{dc_j(s)} = \begin{cases} \beta^{j-1}u'(c_j(s))[P(s) - \gamma(s^+)] - \lambda P(s)/(1+r)^{j-1} & s = s_1 \\ \beta^{j-1}u'(c_j(s))[P(s) - \gamma(s^+) + \gamma(s)] - \lambda P(s)/(1+r)^{j-1} & s_1 < s < s_N \\ \beta^{j-1}u'(c_j(s))[P(s) + \gamma(s)] - \lambda P(s)/(1+r)^{j-1} & s = s_N \end{cases} \quad (1)$$

[34] Vickrey (1939) discusses some mechanics for a period-by-period tax system where taxes paid are based upon an average of past years incomes.

$$\frac{dL}{dl_j(s)} = \begin{cases} \beta^{j-1}v'(l_j(s))\left[P(s) - \gamma(s^+)\dfrac{v'(l_j(s)s/s^+)s/s^+}{v'(l_j(s))}\right] + \frac{\lambda wsP(s)}{(1+r)^{j-1}} & s = s_1 \\ \beta^{j-1}v'(l_j(s))\left[P(s) - \gamma(s^+)\dfrac{v'(l_j(s)s/s^+)s/s^+}{v'(l_j(s))} + \gamma(s)\right] + \frac{\lambda wsP(s)}{(1+r)^{j-1}} & s_1 < s < s_N \\ \beta^{j-1}v'(l_j(s))[P(s) + \gamma(s)] + \frac{\lambda wsP(s)}{(1+r)^{j-1}} & s = s_N \end{cases} \tag{2}$$

Claims 1-4 establish that in a solution to the Kuhn-Tucker conditions all multipliers on incentive constraints are strictly positive: $\gamma(s) > 0, \forall s$. Theorem A1(i) follows from this result.

Claim 1: For $N \geq 2, \gamma(s_N) > 0$.
Claim 2: For $N > 2, \gamma(s^-) = \gamma(s) = 0$ for any s is impossible.
Claim 3: For $N > 2, \gamma(s^-) > 0, \gamma(s) = 0$ for any s is impossible.
Claim 4: For $N > 2, \gamma(s_2) = 0, \gamma(s_3) > 0$ is impossible.

Proof of Claim 1: If $\gamma(s_N) = 0$, then $dL/dc_j(s) = 0$ and u strictly concave implies $c_j(s_N) \leq c_j(s_{N-1}), \forall j$. If $\gamma(s_N) = 0$, then $dL/dl_j(s) = 0$ and v concave implies $l_j(s_N) > l_j(s_{N-1}), \forall j$. Thus, the downward incentive constraint for the agent with shock s_N is violated.

Proof of Claim 2: Suppose $\gamma(s^-) = \gamma(s) = 0$ for some s. Let s be the greatest s such that this holds. Claim 1 implies that $\gamma(s^+) > 0$. Then $dL/dc_j(s) = 0$ and u concave implies $c_j(s^-) > c_j(s), \forall j$. $dL/dl_j(s) = 0$ and v concave implies $l_j(s^-) < l_j(s), \forall j$. Thus, the downward incentive constraint for the agent with shock s is violated.

Proof of Claim 3: Suppose $\gamma(s^-) > 0, \gamma(s) = 0$ for some s. Let s be greatest s such that this holds. Claim 1 and 2 imply $\gamma(s^+) > 0$. Then $dL/dc_j(s) = 0$ and u concave implies $c_j(s^-) > c_j(s), \forall j$. $dL/dl_j(s) = 0$ and v concave implies $l_j(s^-) < l_j(s), \forall j$. Thus, the downward incentive constraint for the agent with shock s is violated.

Proof of Claim 4: Suppose $\gamma(s_2) = 0, \gamma(s_3) > 0$. Then $dL/dc_j(s) = 0$ and u concave implies $c_j(s_1) > c_j(s_2), \forall j$. $dL/dl_j(s) = 0$ and v concave implies $l_j(s_1) < l_j(s_2), \forall j$. This violates the downward incentive constraint for the agent with shock s_2.

(ii) This is implied by $dL/dc_j(s) = 0, \forall j$.

(iii) $dL/dl_j(s) = 0$ and $dL/dc_j(s) = 0$ imply the equation below. The result then follows from the fact that $\gamma(s) > 0$ (Theorem A1(i)) and from the concavity of v. The result for the case $s = s_N$ is obvious.

$$-\frac{v'(l_j(s))}{u'(c_j(s))} = \begin{cases} ws\dfrac{[P(s) + \gamma(s) - \gamma(s^+)]}{\left[P(s) + \gamma(s) - \gamma(s^+)\dfrac{v'\left(l_j(s)\frac{s}{s^+}\right)\frac{s}{s^+}}{v'(l_j(s))}\right]} & : \quad s_1 < s < s_N \\ ws\dfrac{[P(s) - \gamma(s^+)]}{\left[P(s) - \gamma(s^+)\dfrac{v'\left(l_j(s)\frac{s}{s^+}\right)\frac{s}{s^+}}{v'(l_j(s))}\right]} & : \quad s = s_1 \end{cases}$$

||

References

Browning, M., L. Hansen, and J. Heckman. 1999. "Micro Data and General Equilibrium Models." In *Handbook of Macroeconomics,* edited by J. Taylor and M. Woodford. Amsterdam: Elsevier Science B.V.

Conesa, J.C., S. Kitao, and D. Krueger. 2009. "Taxing Capital? Not a Bad Idea After All!" *American Economic Review* 99: 25-48.

Congressional Budget Office. 2004. *"Effective Federal Tax Rates: 1979-2001."* Available at http://www.cbo.gov.

Economic Report of the President. 2004. Washington, DC: United States Government Printing Office.

Feldstein, M., and A. Samwick. 1992. "Social Security Tax Rules and Marginal Tax Rates." *National Tax Journal* 45: 1-22.

Fernandes, A., and C. Phelan. 2000. "A Recursive Formulation for Repeated Agency with History Dependence." *Journal of Economic Theory* 91: 223-247.

Fullerton, D., and D. Rogers. 1993. *Who Bears the Lifetime Tax Burden?* Washigton, DC: The Brookings Institution.

Golosov, M., N. Kocherlakota, and A. Tsyvinski. 2003. "Optimal Indirect and Capital Taxation." *Review of Economic Studies* 70: 569-87.

Golosov, M., and A. Tsyvinski. 2006. "Designing Optimal Disability Insurance: A Case for Asset Testing." *Journal of Political Economy* 114: 257-269.

Golosov, M., A. Tsyvinski, and I. Werning. 2006. "New Dynamic Public Finance: A User's Guide." In *NBER Macroeconomic Annual 2006.* MIT Press.

Gouveia, M., and R. Strauss. 1994. "Effective Federal Income Tax Functions: An Exploratory Analysis." *National Tax Journal* 47: 317-39.

Heathcote, J., K. Storesletten, and G. Violante. 2005. "Two Views of Inequality over the Life-Cycle." *Journal of the European Economic Association Papers and Proceedings* 3: 765-775.

Heathcote, J., K. Storesletten, and G. Violante. 2008. "The Macroeconomic Implications of Rising Wage Inequality in the United States." Manuscript.

