Optimal linear and two-bracket income taxes with idiosyncratic earnings risk[☆]

Minchung Hsu [a], C.C. Yang [b,c,d,*]

[a] National Graduate Institute For Policy Studies (GRIPS), Tokyo, Japan
[b] Institute of Economics, Academia Sinica, Taiwan
[c] Department of Public Finance, National Chengchi University, Taiwan
[d] Department of Public Finance, Feng Chia University, Taiwan

ABSTRACT

This paper quantitatively characterizes optimal linear and two-bracket income taxes. We consider a dynamic-stochastic-general-equilibrium model in which tax design involves redistributing income for both equity and social insurance. Substantive findings include: (i) a significant fraction of agents supply zero labor or hold zero assets at the optimum; (ii) neglecting tax distortion imposed on either of labor–leisure and consumption–saving decisions will lead to the prescription of tax codes that deviate substantially from the optimum; and (iii) the optimal two-bracket tax schedule will turn from regressivity to progressivity in the marginal tax rate once the volatility of idiosyncratic shocks becomes sufficiently large. The last finding is consistent with the results in Apps et al. (forthcoming), and it also reconciles the contradictory results regarding the optimal two-bracket tax schedule between Slemrod et al. (1994) and Strawczynski (1998).

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Mirrlees (1974) pioneered the study of optimal income taxation in a setting where ex ante identical agents face idiosyncratic shocks to their earnings, but relevant insurance markets are missing. This missing-market setting invites a role for income tax to serve as a partial substitute to absorb income fluctuations and share the idiosyncratic risk across agents. The motive for redistributive taxation here is not for equity per se, but rather for social insurance.

Varian (1980) took up the issue addressed by Mirrlees (1974) with the emphasis that a large portion of income differences between agents is attributable to pure luck rather than innate ability. Unlike Mirrlees's static framework where agents make a choice between labor and leisure, Varian considered a dynamic framework where agents make a choice between current and future consumption.

The Mirrlees–Varian model of optimal income taxation is one of the pioneering works in the moral hazard class of the principal-agent problem, in which the key tradeoff involved is between inducing incentives and providing insurance (Laffont and Martimort, 2002). Subsequent works, including Tuomala (1984), Strawczynski (1998), Low and

Maldoom (2004) and Kanbur et al. (2008), have elaborated on Mirrlees–Varian's original idea in a variety of directions. Our paper contributes to this line of the optimal taxation literature mainly on the front that the tax design problem in our model involves "correcting" income distribution across agents for equity as well as providing social insurance to buffer against agents' idiosyncratic risk.[1]

Under plausible assumptions, Mirrlees (1971) found that the optimal non-linear income tax is approximately linear. In contrast to Mirrlees (1974), this 1971 seminal work belongs to the adverse selection class of the principal-agent problem, in which income differences between agents are attributed to innate ability (type) rather than pure luck. The government's tax design in the Mirrlees (1971) framework is to trade off "correcting" income distribution for equity against dulling incentives to work (Laffont and Martimort, 2002).

Subsequent studies following Mirrlees (1971) have further explored the tax schedules of optimal income taxation.[2] Many of them are based on the mechanism design approach, which gives rise to highly nonlinear tax schedules. However, in the real world, virtually all income tax systems are piecewise linear. In this paper we focus on piecewise linear income tax and, in particular, the linear and the two-bracket income tax.

---

* Corresponding author at: Institute of Economics, Academia Sinica, Nankang, Taipei 115, Taiwan.
E-mail addresses: minchunghsu@grips.ac.jp (M. Hsu), ccyang@econ.sinica.edu.tw (C.C. Yang).

[1] Strawczynski (1998) mainly considered the Varian problem in a representative-agent framework; however, he also analyzed a four-agent economy in which two levels of skill apply respectively to two realizations of shock. See also Eaton and Rosen (1980) and Diamond et al. (1980).

[2] Tuomala (2010) provides a recent study on the issue; see the references therein for other studies.

Stern (1976) is perhaps the most celebrated work that quantitatively characterizes the optimal linear income tax. His model is static and deterministic in the framework of Mirrlees (1971), while our model is dynamic and stochastic. In the context of the two-bracket income tax, Slemrod et al. (1994) found quantitatively that the second marginal tax rate is lower than the first at the optimum, whereas Strawczynski (1998) derived the opposite result.[3] It should be noted that the former is framed in a static, deterministic, and ability-driven environment à la Mirrlees (1971), whereas the latter is framed in a dynamic, stochastic, and luck-driven environment à la Varian (1980). The results of our model are driven by agent heterogeneity in both ability and luck. We investigate why the optimal two-bracket income tax may be progressive or regressive in the marginal tax rate.

The work of Slemrod et al. (1994) basically follows that of Stern (1976), but extends the numerical analysis to the two-bracket case. Specifically, it assumes a lognormal wage rate (ability) distribution, the parameters of which are taken from Stern (1976). In a recent paper, Apps et al. (forthcoming) revisited the problem of the optimal two-bracket income tax and presented simulation results based on Pareto wage rate distributions. They numerically discerned the circumstances under which marginal tax rate progressivity or regressivity will arise. Apps et al. (forthcoming) considered their problem in the framework of Mirrlees (1971) and so naturally they did not address the issue of conflicting findings between Slemrod et al. (1994) and Strawczynski (1998). We complement the Apps et al. (forthcoming) results by tackling the left-out issue.

In addition to the studies mentioned above, our paper is closely related to Conesa and Krueger (2006) and Conesa et al. (2009), both of which address optimal income taxation in a dynamic-stochastic-general-equilibrium setting.[4] Besides modeling details and derived results, there are at least three major differences between our paper and theirs. First, we consider the piecewise linear income tax, whereas they considered a three-parameter family of nonlinear income tax schedules. As such, we are able to relate our findings directly to the previous literature on optimal linear and two-bracket income taxation, while they cannot. Second, they addressed optimal taxation in a life-cycle model, while we are in an infinite-horizon framework. Abstracting from life-cycle complications enables us to focus on non-life-cycle elements that are responsible for the design of income tax. Barro (1991) argued that the infinite horizon applies naturally if agents care about their children, who in turn care about their children, and so on. Third, while tax revenues collected are used solely to finance government consumption in Conesa and Krueger (2006) and Conesa et al. (2009), they are used to finance transfer payments as well as government consumption in our paper. In line with the tradition of optimal income taxation à la Mirrlees (1971, 1974), the so-called "tax" schedule in our model actually represents a "tax and transfer" schedule. As Brewer et al. (2010, p. 94) remarked: "Despite its name, optimal tax theory concerns itself just as much with the design of benefits as it does the setting of income tax rates …"

A recent paper by Boadway and Sato (2011) has analytically provided a fairly general treatment of optimal income taxation when differences in individual income are attributed to both ability and luck. For simplicity, they assumed that preferences are quasi-linear in labor so as to eliminate income effects in the demand for consumption. Even in this simplified setting, the derived analytical results seem rather complicated and may fail to prescribe concrete tax structures; see their Proposition 2. As noted by Boadway and Sato (2011), there has been relatively little attention devoted to studying optimal income taxation in the presence of heterogeneity in both ability and luck. Following Stern (1976), Slemrod et al. (1994), Strawczynski (1998), and Apps et al. (forthcoming), we quantitatively characterize optimal linear and two-bracket income taxes but synthesize these previous studies in a framework where both ability and luck matter for the determination of individual income.

The rest of the paper is organized as follows. Section 2 introduces our model. Section 3 calibrates the parameter values of the model. Section 4 considers welfare criteria for optimal taxation. Sections 5–8 report our results and Section 9 concludes.

2. Economic environment

In an important benchmark of the incomplete markets model,[5] Aiyagari (1994) considered a dynamic-stochastic-general-equilibrium (DSGE) setting in which agents face idiosyncratic earnings risk that cannot be insured. Our model follows his model closely. In the Aiyagari economy, labor hours are exogenously given and income is not subject to taxation. We allow for the choice of labor hours and the imposition of income taxes. The Aiyagari model is interesting from the viewpoint of taxation, in that it generates an endogenous cross-sectional distribution of income and of wealth, which is conditional upon tax parameters.