Hopenhayn, H., and J. Nicolini. 1997. "Efficient Unemployment Insurance." *Journal of Political Economy* 105: 412-38.

Huggett, M. 1996. "Wealth Distribution in Life-Cycle Economies." *Journal of Monetary Economics* 38: 469-94.

Huggett, M., and G. Ventura. 1999. "On the Distributional Effects of Social Security Reform." *Review of Economic Dynamics* 2: 498-531.

Huggett, M., G. Ventura, and A. Yaron. 2007. "Sources of Lifetime Inequality." NBER Working Paper 13224.

Imrohoroglu, A., S. Imrohoroglu, and D. Joines. 1995. "A Life Cycle Analysis of Social Security." *Economic Theory* 6: 83-114.

Kaplan, G. 2007. "Inequality and the Lifecycle." Manuscript, New York University.

Kocherlakota, N. 2005. "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation." *Econometrica* 73: 1587-1621.

Mas-Colell, A., M. Whinston, and J. Green. 1995. *Microeconomic Theory.* Oxford: Oxford University Press.

Mirrlees, J. 1971. "An Exploration into the Theory of Optimum Income Taxation." *Review of Economic Studies* 38: 175-208.

Mirrlees, J. 1995. "Private Risk and Public Action: The Economics of the Welfare State." *European Economic Review* 39: 383-97.

Nishiyama, S., and K. Smetters. 2007. "Does Social Security Privatization Produce Efficiency Gains?" *Quarterly Journal of Economics* 122: 1677-1719.

Press, W., S. Teukolsky, W. Vetterling, and B. Flannery. 1994. *Numerical Recipes in Fortran: The Art of Scientific Computing,* 2nd ed. Cambridge: Cambridge University Press.

Rogerson, W. 1985. "Repeated Moral Hazard." *Econometrica* 53: 69-76.

Rosen, H. 2002. *Public Finance*, 6th ed. New York: McGraw-Hill.

Saez, E. 2001. "Using Elasticities to Derive Optimal Tax Rates." *Review of Economic Studies* 68: 205-29.

Siegel, J. 2002. *Stocks for the Long Run*, 3rd ed. New York: McGraw-Hill.

Social Security Handbook. 2003. Available at http://www.socialsecurity.gov/.

Storesletten, K., C. Telmer, and A. Yaron. 1999. "The Risk-Sharing Implications of Alternative Social Security Arrangements." *Carnegie-Rochester Conference Series on Public Policy* 50: 213-59.

Tauchen, G. 1986. "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions." *Economics Letters* 20: 177-81.

Tuomala, M. 1990. *Optimal Income Tax and Redistribution.* Oxford: Clarendon Press.

Vickrey, W. 1939. "Averaging of Income for Income-Tax Purposes." *Journal of Political Economy* 47: 379-97.

Wang, C., and S. Williamson. 2002. "Moral Hazard, Optimal Unemployment Insurance, and Experience Rating." *Journal of Monetary Economics* 49: 1337- 71.

Werning, I. 2007. "Optimal Fiscal Policy with Redistribution." *Quarterly Journal of Economics* 122: 925-967.

Table 1: Parameter Values

Definition	Symbol	Value	Source
Model Periods	J	$J = 56$	Age 25 - 80
Retirement Period	R	$R = 41$	Kaplan (2007)
Labor Productivity	$\omega(s_j, j)$	$\omega(s_j, j) = \mu_j \; exp(s_j^1 + s_j^2 + s_j^3)$ $s_j^1 \sim N(-\sigma_1^2/2, \sigma_1^2)$ $s_j^2 = \rho s_{j-1}^2 + \eta_j, \eta_j \sim N(0, \sigma_2^2)$ $s_j^3 \sim N(-\sigma_3^2/2, \sigma_3^2)$ Permanent-Shock Model $(\sigma_1^2, \sigma_2^2, \sigma_3^2, \rho) = (.056, 0, 0, 0)$ Full Model $(\sigma_1^2, \sigma_2^2, \sigma_3^2, \rho) = (.056, .019, .072, .946)$	Kaplan (2007)
Mean Productivity	μ_j	Figure 3	Kaplan (2007)
Preferences	$u(c, l)$	$u(c, l) = \frac{c^{(1-\nu)}}{(1-\nu)} + \phi \frac{(1-l)}{(1-\gamma)}^{(1-\gamma)}$ $(\nu, \gamma, \phi) = (1.66, 5.55, 0.13)$	Kaplan (2007)
Social Security Tax	τ	$\tau = .106$	OASI tax rate
Benefit Function	$b(x)$	Figure 1	SS Handbook (2003)
Income Tax	T^{inc}	Figure 2	CBO (2004)
Interest Rate	r	$r = 0.042$	Siegel (2002)
Discount Factor	β	$\beta = .98803$	See Text

Table 2: Welfare Gains to Optimal Parametric Reforms

Type of Reform	Permanent-Shock Model	Full Model
Reform 1: Change the Benefit Function	0.18	1.15
Reform 2: Tax the Present Value of Earnings	3.95	-0.07
Reform 3: Eliminate Capital Income Taxation	0.22	-0.22
Maximum Possible Gain	4.09	unknown

NOTE: The benefit function is $b(x;\alpha) = \sum_{i=1}^{3} \alpha_i x^{i-1}$, where x is average lifetime earnings. The present-value tax function $T(pv;\alpha)$ is a class of step functions in the permanent-shock model and is a class of piecewise-linear functions in the full model. See the Appendix.

Figure legends

Figure 1: US Social Security Benefit Formula. Source: Social Security Handbook (2003). Average earnings and benefit payments are both expressed as a multiple of average economy wide earnings.

Figure 2: Average Federal Income Tax Rates. Source: Congressional Budget Office (2004).

Figure 3: US Wage Profile. Source: Kaplan (2007).

Figure 4: Inequality Over the Life Cycle - Full Model.

Figure 5a: Maximum Welfare Gains - Private Information. The bold vertical line in Figure 5 highlights the location of the point estimate of the variance described in the text.

Figure 5b: Maximum Welfare Gains - Private vs Public Information. The bold vertical line in Figure 5 highlights the location of the point estimate of the variance described in the text.

Figure 6: Lifetime Taxation.

Figure 7a: Work Hours Profiles - Planning Problem - Permanent Shocks. Labor productivity $w(s, j)$ increases in the shock s. There are five possible shock values $s_1 < s_2 < s_3 < s_4 < s_5$.

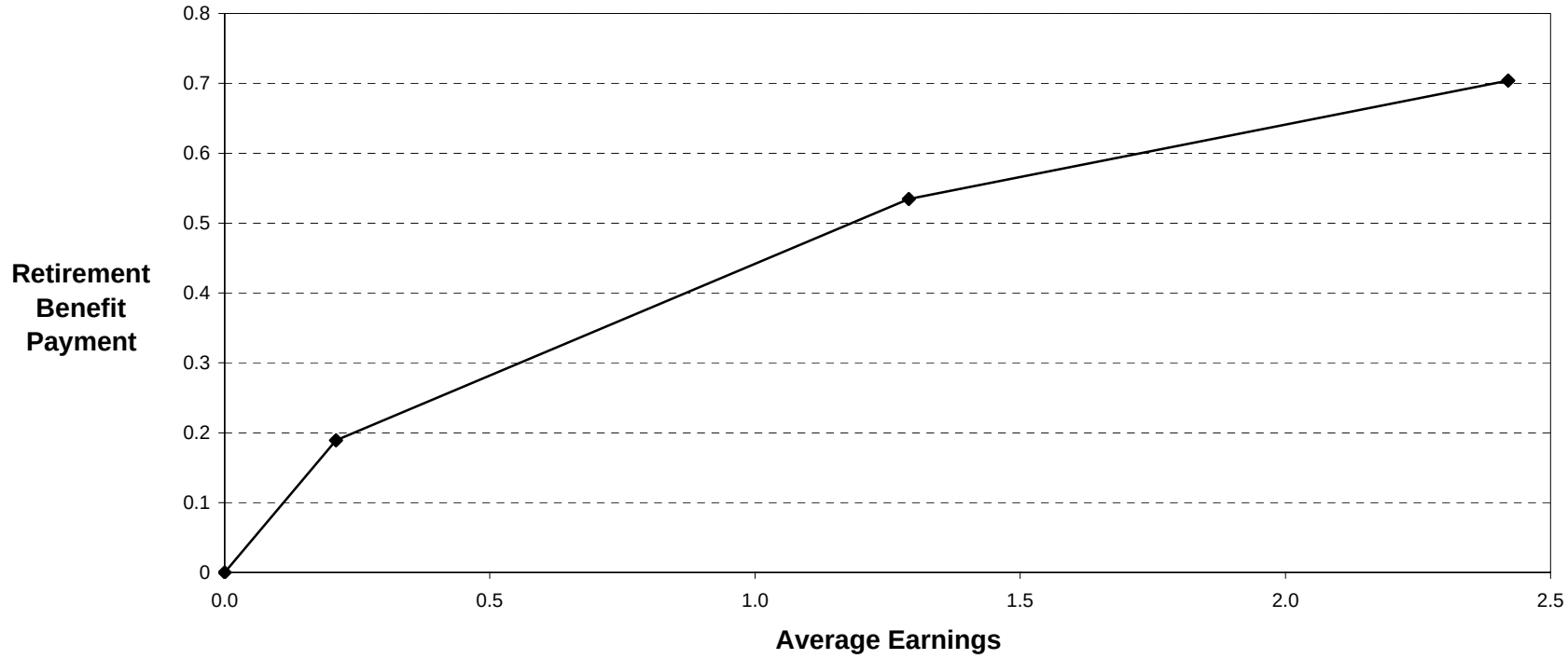
Figure 7b: Work Hours Profiles - Social Security with Income Tax - Permanent Shocks. Labor productivity $w(s, j)$ increases in the shock s. There are five possible shock values $s_1 < s_2 < s_3 < s_4 < s_5$.

Figure 8: Consumption - Labor Wedge Social Insurance. Labor productivity $w(s, j)$ increases in the shock s. There are five possible shock values $s_1 < s_2 < s_3 < s_4 < s_5$.

Figure 9: Consumption - Labor Wedge Planning Problem. Labor productivity $w(s, j)$ increases in the shock s. There are five possible shock values $s_1 < s_2 < s_3 < s_4 < s_5$.

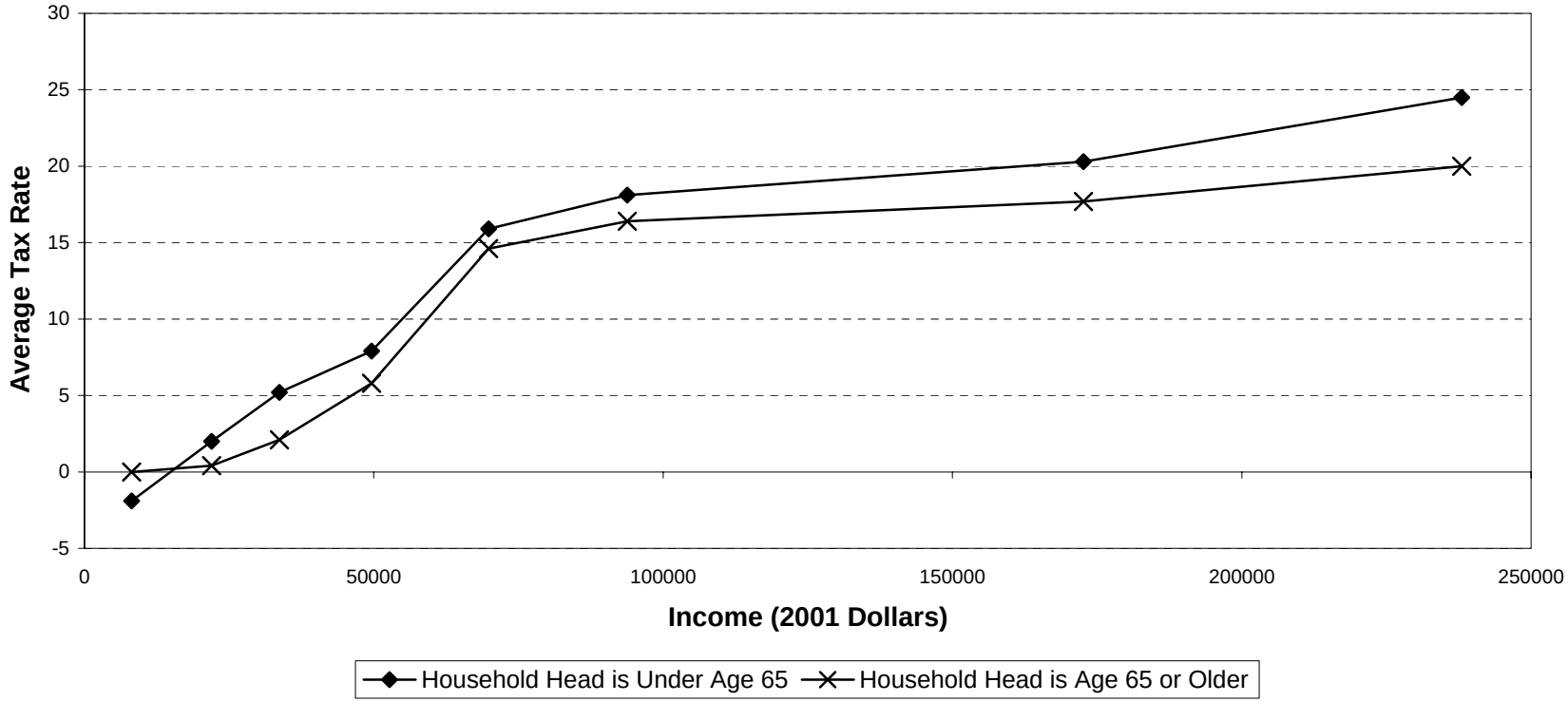
Figure 10: Lifetime Taxation: Full Model. The results for the Benchmark model and Reform 1 are constructed by calculating the average present value of taxes paid for agents whose lifetime earnings fall in different lifetime earnings bins.