2.1. Setting

Time is discrete and runs from $t = 0, 1,..., \infty$. The economy is populated by a continuum of infinitely-lived agents (households) of unit mass. Each agent is atomistic and so a price taker. Agents are heterogeneous in that they face different histories of realizations of idiosyncratic shocks to their labor productivity. This is the only source of heterogeneity across agents in the model.

There are three sectors in the economy: households, firms, and the government. There are three goods: the service of labor, the service of capital, and a final good that can be used for either consumption or investment. We let the final good be the numeraire.

2.2. Labor productivity shocks

There is no aggregate risk in the economy. All agents are subject to idiosyncratic labor productivity shocks, which are realized at the beginning of each period $t > 0$ (each agent starts identically at time 0 with some initial productivity shock). There are no viable insurance markets or state-contingent securities available for agents to insure against the risk of the shocks. The realized shocks take a finite number of possible values, which are observed by agents before making their labor–leisure and consumption–saving decisions in each period. The stochastic process of the shocks is identical and independent across agents, and follows a Markov chain with stationary transitions over time. The Markov chain is parameterized by appealing to econometric studies based on micro-level data. The details of this process will be deferred to the next section when we calibrate the model.

We let z denote the generic realization of the labor productivity shock, and normalize the mean of z to be unity. The effective labor supply for an agent equals zn, where n is her labor hours chosen.

2.3. Asset market

There are no state-contingent assets but a single risk-free, one-period asset. Agents have no asset at time zero; however, they can accumulate their asset holdings by saving. Saving will be channeled to become capital, which is used by firms in production. Since there is only one asset held by agents, the distribution of this asset represents the distribution of wealth in the economy.

---

[3] See also Sheshinski (1989), who presented a proof in the framework of Mirrlees (1971) that a regressive two-bracket tax code can never be optimal. However, Slemrod et al. (1994) showed that Sheshinski's proof is flawed since it ignores a possible discontinuity in the tax revenue function.

[4] There is also a literature called "new dynamic public finance," in which the emphasis is on the implications of information frictions for optimal taxes in dynamic settings; see Golosov et al. (2006) and Kocherlakota (2010) for reviews.

[5] For introductions to the incomplete market model, see Heathcote et al. (2009) and Guvenen (2011).

Agents may trade in the asset to absorb idiosyncratic risk. However, agents face a borrowing constraint so that the amount of their next period's asset cannot fall below a lower bound b.[6] This borrowing constraint not only rules out Ponzi schemes, but also serves as an effective way of limiting the ability of agents to fully smooth consumption over time. As a result of the constraint, agents may save in a precautionary manner to self-insure against future income drops. Agents may also supply labor in a precautionary manner as emphasized in the static model of Low and Maldoom (2004). It should be noted that while self-insurance behavior arises in static models as in Low and Maldoom (2004) if the third derivative of the period utility function is positive, it arises in dynamic models as long as agents are risk-averse and face a borrowing constraint that can bind due to adverse shocks.[7]

At any given time each agent is characterized by a state $s = (z, a)$, where z is her realization of idiosyncratic labor productivity shocks and a is her asset holdings. Different states represent different types of agents. We let $F_t(s)$ denote the type distribution of agents at time t.

2.4. Production technology

The aggregate output Y of the economy at time t is given by a Cobb–Douglas production function

$$Y_t = K_t^\theta L_t^{1-\theta},$$

where K and L are the aggregate capital and effective labor, and θ is the capital share. Capital depreciates at an exogenous rate of $\delta \epsilon (0,1)$ in each and every period. All markets behave competitively. With a constant returns to scale technology and perfect competition, we assume without loss of generality the existence of a representative firm operating this technology.

Let w denote the real wage rate per unit of effective labor, and r the net-of-depreciation real rental rate per unit of capital. Given w and r, the representative firm's maximizing profit yields the following condition:

$$w_t = (1-\theta)(K_t/L_t)^\theta;$$
$$r_t = \theta(K_t/L_t)^{\theta-1} - \delta.$$

2.5. Government and income tax schedules

The government via taxes and transfers engages in redistributing income across agents and, at the same time, it must collect enough tax revenue to finance a sequence of exogenously given government consumption $\{G_t\}_{t=0}^\infty$.

Let y denote agent pre-tax income, i.e., $y = wzn + ra$, which equals labor income wzn plus capital income ra. Let T be the tax-transfer schedule imposed on y by the government. Following the convention, we shall refer to $T(y)$ simply as a tax schedule. The government is committed to applying the same tax schedule in each period, and the tax schedule does not discriminate between labor and capital income. Both features approximately hold in the real world.[8] We focus on the piecewise linear income tax and, in particular, on its two simplest forms: linear and two-bracket.

A linear income tax is defined as

$$T(y) = \tau y - g,$$

where τ is the marginal tax rate applied to pre-tax income, and g is the uniform per capita grant. A two-bracket income tax is defined as

$$T(y) = \begin{cases} \tau_1 y - g & \text{if } y \le y_0 \\ \tau_1 y_0 + \tau_2(y-y_0) - g & \text{if } y > y_0 \end{cases},$$

where y_0 is the cutoff point, τ_1 is the marginal tax rate applied to the first income bracket ($y \le y_0$), and τ_2 is the marginal tax rate applied to the second income bracket ($y > y_0$). Note that $T(y = 0) = -g$, which is the transfer role of the tax schedule.

Following Conesa and Krueger (2006) and Conesa et al. (2009), the government is required to balance its budget in each period and so we have

$$G_t = \int T[y_t(s)]dF_t(s).$$

Given G, the parameters of the tax schedule to be determined at the optimum are τ in the case of the linear tax, and τ_1, τ_2, and y_0 in the case of the two-bracket tax. The amount of transfer g is determined residually from the government's balanced budget.

2.6. Households

2.6.1. Preferences

In each period each agent (household) is endowed with one unit of time, which is divided between labor, n, and leisure, $1-n$. Agent preferences over consumption and leisure in each period are represented by a period utility function of the form:

$$u(c, 1-n) = \frac{\left[c^\phi(1-n)^{1-\phi}\right]^{1-\mu}}{1-\mu},$$

where ϕ is a parameter denoting the relative importance of consumption versus leisure, and μ is a parameter related to the risk aversion of the agent. Given the intertemporal budget constraint specified later, labor hours chosen can be derived in closed form:

$$n = 1 - \frac{(1-\phi)c}{\phi(1-\tau_y)wz}, \tag{1}$$

where τ_y is the marginal income tax rate that an agent faces. Note that the value of n derived from the above equation may be negative. Whenever this occurs, we set $n = 0$.

The period utility function specified above is widely used in the macro literature. It is consistent with balanced growth, which is a broad fact about the growth of advanced industrial economies; see Heer and Maussner (2009, Chapter 1).[9]

2.6.2. Household problem

Given an income tax schedule $T(y)$ and factor prices w and r, agent i's objective is to maximize her expected discounted lifetime utility

$$E_0\left[\sum_{t=0}^\infty \beta^t u\left(c_{i,t}, 1-n_{i,t}\right)\right],$$

---

[6] For arguments in favor of borrowing constraint settings see, for example, Aiyagari (1994).

[7] Huggett and Ospina (2001) showed that precautionary saving occurs in an infinite horizon setting if and only if the borrowing constraint binds for some agents; it is not necessary to have convexity of marginal utility to generate precautionary behavior.

[8] For example, the current US personal income tax code is applied year after year (unless there is some tax reform), and it does not distinguish sources of income in general when computing tax liabilities. A notable exception with discrimination between sources of income is the so-called "dual income tax" (see Sorensen, 1994, for the detail).