Figure 1: US Social Security Benefit Formula



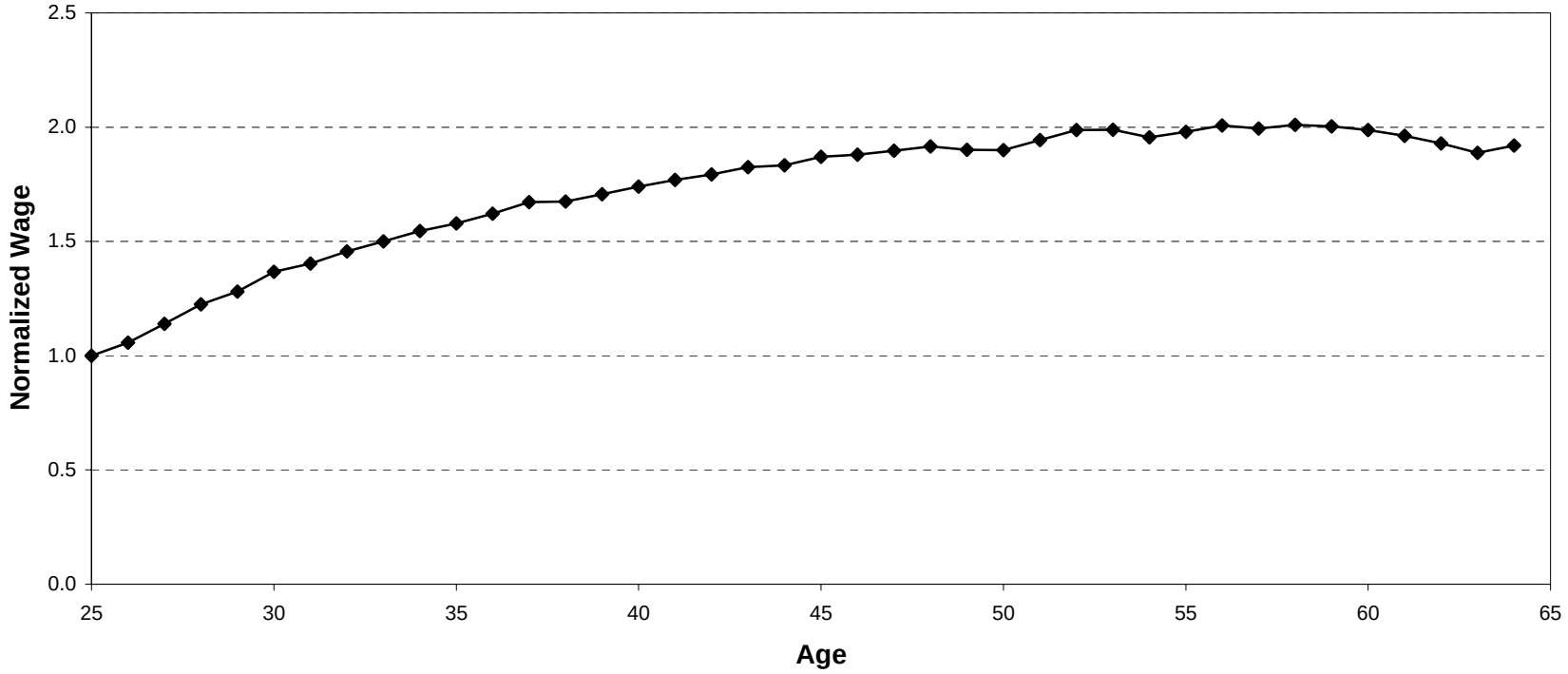
Source: Social Security Handbook (2003).
Average earnings and benefit payments are both expressed as a multiple of average economy wide earnings.

Figure 2: Average Federal Income Tax Rates



Source: Congressional Budget Office (2004)

Figure 3: US Wage Profile



Source: Kaplan (2007).