[9] The implied (Frisch) labor supply elasticity is around unity, which may be relatively high compared to those found in microeconometric studies; see Chetty et al. (2011). However, Keane and Rogerson (2011) argued for several possible routes for reconciling micro and macro labor supply elasticities; see also Chetty (2012). In an earlier version of the paper, we also consider an alternative utility function that is widely used in the micro literature, $u(c, 1-n) = \frac{c^{1-\mu_1}}{1-\mu_1} + n\frac{(1-n)^{1-\mu_2}}{1-\mu_2}$ and address the sensitivity of our results by allowing for lower labor supply elasticities. The main results remain qualitatively unchanged.

where β is the time discount factor, and the expectation E_0 is taken with respect to the stochastic process governing labor productivity shocks z as of time 0.

We apply standard dynamic programming techniques to solve the agent's problem. Let $s = (a,z)$ summarize the state of an agent and $V(s)$ denote the agent's value function. The problem facing an agent can be expressed as follows:

$$V(s) = \max_{c,n,a'} u(c, 1-n) + \beta E V\left(s'\right) \qquad (2)$$

subject to

$c + a' = wzn + (1 + r)a - T(y);$
$T(y)$ is either linear or two-bracket;
$y = wzn + ra;$
$a' \geq -b; c \geq 0; 0 \leq n \leq 1, a(t = 0) = 0, z(t = 0) = 1$

where, as is common in dynamic programming, we use a prime to denote all next-period variables. The weak inequality $a' \geq -b$ denotes the borrowing constraint faced by the agent. We allow for the possibility of corner solutions with $a' = -b$ or $n = 0$. The corner solution where $c = 0$ or $n = 1$ will not arise because the marginal period utility will go to infinity (the Inada condition) under our parameterization.

2.6.2.1. Linear income tax. There are two possibilities to consider in the case of the linear income tax, depending on whether the borrowing constraint is binding or not.

(i) Not binding ($a' > -b$): Using the budget constraint $c = wzn + (1 + r)a - T(y) - a'$ gives the Euler equation:

$$u_c(c, 1-n) = \beta E V_{a'}\left(s'\right).$$

Calculate u_c and substitute in the closed-form solution of n as expressed by (1) if $n > 0$. We can then write the above Euler equation as

$$\phi\left[\frac{(1-\phi)}{\phi(1-\tau)wz}\right]^{(1-\phi)(1-\mu)} c^{-\mu} = \beta E V_{a'}\left(s'\right);$$

$$\Rightarrow c = \left[\frac{\beta E V_{a'}\left(s'\right)}{\phi\left[\frac{(1-\phi)}{\phi(1-\tau)wz}\right]^{(1-\phi)(1-\mu)}}\right]^{\frac{-1}{\mu}}.$$

Applying the envelope theorem yields:

$$E V_{a'}\left(s'\right) = E\left[\left(1 + r'(1-\tau)\right) u_{c'}\left(c', 1-n'\right)\right].$$

Thus we have

$$c = \left\{\frac{\beta E\left[\left(1 + r'(1-\tau)\right) u_{c'}\left(c', 1-n'\right)\right]}{\phi\left[\frac{(1-\phi)}{\phi(1-\tau)wz}\right]^{(1-\phi)(1-\mu)}}\right\}^{\frac{-1}{\mu}}.$$

If $n = 0$, we simply substitute in $n = 0$ instead of the closed-form solution of n.

(ii) Binding ($a' = -b$): We use the budget constraint to solve for c:

$$c = wzn + (1 + r)a - T(y) + b$$

$$\Rightarrow c = (1-\tau)wz\left[1 - \frac{(1-\phi)c}{\phi(1-\tau)wz}\right] + (1+r)a - \tau ra + g + b$$

$$\Rightarrow c = \phi[(1-\tau)wz + (1 + r(1-\tau))a + g + b].$$

If $n = 0$, we simply substitute in $n = 0$ instead of the closed-form solution of n.

2.6.2.2. Two-bracket income tax. There are also two possibilities to consider in the case of the two-bracket income tax, depending on whether the borrowing constraint is binding or not. Similar to the case of the linear income tax, we make use of the Euler equation to solve for c if the borrowing constraint is not binding, and the budget constraint to solve for c if the borrowing constraint is binding. However, each possibility now has three subcases: $\tau_y = \tau_1$ if $y < y_0$, $\tau_y = \tau_2$ if $y > y_0$, and τ_y is undefined if $y = y_0$. When $y = y_0$, the closed-form solution of n fails to apply even if $n > 0$. The choice of labor hours n in this subcase is solved by making use of the equality $wzn + ra = y_0$.

2.7. Stationary equilibrium

Our analysis focuses on the stationary equilibrium, which is the counterpart of the steady state in our stochastic economy. While all variables are constant in the steady state, the prices and the distribution of the state variables are all constant in the stationary equilibrium. In particular, we have $F_t = F_{t+1} = F$, that is, the type distribution of agents is time invariant.[10] The focus on stationary equilibrium is useful if we want to concentrate on the long-run effects of taxes imposed.

Let $G_t = G$ from some period t onward. Conditional on government consumption G and the income tax schedule $T(y)$, a stationary equilibrium for the economy consists of a value function $V(s)$ and decision rules $(c(s), n(s), a'(s))$ for agents, an allocation of factors (K, L) for the representative firm, a time-invariant type distribution $F(s)$, and relative prices of labor and capital (w, r), such that the following conditions are satisfied:

- Given (w, r), the value function $V(s)$ is a solution to the agent's decision problem (2), and $(c(s), n(s), a'(s))$ are the associated optimal decision rules.
- Given (w,r), the firm maximizes profits satisfying:

$w = (1-\theta)(K/L)^{\theta};$
$r = \theta(K/L)^{\theta-1} - \delta.$

- Given (w, r), $F(s)$ is time-invariant or stationary, and consistent with the optimal decision rule $a'(s)$ and the Markov chain for the labor productivity shock z.
- Given (w, r), the government's budget constraint is met, i.e., $G = \int T[y(s)]dF(s)$.
- All markets are clear:

$K = \int a dF(s);$
$L = \int zn(s)dF(s)$
$C + I + G = Y = K^{\theta}L^{1-\theta},$

where $C = \int c(s)dF(s)$ and $I = K' - (1 - \delta)K.$

---

[10] Let $f_t(z, a)$ be the probability density function of $F_t(z, a)$. Note that $f_t(z,a) = Pr(z_t = z, a_t = a)$. When we constrain asset holdings (a) to a grid $A = [0 < a_1 < a_2... < a_n]$, the exogenous Markov chain on labor productivity shocks (denoted by $P(z, z')$) and the optimal saving decision by agents (denoted by $a' = g(z, a)$) induce a law of motion for $f_t(z, a)$:

$$f_{t+1}\left(z', a'\right) = \sum_z \sum_{\{a: a' = g(z,a)\}} f_t(z, a) P\left(z, z'\right),$$

where a prime denotes next-period variables and $\{a : a' = g(z, a)\}$ refers to all states of a that satisfy $a' = g(z, a)$, given a specific z. A stationary distribution F solves the above equation with $f_{t+1} = f_t$ (and so $F_{t+1} = F_t$). A direct way to compute F is to iterate to convergence on the equation. To prove the convergence, the so-called "supermartigale convergence theorem" is applied. Ljungqvist and Sargent (2012, Chapters 17–18) provided a detailed treatment on the issue. See also Heer and Maussner (2009, Chapter 7) for the detail on the concept of the stationary equilibrium.

Conditional on government consumption G and the income tax schedule $T(y)$, the stationary type distribution of agents does not change over time according to the definition of stationary equilibrium. This implies that the aggregates L and K are constant in stationary equilibrium and so are the factor prices w and r. Crucially, although the type distribution of agents is time invariant in stationary equilibrium, the state facing an agent may vary from one period to the next simply because idiosyncratic shocks still evolve stochastically. As a consequence, agents may move around in the time-invariant cross-sectional distribution even in stationary equilibrium. As long as the so-called "monotone mixing condition" is satisfied, there will be a unique stationary type distribution in equilibrium; see Heathcote et al. (2009) for the detail. This condition, which is met in our model, basically requires sufficient upward and downward social mobility for agents.