Figure 4: Inequality Over the Life Cycle - Full Model

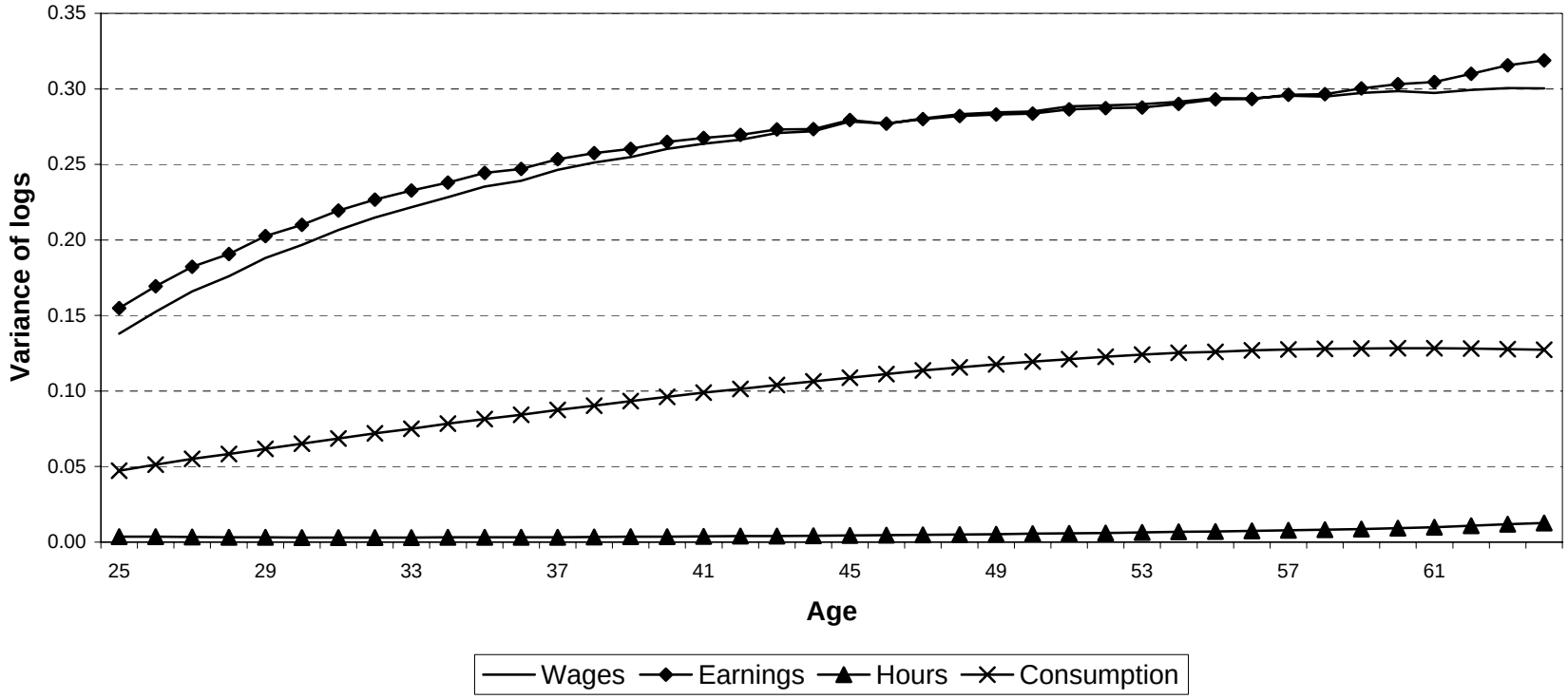
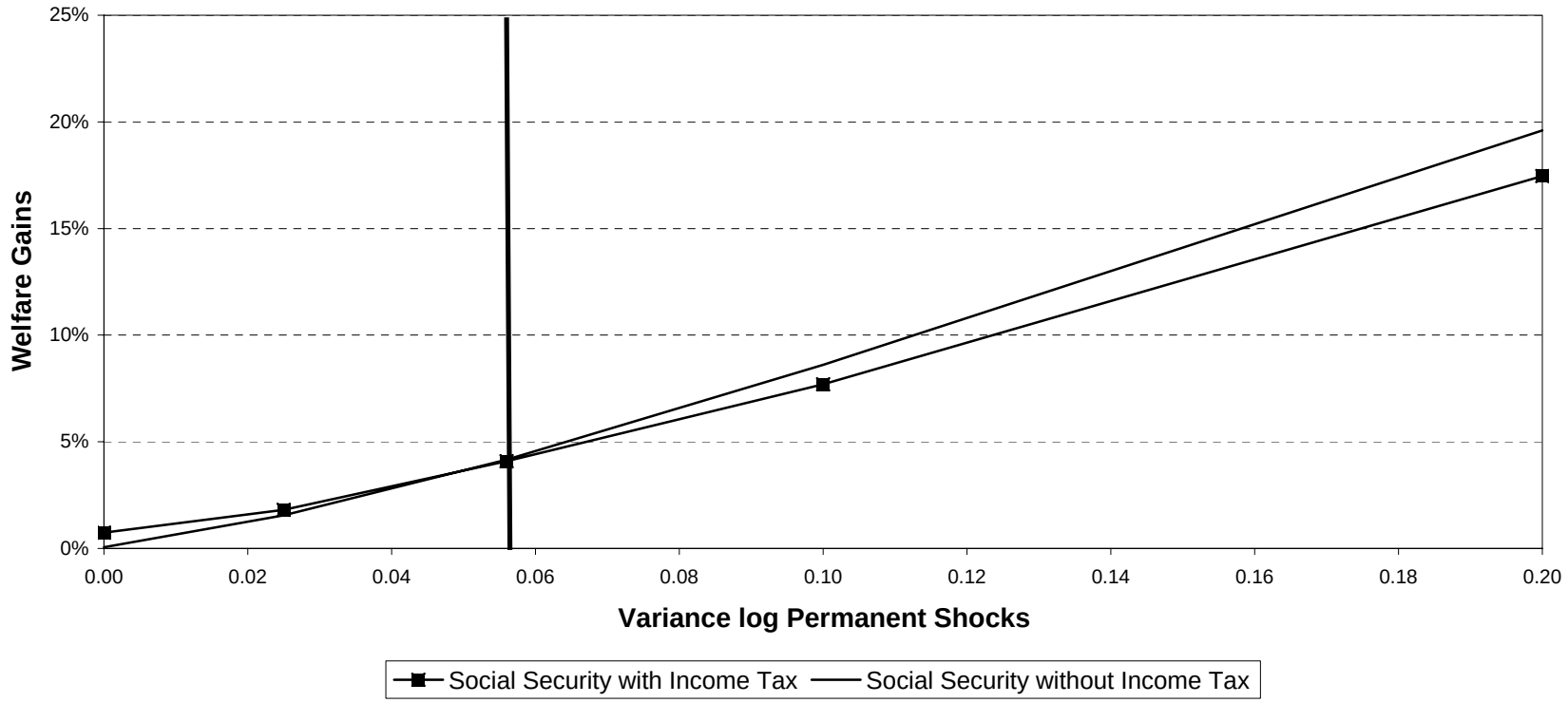
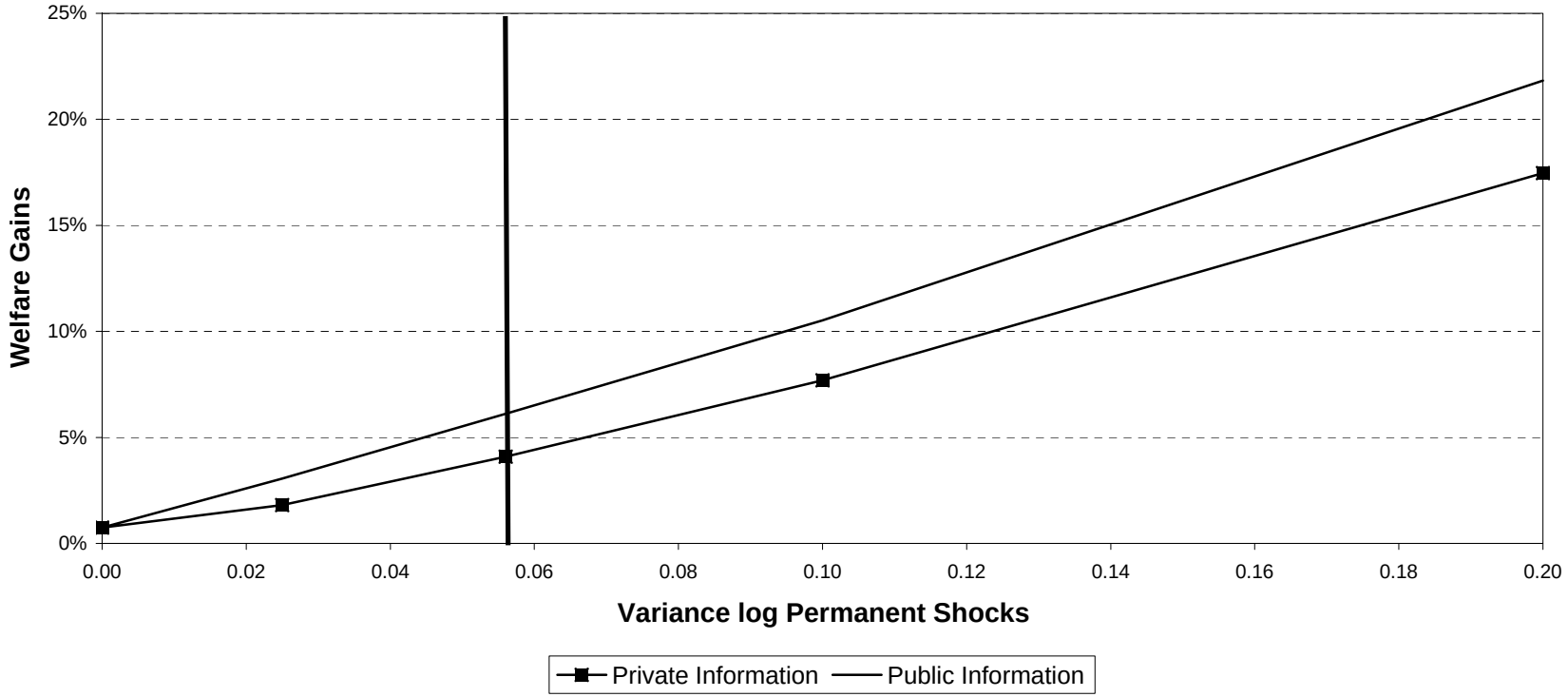


Figure 5a: Maximum Welfare Gains - Private Information



The bold vertical line in Figure 5 highlights the location of the point estimate of the variance described in the text.