In Mirrlees (1971), taxation is imposed after each agent learns her own labor productivity (type) and so income taxes serve as a device for correcting the "inequitable" income distribution of ex post heterogeneous agents. Cremer and Gahvari (1999) and Chetty and Saez (2010) considered optimal taxation in the framework of Mirrlees (1971) but with a modification—all agents are behind the veil of ignorance, that is, taxation is imposed before each agent learns her own labor productivity. Income taxes then serve as a device for providing social insurance for ex ante homogeneous agents. Our model is in some sense a dynamic extension of the later models with two distinct characteristics. First, shocks occur only once in these models, whereas shocks occur repeatedly in our model. Secondly, agents face no risk ex post in these models, whereas agents face risk even in stationary equilibrium in our model since shocks still stochastically evolve in stationary equilibrium. At any time t in stationary equilibrium, agents have different histories of their realized labor productivity *up to time t*, but at the same time they are behind the veil of ignorance regarding the possible realization of their labor productivity *beyond time t*. As a result, our tax design problem involves both the ex post feature of the Mirrlees model and the ex ante feature of the Cremer–Gahvari–Chetty–Saez model and, therefore, it involves redistributing income for both equity and social insurance.

Only very few exceptions allow for the derivation of analytical results in a heterogeneous-agent DSGE model (Heer and Maussner, 2009, chapter 7). We numerically solve for the stationary equilibrium of our model.[11]

3. Parameterization

3.1. Production technology and preferences

The model will be calibrated to the U.S. economy. One period in our model is taken to be one year of calendar time. Except for the endogenous choice of labor hours and the imposition of income taxes, our model follows Aiyagari (1994) closely. As such, the chosen values of our parameters follow Aiyagari (1994) closely, too. Aiyagari noted that these parameter values for technology and preference are consistent with aggregate features of the postwar U.S. economy and are commonly used in macro models.

On the production side, following Aiyagari (1994), the capital share θ is set to 0.36 and the depreciation rate of capital δ is set to 0.08 per year. The parameter μ in the utility function is set to 2 in the baseline. We also consider higher values of μ. The borrowing constraint b is set to zero as in Aiyagari (1994) and Huggett (1997); see Heathcote (2005) for a discussion on this parameterization.

We set the discount factor $\beta = 0.96$ as in Aiyagari (1994) and choose the parameter $\phi = 0.4$ in the utility function so that average hours worked in the economy equal one-third of the time endowment when the income tax is only used to finance government consumption. The resulting capital-output ratio equals 3.3, which is in the range of values in the U.S. data; see, for example, Cooley and Prescott (1995) in which they match the capital-output ratio to 3.32.

3.2. Labor productivity shocks

In addressing optimal taxation in the static, deterministic framework of Mirrlees (1971), Saez (2001) advanced the literature by parameterizing the distribution of unobservable innate ability to match real-world data. Our approach here is in a similar vein, though in a dynamic, stochastic rather than static, deterministic framework.

Let $\omega_{i,t}$ denote the log of the real wage for agent i in period t. A fairly common setup to estimate the process for $\omega_{i,t}$ is as follows[12]:

$$\omega_{i,t} = \gamma x_{i,t} + \alpha_i + \chi_{i,t} + v_{i,t}; \quad v_{i,t} \sim N\left(0, \sigma_v^2\right), \tag{3}$$

where $x_{i,t}$ is a vector of observable characteristics of agents, α_i reflects agent i's unobserved fixed effect, $\chi_{i,t}$ represents a shock, and $v_{i,t}$ is a residual term that may capture measurement errors. The primal object of interest is the shock χ, which is related to our model shock z with $z = \exp(\chi)$. Other effects on $\omega_{i,t}$ are included so that $\chi_{i,t}$ can be better estimated. The shock $\chi_{i,t}$ is assumed to evolve according to a first-order autoregression (AR(1)) process:

$$\chi_{i,t} = \rho \chi_{i,t-1} + \varepsilon_{i,t}; \quad \varepsilon_{i,t} \sim N\left(0, \sigma_\varepsilon^2\right), \chi_{i,0} = 0,$$

where ρ determines the degree of persistence in shocks. The variates $v_{i,t}$ and $\epsilon_{i,t}$ are assumed to be independent of each other and of other variables across agents and over time.

The empirical estimates regarding the shock $\chi_{i,t}$ vary somewhat. Using the U.S. Panel Study of Income Dynamics (PSID) data and letting $\omega_{i,t}$ be annual labor earnings, Storesletten et al. (2004) found that $\rho = 0.99$ and $\sigma_\epsilon = 0.13$. Using PSID data as well but letting $\omega_{i,t}$ be the hourly wage relative to others, Floden and Linde (2001) found a less persistent but more volatile process with $\rho = 0.92$ and $\sigma_\epsilon = 0.21$. Our parameterization adopts Floden and Linde's estimates, mainly because the labor productivity term z in our model is more in alignment with the hourly wage than annual labor earnings (while labor hours are endogenously determined in Floden and Linde, they are supplied inelastically in Storesleten et al.).

Given the estimated AR(1) process of χ, we approximate $z = \exp(\chi)$ using a five-state Markov chain, on the range of plus and minus 2 times the standard deviation from the mean of z, as described in Tauchen (1986). The resulting labor productivity shock z takes five possible values:

$$z \in \{0.3424, 0.5852, 1.0000, 1.7089, 2.9202\},$$

with the transition matrix:

$$Pr\left(z'|z\right) = \begin{pmatrix} 0.8072 & 0.1925 & 0.0003 & 0.0000 & 0.0000 \\ 0.0694 & 0.7886 & 0.1418 & 0.0001 & 0.0000 \\ 0.0001 & 0.1010 & 0.7980 & 0.1010 & 0.0001 \\ 0.0000 & 0.0001 & 0.1418 & 0.7886 & 0.0694 \\ 0.0000 & 0.0000 & 0.0003 & 0.1925 & 0.8072 \end{pmatrix}.$$

---

[11] The algorithm for the numerical solution is standard and similar to that in Aiyagari (1994). To speed up computations, we apply the endogenous grid method for solving an agent's dynamic stochastic optimization problems, as introduced in Carroll (2006) and Barillas and Fernandez-Villaverde (2007); we also use the approach of non-stochastic simulation for generating model equilibrium features, as introduced in Young (2010).

[12] There is an alternative approach in which α_i in Eq. (3) is replaced by $\alpha_i + \beta_i t$, where β_i denotes differences in learning ability across agents. This approach allows for heterogeneous trends and usually leads to estimates with a lower persistency in labor productivity shocks; see Guvenen (2011).

The Markov chain above can be shown to approach a unique stationary distribution:

$$\Pi(z) = (0.0874, 0.2424, 0.3405, 0.2424, 0.0874).$$

Table 1 summarizes the parametric values used in our baseline model. Alternative values of some parameters will also be explored for the purpose of better understanding their impacts on optimal taxation.

4. Welfare criterion

We assume a utilitarian welfare criterion, which is the average of the expected lifetime values of all agents in stationary equilibrium:

$$SW = \int V(s)dF(s).$$

What is the welfare difference between an economy in autarky (no taxes imposed) and an economy with the operation of optimal taxes? Following Lucas (1987) and analogous to Low and Maldoom (2004), we report a measure for the welfare difference. This measure, ζ, known as the consumption-equivalent variation (CEV) in the literature, is implicitly defined by

$$\int_i E_0 \left[\sum_{t=0}^{\infty} \beta^t u\left(c_{i,t-\text{opt}}, 1-n_{i,t-\text{opt}}\right) \right]$$
$$= \int_i E_0 \left[\sum_{t=0}^{\infty} \beta^t u\left((1+\zeta)c_{i,t-\text{aut}}, 1-n_{i,t-\text{aut}}\right) \right],$$

where $c_{i,t-\text{opt}}$ and $n_{i,t-\text{opt}}$ are agent i's consumption and labor hours chosen at time t in the optimal-tax economy, and $c_{i,t-\text{aut}}$ and $n_{i,t-\text{aut}}$ are agent i's corresponding choices in the autarky economy. In words, ζ represents a markup (if $\zeta > 0$) or markdown (if $\zeta < 0$) of agent consumption in the autarky economy such that the utilitarian autarky society would be as well off as that in the utilitarian optimal-tax society. Note that ζ measures a proportionate increase or decrease in consumption in all periods and for all agents.

Utilizing our specified period utility function gives rise to

$$\int_i E_0 \left[\sum_{t=0}^{\infty} \beta^t u\left(c_{i,t-\text{opt}}, 1-n_{i,t-\text{opt}}\right) \right]$$
$$= (1+\zeta)^{\phi(1-\mu)} \int_i E_0 \left[\sum_{t=0}^{\infty} \beta^t u\left(c_{i,t-\text{aut}}, 1-n_{i,t-\text{aut}}\right) \right].$$

This then leads to

$$\zeta = \left(\frac{SW_{\text{opt}}}{SW_{\text{aut}}}\right)^{\frac{1}{\phi(1-\mu)}} - 1,$$

where SW_{opt} and SW_{aut} denote SW under the optimal tax and the autarky economy, respectively. As a reference, we also report ζ resulting from a Rawlsian social welfare criterion.

In addition to ζ, we report Gini coefficients for income and wealth distributions resulting from the imposition of taxes.

Table 1
Summary of parameters.

Parameters	θ	δ	μ	ϕ	β	b	ρ	σ
Baseline	0.36	0.08	2	0.4	0.96	0	0.92	0.21
Variants	–	–	{3,4}	–	–	–	{0,1}	{0.315, 0.42}

5. Optimal linear income tax

To present the results more clearly, we first consider a pure tax-transfer system without government consumption ($G = 0$). This means that the total tax revenue collected by the government will uniformly be transferred back to all agents in a lump-sum manner.

The resulting social welfare with respect to the income tax rate, in terms of CEV, displays a hump shape as plotted in Fig. 1. The value of CEV reaches its highest level at the tax rate 16%, and the welfare improvement from the autarky to the optimal-tax economy equals a 0.39% increase in lifetime consumption. This improvement in welfare is rather moderate. However, as will be shown later, if the agent risk aversion, μ, or the standard deviation of shocks, σ_ϵ, increases, the resulting values of CEV will become much higher.

Taxation distorts both labor–leisure and consumption–saving decisions in our model. From Figs. 2 and 3, we see that aggregate labor hours and asset holdings are both monotonically decreasing in the tax rate. This is the cost side of imposing income tax. However, imposing income tax has its beneficial side in that it redistributes income for equity and social insurance. The resulting social benefit against the cost of tax distortions leads to a hump shape for CEV as displayed in Fig. 1.

Fig. 4 presents the patterns of the Gini coefficient for after-tax total income, pre-tax labor income, and wealth as the income tax rate varies. The distribution of after-tax total income in terms of the Gini coefficient is improving in the tax rate. However, the distributions of pre-tax labor income and of wealth are both deteriorating in the tax rate. We report in Table 2 that our model generates the right empirical ranking between wealth and income inequality as in Aiyagari (1994): wealth is more dispersed than income.

Fig. 5 shows that the fraction of agents choosing to supply zero labor and those choosing to hold zero assets are both increasing in the tax rate. At the optimal linear income tax rate (16%), 7.15% of agents supply zero labor and 12.38% of agents hold zero assets.[13] It has been known that optimal income taxation is compatible with the outcome that a non-trivial fraction of agents do not work (Kaplow, 2008, Chapter 4). Hubbard et al. (1995) argued that nil asset holdings can be explained as a utility-maximizing response to the government's welfare programs. We show here that optimal income taxation is also compatible with the outcome that a non-trivial fraction of agents do not accumulate any wealth.

It is worth knowing how distributions of labor hours and asset holdings change as tax rates increase. We report changes for the top and the bottom 20% of the distributions. As expected, increases in tax rates discourage labor hours and asset holdings of both the top 20% and the bottom 20% of agents. However, the magnitudes are quite different in general. For example, when the tax rate is increased from 0 to 20%, the average labor hours (asset holdings) of the top 20% will decrease by 8.5% (22.9%), whereas they will decrease by 61.5% (59.6%) for the bottom 20%. In particular, as shown in Fig. 5, the fractions of the bottom agents who supply zero labor or hold zero assets become higher as tax rates increase. These lead to increasing Gini coefficients of assets and of gross earnings in the tax rate; see Fig. 4. The results are mainly due to: (i) while the tax-transfer system does not provide sufficient insurance for the rich, it does smooth consumption fairly well for the poor and so substantially reduces their precautionary-saving motives; (ii) the tax-transfer system overall transfers income from the rich to the poor and so it delivers negative income effects to the rich but positive effects to the poor; these income effects enhance the labor supply of the rich but reduce the labor supply of the poor.

Table 2 summarizes the results for the autarky economy and the optimal-tax economy with a utilitarian social welfare function. For reference, we also report the optimal results for a Rawlsian social welfare function, in that the government's aim is to maximize the

---

[13] It is interesting to note from our simulation that no agents supply zero labor and hold zero assets at the same time.
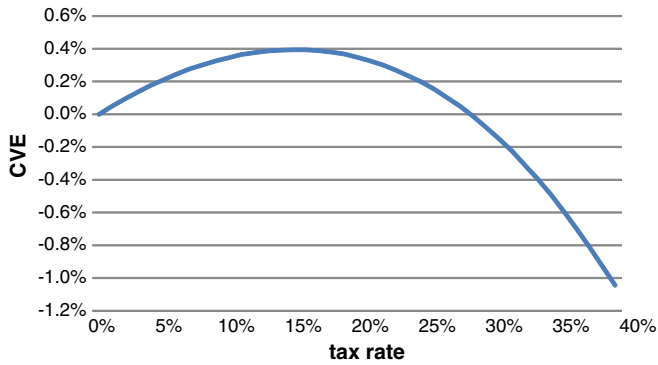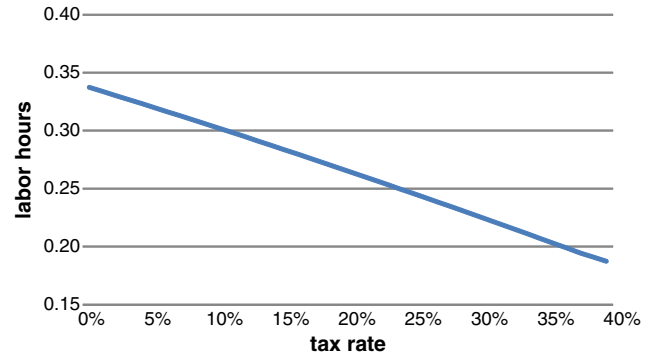
Fig. 1. Utilitarian CEV.

Fig. 2. Aggregate labor hours.

expected lifetime values of agents who accumulate zero wealth in stationary equilibrium.[14] It is not surprising to find that, with the Rawlsian objective, the optimal tax rate becomes higher and the distribution of after-tax total income improves, while aggregate labor hours and asset holdings become lower and the distributions of pre-tax labor income and of wealth deteriorate relative to the utilitarian objective. At the optimum, 8.46% of agents supply zero labor and 12.85% of agents hold zero assets with the Rawlsian objective. Both are higher compared to the utilitarian objective.

5.1. Self-insurance through increasing saving and labor supply

When facing uninsurable shocks, agents will exhibit precautionary incentives and self-insure against income drops. Following Low and Maldoom (2004), we address the comparative static effects of increasing risk aversion (μ) and of increasing shock volatility (σ_ϵ). A main difference between Low and Maldoom (2004) and our paper is that while self-insurance must take place through the channel of increasing labor supply in the Low–Maldoom model, it can take place through the channel of increasing saving as well as that of increasing labor supply in our model.