Figure 5b: Maximum Welfare Gains - Private vs Public Information



The bold vertical line in Figure 5 highlights the location of the point estimate of the variance described in the text.

Figure 6: Lifetime Taxation

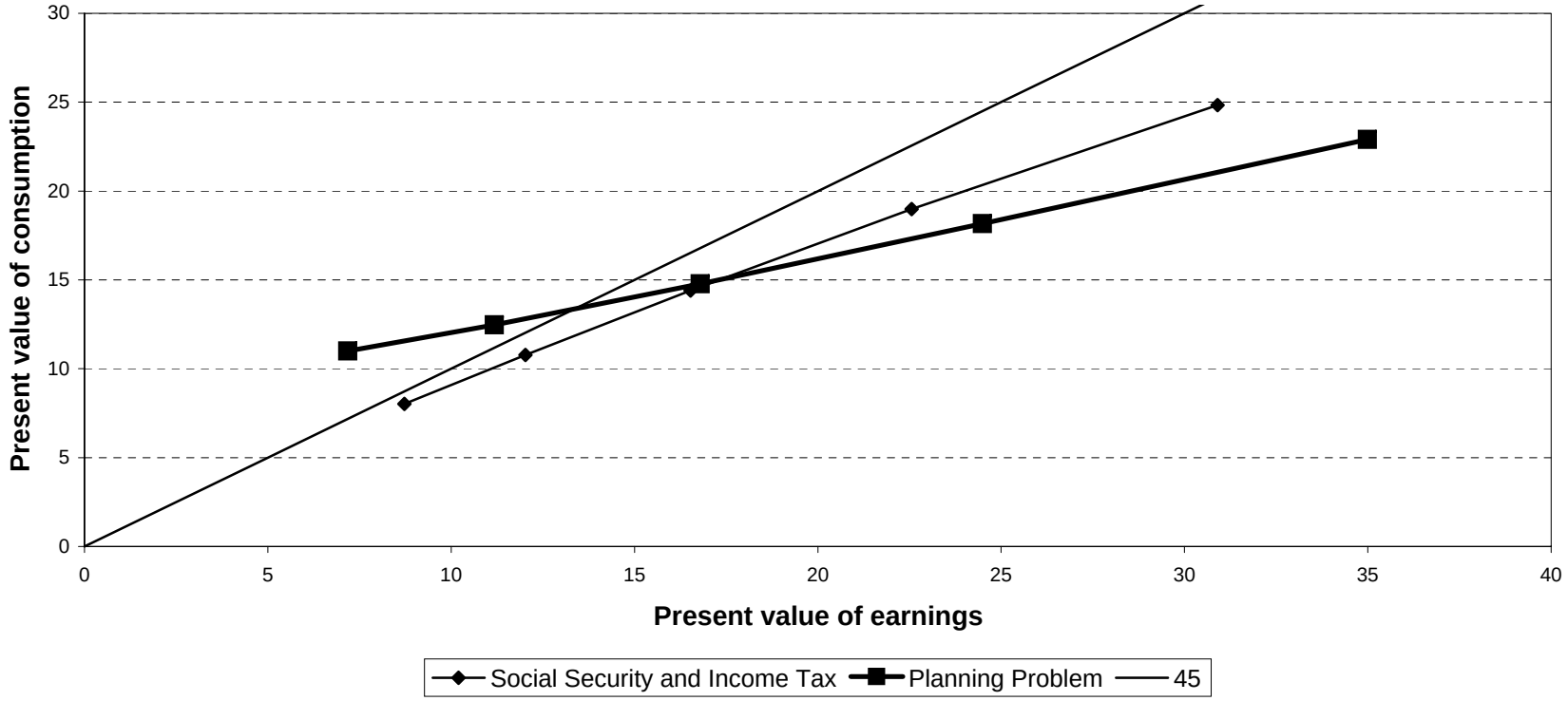
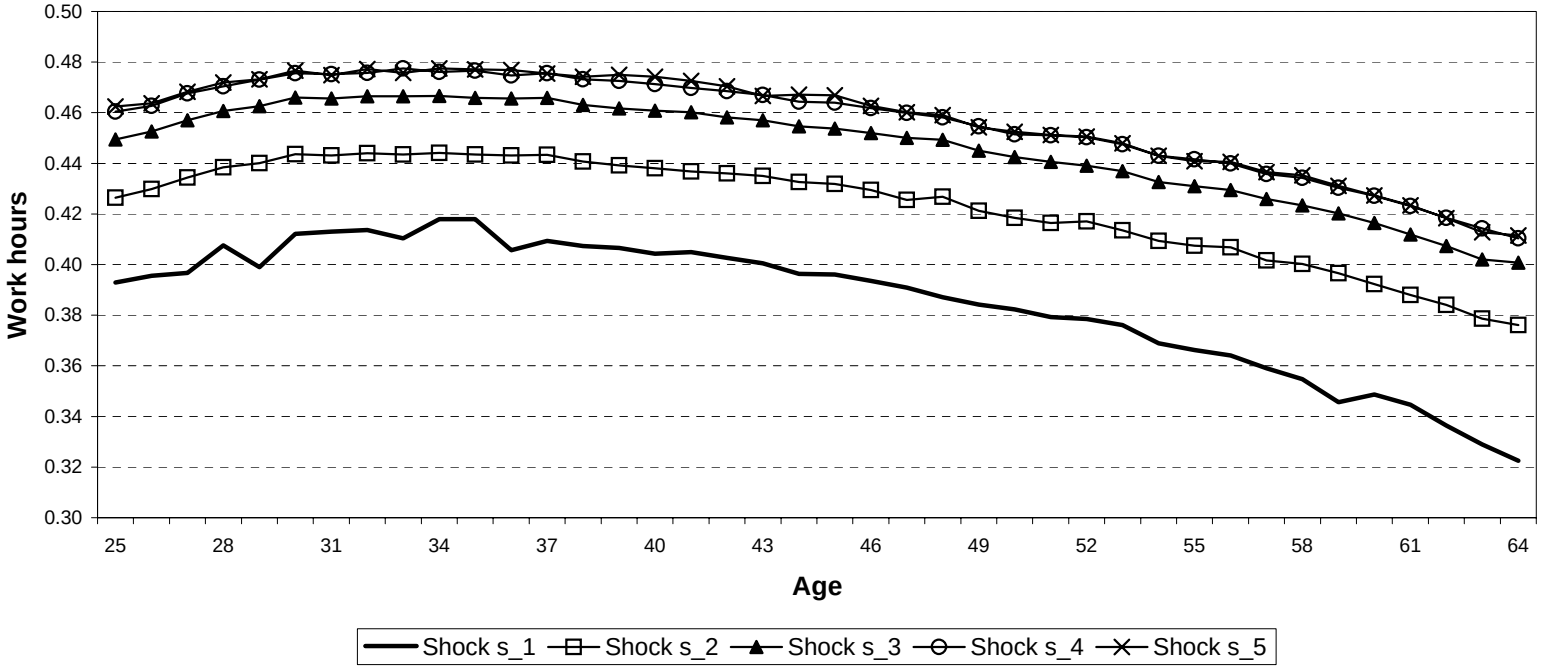
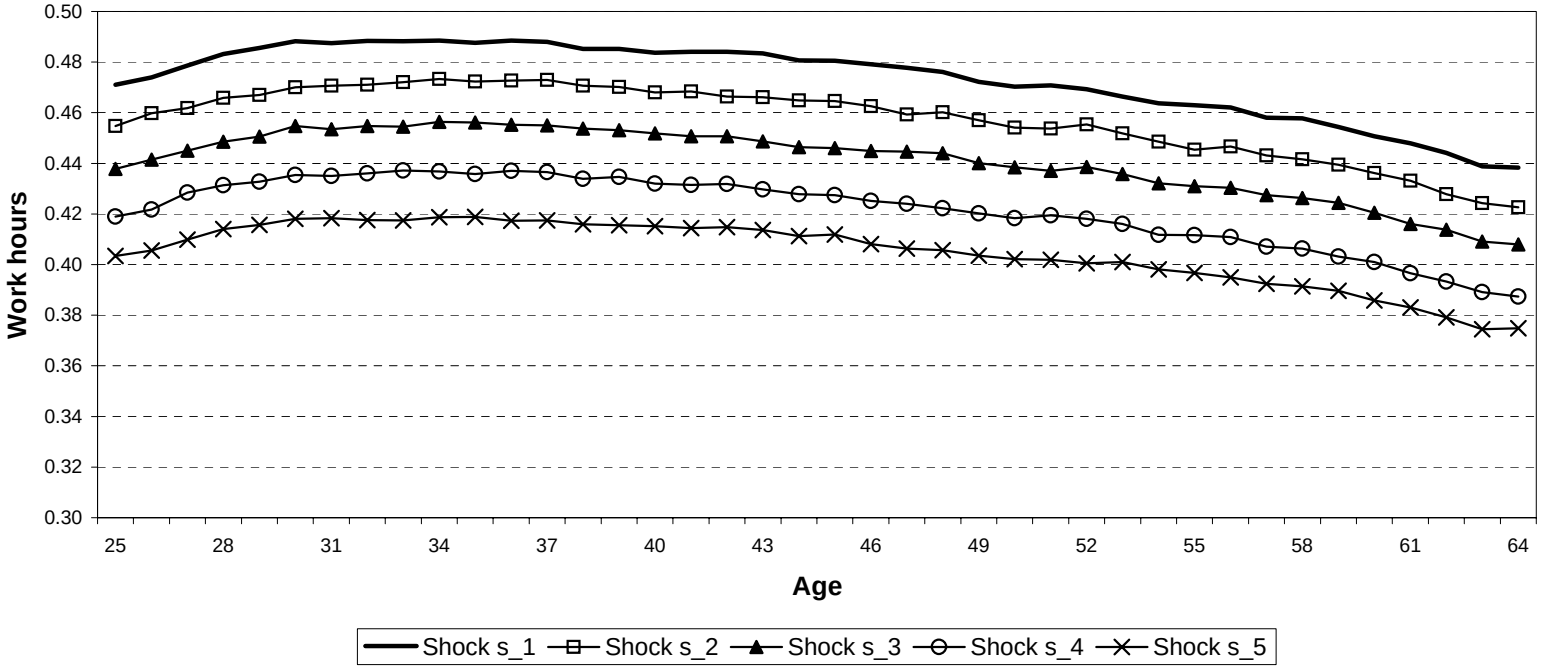


Figure 7a: Work Hours Profiles - Planning Problem - Permanent Shocks



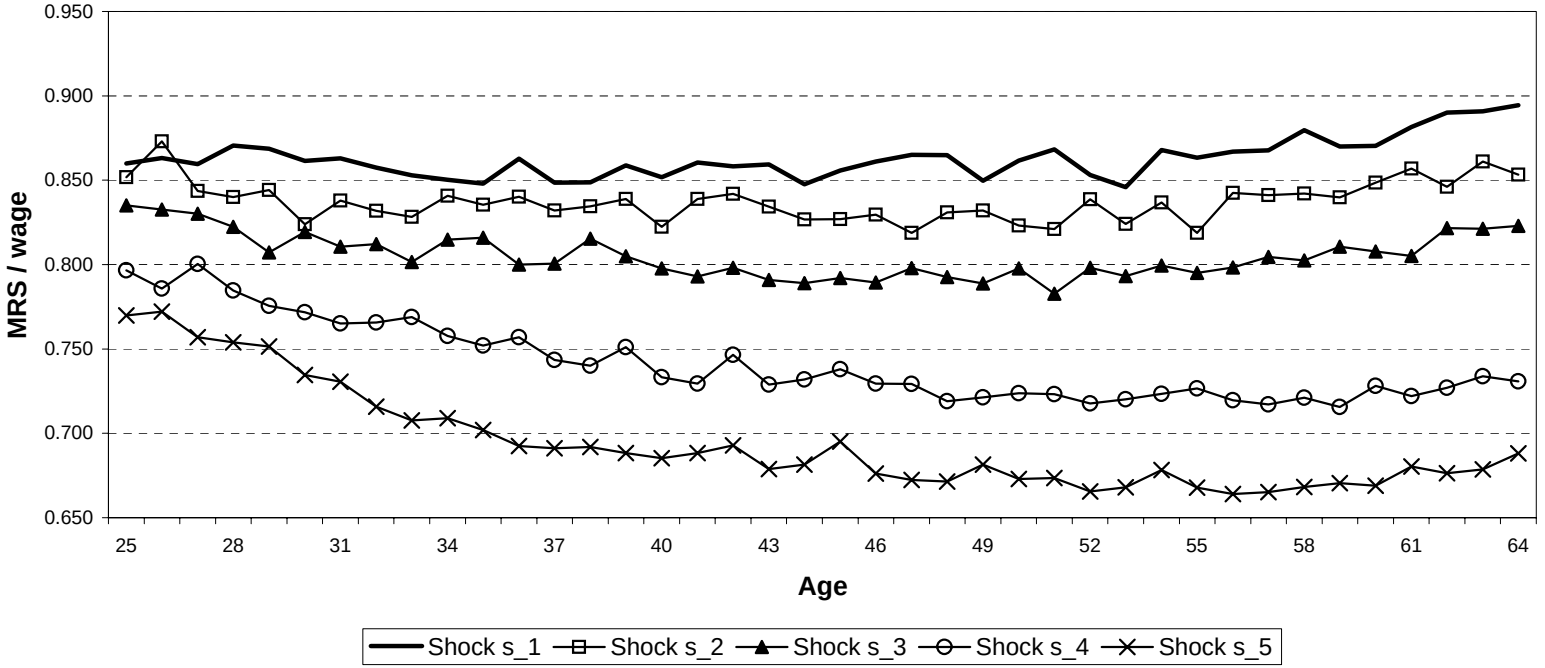
Labor productivity w(s, j) increases in the shock s. There are five possible shock values s_1 < s_2 < s_3 < s_4 < s_5.