Papers including Low (2005), Floden (2006), Pijoan-Mas (2006) and Marcet et al. (2007) explored the insurance value of increasing labor supply, and the interaction between labor supply and saving as self-insurance mechanisms. We complement their findings in the context of optimal taxation.

5.1.1. Increasing risk aversion

When agents become more risk averse, sharing idiosyncratic risk between agents should be more valuable. We perform the same numerical simulations as in the baseline economy but allow for higher degrees of risk aversion for agents. The results are reported in Table 3.

The optimal tax rates and associated asset holdings become somewhat higher if $\mu = 3, 4$ rather than $\mu = 2$. However, the resulting CEV compared to the baseline economy improves significantly, jumping from a 0.39% increase in lifetime consumption to 1.26% if $\mu = 3$ and to 2.66% if $\mu = 4$. This finding is consistent with that in Low and Maldoom (2004), which shows a higher value of income tax acting as an insurance device against income fluctuations as agents become more risk averse.

5.1.2. Increasing shock volatility

When agents are subject to larger volatile shocks, sharing idiosyncratic risk between agents should also be more valuable. We perform the same numerical simulations as in the baseline economy, except that we allow for standard deviations of productivity shocks at 1.5 and 2 times the original (i.e., $1.5\sigma_\epsilon$ and $2\sigma_\epsilon$). The results are reported

in Table 4. Note that both the optimal tax rate and the CEV increase substantially compared to the baseline economy. This result is consistent with Mirrlees (1990), which shows that adding uncertainty always increases the marginal tax rate of the linear income tax.

Low and Maldoom (2004) observed that there are two opposite effects on incentives as shock volatility increases. On the one hand, larger risk directly increases agents' precautionary incentives (direct effect). On the other hand, larger risk raises the value of insurance, and the resulting higher social insurance provided by a benevolent government indirectly lowers agents' precautionary incentives (indirect effect). It is interesting to note from Table 4 that, along with increased optimal tax rates, aggregate labor hours decrease significantly while aggregate asset holdings increase significantly relative to the baseline. Thus, in response to increased shock volatility (σ_ϵ), the indirect effect dominates the direct effect in the case of self-insurance via labor hours, but the opposite occurs in the case of self-insurance via asset holdings. Low and Maldoom (2004) also found that the indirect effect dominates the direct effect when individuals self-insure by increasing labor hours. However, due to the static nature of their model, they did not have the opposite result that the direct effect dominates the indirect effect when individuals self-insure by increasing asset holdings.

5.2. Labor–leisure versus consumption–saving tax distortion

Income tax distorts both labor–leisure and consumption–saving decisions in our model. How important is one distortion relative to the other in the design of income tax? To answer the question, we perform counter-factual exercises by shutting down one distortion at a time.

5.2.1. Fixed labor hours

To shut down the tax distortion imposed on the labor–leisure decision, we fix labor hours for all agents at the equilibrium average level of the optimal-tax baseline economy. That is, we let labor hours be fixed at $n = 0.28$ for all agents and all the time (remember that
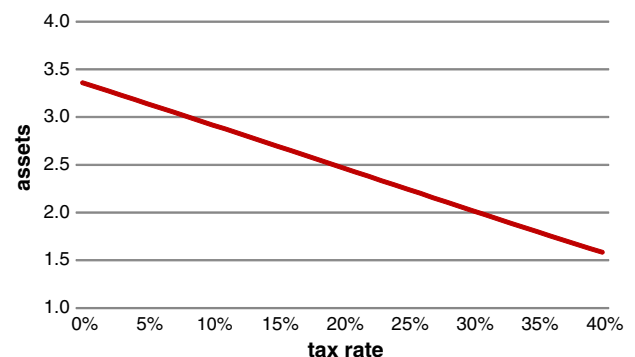
---

[14] Agents who supply zero labor hours have higher expected lifetime values than those who hold zero assets in our model.

Fig. 3. Aggregate asset holdings.

Fig. 4. Gini coefficient.

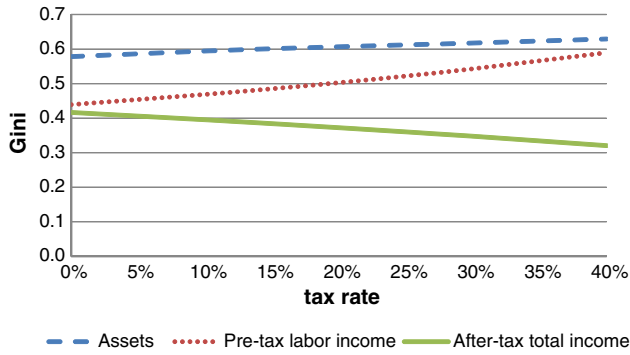| Economy	Tax τ	Incentives		Gini coefficient			Output	
		Hours n	Assets a	Wealth	Gross earnings	Net income	Y	CEV
No tax	0%	0.34	3.36	0.58	0.44	0.42	0.92	NA
Opt. tax (Utilitarian)	16%	0.28	2.65	0.60	0.49	0.38	0.78	0.39%
Opt. tax (Rawlsian)	21%	0.26	2.43	0.61	0.51	0.37	0.73	0.82%

our economy is populated by a unit mass of agents).[15] Income tax would not affect the labor–leisure decision at all in such a counter-factual case. The second row of Table 5 reports the results. When labor hours were fixed, the optimal tax rate would increase significantly from 16% in the baseline economy to 30%.

The analytical investigations conducted by Floden (2006) show that the flexibility of labor supply will raise precautionary saving in the face of future uncertainty when preferences are consistent with balanced growth. Low (2005) obtained a similar result in a calibrated life-cycle model. Table 5 shows that agents' asset holdings would decrease relative to the baseline if labor hours were fixed rather than flexible (1.68 vs. 2.65). Our result is consistent with these previous findings.

5.2.2. Fixed asset holdings

To shut down the tax distortion imposed on the consumption–saving decision, we fix asset holdings for all agents at the equilibrium average level of the optimal-tax baseline economy.[16] That is, we let asset holdings be fixed at $a = 2.65$ for all agents and all the time. Income tax would not affect the consumption–saving decision at all in such a counter-factual case. The third row of Table 5 reports the results. When asset holdings were fixed, the optimal tax rate would increase substantially from 16% in the baseline economy to 34%.

To sum up, we conclude that neglecting either of the tax distortions imposed on labor–leisure and consumption–saving decisions will substantially underestimate the extent of tax distortion and hence substantially overestimate the optimal degree of progressivity for income taxation. This highlights the importance of taking into consideration both incentive margins at the same time in a model.

The tax-distortion view above can also be interpreted from a different angle. Agents self-insure against risk through the channel of adjusting labor supply or saving in our model. If labor hours or asset holdings were fixed, agents would not be able to hedge their risk as much as in the situation where both labor hours and asset holdings are flexible. This largely explains why the optimal tax rate would become much higher as shown in Table 5: the inflexibility of the private sector against risk expands the scope of income tax as a device for insurance. This also explains why *CEV* from optimal taxation in the baseline economy is rather small when both labor hours and asset holdings are flexible.

5.3. Taxation for equity versus for social insurance

Agents are heterogeneous in our model, in the sense that they have different histories of realizations of idiosyncratic shocks to their labor productivity and that these shocks are highly persistent. Below we compare our model with two counter-factual polar cases: (i) the occurrence of productivity shocks were to follow an i.i.d. process (i.e., no persistency at all) all the time, and (ii) the occurrence of productivity shocks were to be once and for all (i.e., perfect persistency plus no volatility) from the very beginning. The comparison will enable us to address the relative importance of redistributive taxation for equity versus for social insurance.

5.3.1. Transitory (i.i.d.) shock

Let $\rho = 0$ in the AR(1) process, so that labor productivity shocks would exhibit no persistency at all. The labor productivity states are assumed to remain the same as before, but their occurrence in each and every period is to be determined by an i.i.d. process, for which we use the same stationary distribution of labor productivity implied by the original Markov chain, that is, $\Pi(z)$. In such a counter-factual case where labor productivity shocks are i.i.d. so as to be purely transitory, agents become more or less homogeneous since they all have more or less the same expected lifetime value at any point of time. As a consequence, the tax scheme acts mainly as a device of providing social insurance for agents, since shocks still stochastically evolve in stationary equilibrium.