Figure 7b: Work Hours Profiles - Social Security with Income Tax - Permanent Shocks



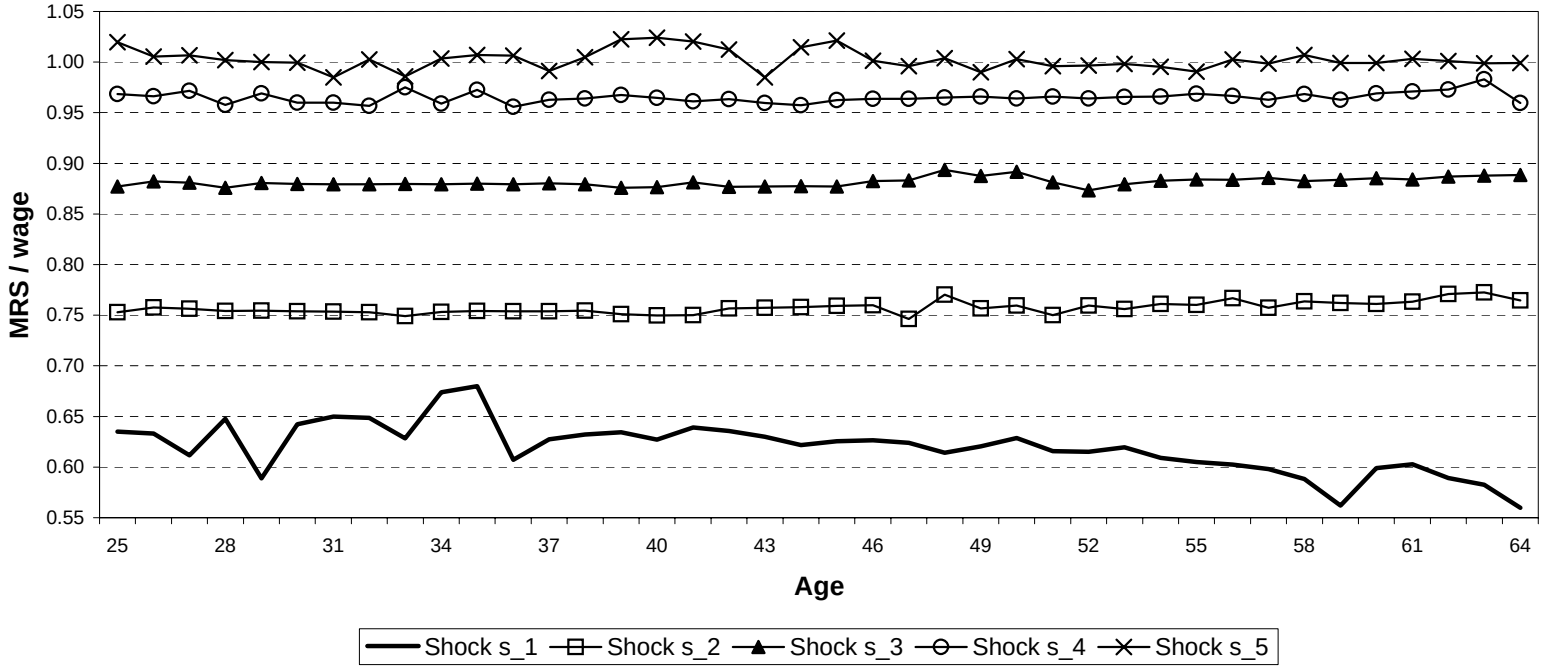
Labor productivity w(s, j) increases in the shock s. There are five possible shock values s_1 < s_2 < s_3 < s_4 < s_5.

Figure 8: Consumption - Labor Wedge Social Insurance



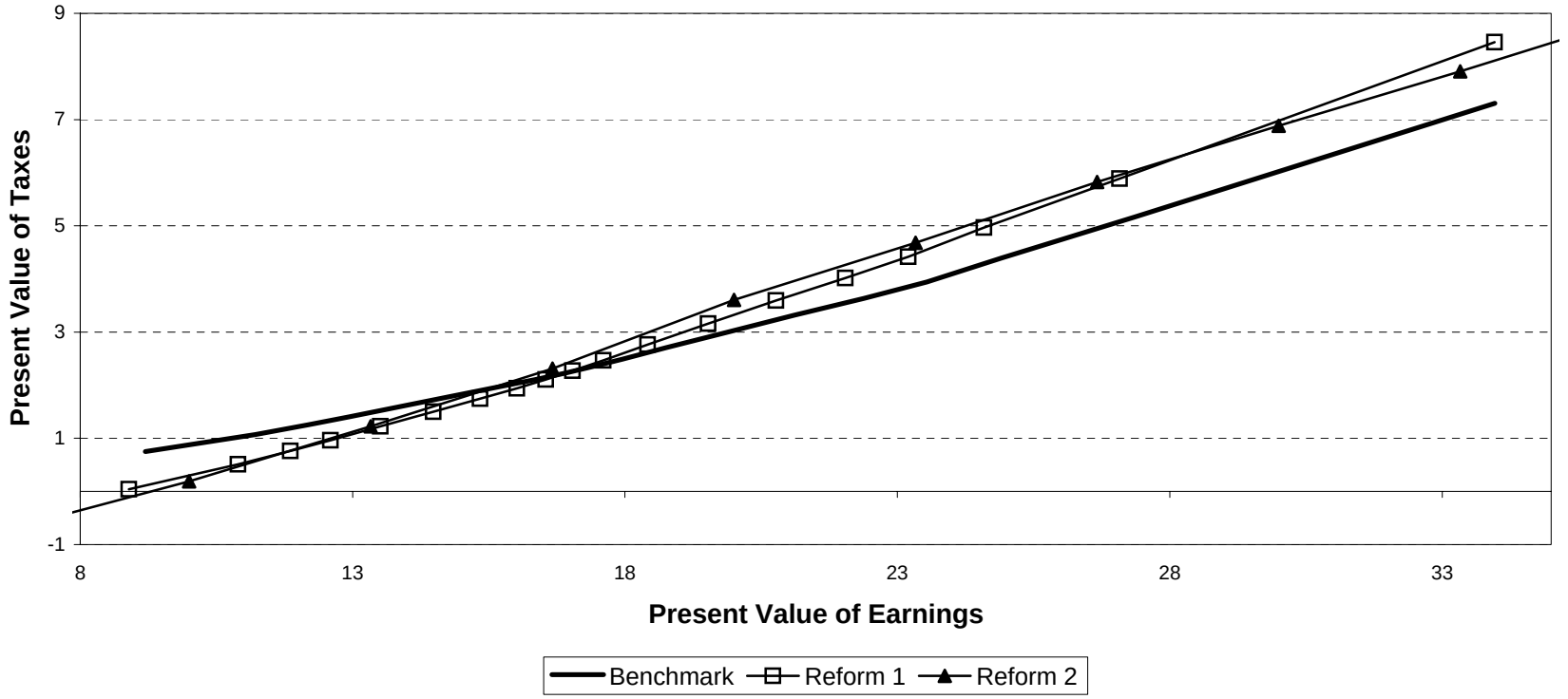
Labor productivity w(s, j) increases in the shock s. There are five possible shock values s_1 < s_2 < s_3 < s_4 < s_5.

Figure 9: Consumption - Labor Wedge Planning Problem



Labor productivity w(s, j) increases in the shock s. There are five possible shock values s_1 < s_2 < s_3 < s_4 < s_5.

Figure 10: Lifetime Taxation: Full Model



The results for the Benchmark model and Reform 1 are constructed by calculating the average present value of taxes paid for agents whose lifetime earnings fall in different lifetime earnings bins.