The results are reported in Table 6 in the second row, which indicates that the optimal tax rate would equal 0%, a corner solution. This outcome implies that the cost of tax distortion would dominate the benefit of risk sharing provided by income tax and, therefore, self-insurance rather than income tax should be relied on to take care of idiosyncratic risk if the shocks were to follow an i.i.d. process. Intuitively, when idiosyncratic shocks are transitory in nature, they are easier to smooth through self-insurance by individuals and so greatly reduce the value of income tax as a device to insure against income fluctuations.

Fig. 5. Ratio of zero labor supply and of zero asset holdings.

---

[15] We also consider the other case where labor hours for all agents were fixed at the equilibrium average level of the no-tax economy. The resulting outcomes remain the same qualitatively.

[16] We also consider the other case where asset holdings for all agents were fixed at the equilibrium average level of the no-tax economy. The resulting outcomes remain the same qualitatively.

Table 3
Optimal linear tax: various risk aversions.

Economy	Opt. tax τ	Incentives		Gini coefficient			Output	
		Hours n	Assets a	Wealth	Gross earnings	Net income	Y	CEV
$\mu = 2$ (baseline)	16%	0.28	2.65	0.60	0.49	0.38	0.78	0.39%
$\mu = 3$	18%	0.27	2.69	0.58	0.49	0.37	0.78	1.26%
$\mu = 4$	19%	0.27	2.79	0.56	0.49	0.37	0.79	2.66%

5.3.2. Permanent shock

We consider the other counter-factual polar case where productivity differences between agents were permanent from the very beginning. The same five labor productivity states are employed as before and their distribution remains the same as the stationary distribution implied by the original Markov chain, that is, $\Pi(z)$. The only difference from the baseline is that there is no longer a transition of labor productivity shocks and each agent will carry her initial realization of productivity shocks forever. This implies no volatility of shocks, either. In such a case where labor productivity shocks are permanent from the very beginning, our model reduces to a model of pure adverse selection as in Mirrlees (1971) and the tax scheme acts solely as a device for equity.

The results are reported in Table 6 in the third row, which indicates that the optimal tax rate would equal 17%, a result that is pretty close to the optimal tax rate of 16% in the baseline economy. Overall, the optimal taxation in the baseline economy seems to mainly serve as a device for equity rather than for social insurance. This finding is reminiscent of Levine and Zame (2002), who showed that market incompleteness in the presence of idiosyncratic shocks has little effect on welfare, provided that agents are long-lived and patient to self-insure against the risk.

It is worth noting that the aggregate amount of asset holdings in equilibrium will be reduced significantly compared to the baseline (1.81 vs. 2.65). This is largely due to the absence of precautionary incentives to accumulate assets once shocks become permanent at the very beginning.

5.4. Government consumption

So far we have considered a pure tax-transfer system without government consumption ($G = 0$). In the real world, government consumption does exist and is required to be financed by tax revenue. We hereby set G as of 17% of GDP (equal to output Y in our model) as in Conesa et al. (2009), and investigate the impact of incorporating G on the optimal linear tax scheme. A minimum linear tax rate to support G as of 17% of GDP (Y) is equal to 24% of income (y).

In Fig. 6 we plot the resulting pattern of CEV with respect to the income tax rate in the presence of G. The optimal tax rate equals 16% without G, while it equals 33% as G is 17% of GDP. Table 7 shows that the amount of the transfer as a percentage of income y at the optimum decreases from 16% in the absence of G to 10% in the presence of G. It has long been known in the literature (Inman, 1987) that the provision of government consumption will crowd out the provision of government transfers. Our finding here corroborates this result. Table 7 also shows that Gini coefficients remain roughly the same with and without G,

Table 4
Optimal linear tax: various shock volatilities.

Economy	Opt. tax τ	Incentives		Gini coefficient			Output	
		Hours n	Assets a	Wealth	Gross earnings	Net income	Y	CEV
σ_ε (baseline)	16%	0.28	2.65	0.60	0.49	0.38	0.78	0.39%
$1.5\sigma_\varepsilon$	31%	0.18	2.98	0.63	0.70	0.45	0.87	3.00%
$2\sigma_\varepsilon$	41%	0.12	3.30	0.65	0.80	0.44	1.05	7.11%

Table 5
Optimal linear tax: fixed hours vs. asset holdings.

Economy	Opt. tax τ	Hours n	Assets a	Output Y
Baseline	16%	0.28	2.65	0.78
Hours fixed	30%	0.28	1.68	0.59
Assets fixed	34%	0.24	2.65	0.71

and that a higher optimal tax rate exerts a relatively larger effect on capital accumulation than labor supply if compared to the baseline economy. Finally, the resulting CEV has a 0.27% increase in lifetime consumption.

The most celebrated work on the quantitative characterization of optimal linear income tax is perhaps Stern (1976), who considered a static, deterministic model à la Mirrlees (1971). Because there are several important differences in modeling, it may not be so meaningful to make a comparison between his findings and ours. Nevertheless, it is worth noting that, with a utilitarian social welfare criterion and G as 20% of GNP, his central estimate yields an optimal tax rate of 25%. This level is lower than our optimal tax rate of 33% with G as 17% of GDP.[17]

6. Alternative settings

This section considers several alternative settings and examines the robustness of the results found in the previous section.

6.1. Government debt

We rule out government debt in our model. Allowing for its presence, the government's intertemporal budget constraint becomes

$$G_t + (1 + r)D_t = \int T[y_t(s)]dF_t(s) + D_{t+1},$$

where D_t is the level of government debt at the beginning of period t.

Following Aiyagari and McGrattan (1998), we assume that government debt is risk free and a perfect substitute for the private risk-free asset. As such, the asset held by agents, a, now also includes the debt issued by the government.

In the presence of government debt, the main changes of a stationary equilibrium are in the government budget constraint and the market clearing conditions:

- Given (w, r), the government's budget constraint is met, i.e., $G + rD = \int T[y(s)]dF(s)$.
- All markets are clear:[18]

$$K + D = \int adF(s);$$
$$L = \int zn(s)dF(s);$$
$$C + I + G + rD = Y = K^\theta L^{1-\theta}.$$

This revised model is close to Aiyagari and McGrattan (1998) but with an important difference: the transfer g is endogenous in our model, while it is exogenously fixed in the Aiyagari–McGrattan model.

---

[17] It should be noted that Stern's (1976) simulations are mainly designed to explore the sensitivity of optimal linear tax rates with respect to the revenue requirement, the elasticity of substitution between consumption and leisure, and the form of the social welfare function.

[18] Let κ denote the debt to GDP ratio and so $D = \kappa Y$. We know capital income $(r + \delta)K = \theta Y$ and hence $Y = \frac{r+\delta}{\theta}K$. From the asset market clearing condition,

$$K + \frac{\kappa(r + \delta)}{\theta}K = \int adF(s) \Rightarrow K = \frac{\theta}{\theta + \kappa(r + \delta)}\int adF(s). \qquad (4)$$

Table 6
Optimal linear tax: transitory vs. permanent shock.

Economy	Opt. tax τ	Hours n	Assets a	Output Y
Baseline	16%	0.28	2.65	0.78
Transitory (i.i.d) shock	0%	0.26	2.76	0.86
Permanent shock	17%	0.30	1.81	0.65

The findings, reported in Table 8, are summarized below.[19]

- The optimal debt level is zero and, therefore, the optimal tax is the same as that in the baseline model. This result differs from the positive level of optimum debt reported by Aiyagari and McGrattan (1998). The benefit of having positive debt in the Aiyagari type of models is to loosen the agent's borrowing constraint and drive up the interest rate (due to the crowding-out of private capital) so that agents can use savings to smooth consumption more effectively; see Aiyagari and McGrattan (1998) for the detail. However, in our environment the presence of government debt will also crowd out transfer g so as to mitigate risk sharing provided by the tax-transfer system.

- If we take the debt to output ratio as given, then the resulting optimal tax rate does not change significantly. For example, the optimal tax rate will be 15% rather than 16% in the baseline economy as the debt to output ratio equals 60%.

- If we consider the same setting as in Aiyagari and McGrattan (1998), that is, we fix a transfer level, it is possible to yield a positive level of government debt at the optimum as in Aiyagari and McGrattan (1998). However, the optimal debt level is sensitive to the discount factor β, as found in Aiyagari and McGrattan. We also confirm that moving to an optimal debt level does not gain much social welfare as noted by Aiyagari and McGrattan. We find in general that the welfare gain is lower than 0.1% in *CEV*.

Overall, our investigation suggests that there is not much loss of generality if we abstract our model economy from government debt.

6.2. Separate taxation of labor and capital income

The tax schedule does not discriminate between labor and capital income in our model. Allowing for discrimination leads to the following tax schedule

$T = \tau_l wzn + \tau_k ra - g,$

where τ_l is the marginal tax rate applied to pre-tax labor income, and τ_k is the marginal tax rate applied to pre-tax capital income.

Table 9 reports the optimal tax rates on labor income under various given tax rates imposed on capital income. We find that the optimal capital tax rate is equal to zero ($\tau_k = 0$) and the corresponding optimal labor tax rate is 18%. The resulting *CEV* from the no-tax economy is 0.47%, which is higher than that in the baseline model (*CEV* = 0.39%). It can be seen from the table that a higher capital tax results in both a lower optimal labor tax and a lower welfare gain.

In their startling work, Chamley (1986) and Judd (1985) showed that government policy should set the tax rate on capital to zero in the steady state.[20] Conesa et al. (2009) argued that incomplete markets or life cycle features can cause optimal taxation on capital to deviate from the Chamley/Judd type of results. By incorporating both incomplete markets and the life cycle in a framework, they found the optimal
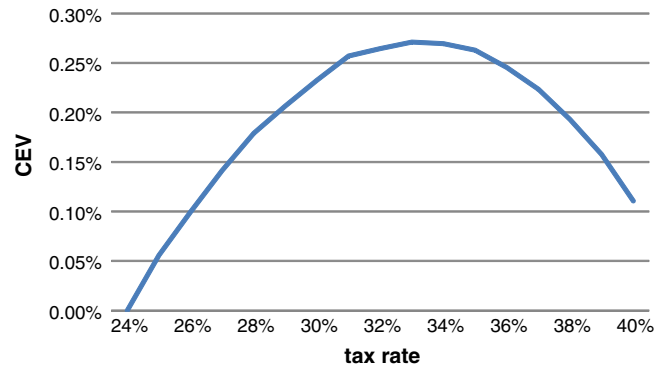
Fig. 6. Utilitarian CEV with government consumption.

capital tax rate to be as high as 36%. Our result suggests that the life cycle rather than incomplete markets could be the key to their finding. Indeed, in offering an intuition for their result, Conesa et al. (2009, p. 41) explicitly explained: "in a life-cycle model in which household labor supply changes with age, if the government cannot condition the tax function on age, it optimally uses the capital income tax to mimic age-dependent labor income taxes."[21]

6.3. Permanent productivity differences

We focus on idiosyncratic earnings risk and exclude permanent productivity differences between agents in our model. To allow for a permanent component of labor productivity, we modify $z_{i,t} = \exp(\chi_{i,t})$ to become

$$z_{i,t} = \exp\left(\chi_{i,t} + \psi_i\right),$$

where ψ_i is a permanent component of agent i's labor productivity. ψ_i is assumed to be i.i.d. distributed with mean zero and variance σ_ψ^2. $\chi_{i,t}$, a non-permanent productivity shock, is assumed to remain the same as before. χ and ψ are independent.

Floden and Linde (2001) have estimated the permanent labor productivity and reported that $\sigma_\psi^2 = 0.1175(\sigma_\psi = 0.34)$ for the U.S. economy. We consider a simple extension of our model in which there exist two permanent-productivity groups with $\psi \in \{-0.17, 0.17\}$ to match $\sigma_\psi = 0.34$. The resulting labor productivity shock z becomes:

$z(\Psi = \text{-}0.17) \in \{0.2889, 0.4937, 0.8437, 1.4417, 2.4637\};$
$z(\Psi = 0.17) \in \{0.4059, 0.6936, 1.1853, 2.0255, 3.4613\}.$

Table 10 reports the results. Because the distribution of the permanent productivity is symmetric and the productivity shock volatility faced by agents remains the same, we find that the incorporation of the permanent component does not affect the results significantly. In particular, the resulting optimal tax rates and *CEV*'s become higher but both are not substantially different from those in the baseline economy (compared with Table 2).

---

[19] Values of *CEV* in the table are to measure welfare gains from the no-tax and no-debt economy.

[20] When preferences are weakly separable between consumption and labor, Atkinson and Stiglitz (1976) have implied the same result. However, their result is not applicable here, since our utility function rules out the separability.

[21] A critical element of the life-cycle model is to explicitly account for the age dimension of agents. The recent literature has suggested that it may be desirable to consider age-dependent tax schedules; see, for example, Blomquist and Micheletto (2008) and Weinzierl (2011). In the incomplete market framework of Aiyagari (1994), Aiyagari (1995) showed that if a government can choose its consumption optimally, the optimal tax on capital income will be positive. This result is applicable neither to Conesa et al. (2009) nor to our paper, since government consumption is assumed exogenously given in both papers.

Table 7
Optimal linear tax: plus government consumption.

Economy	Opt. τ	Transfer/income	Incentives		Gini coefficient			Y	CEV
			Hours n	Assets a	Wealth	Gross earnings	Net income		
Baseline	16%	16%	0.28	2.65	0.60	0.49	0.38	0.78	0.39%
With G	33%	10%	0.28	2.33	0.59	0.49	0.38	0.75	0.27%

7. Optimal two-bracket income tax

The shape of the income tax schedule is a central focus in the literature on optimal income taxation. Specifically, many attempts have been made to yield an optimal income tax schedule that is consistent with what we typically observe in the real world, that is, progressive marginal tax rates; see Salanie (2003, Chapter 4) for a review. The two-bracket income tax – a direct extension of the linear income tax – could offer important insights into the issue.

Table 11 summarizes the results of the optimal two-bracket tax for the baseline economy with the utilitarian and the Rawlsian social welfare criterion, respectively. Several features of the results are worth noting.

First, in the case of the utilitarian, $CEV = 0.44\%$ at the optimum, which is close to $CEV = 0.39\%$ in the linear case; in the case of the Rawlsian, $CEV = 0.86\%$ at the optimum, which is also close to $CEV = 0.82\%$ in the linear case (compared with Table 2). Our result is thus similar to Slemrod et al. (1994), which finds that the social welfare gain from using a two-bracket tax rather than being confined to a linear tax is minor. However, it is important to recognize that our model, though complex enough, is after all a stylized and simple model of any real economy. For this reason we would regard the result as a reference rather than a verdict.

Second, the optimal marginal tax rates are regressive since the second marginal tax rate is lower than the first one. The result holds, regardless of whether social welfare is utilitarian or Rawlsian. We shall explore later why this regressive marginal tax schedule arises at the optimum.

Third, the optimal cut points equal 1.43 and 1.36 of mean income (y), respectively, for the utilitarian and the Rawlsian social welfare functions. It is found that at the optimum the fraction of agents in the first bracket is 70.03% and that in the second bracket is 29.97% for the utilitarian function (those for the Rawlsian function are 67.94% and 32.06%).

Fourth, aggregate labor hours, aggregate asset holdings, outputs, and Gini coefficients remain roughly the same as those in the optimal linear tax (compared with Table 2). This may explain why CEV improvements are small from the extension of the linear to the two-bracket tax.

In the case of the linear income tax, it is found that 7.15% of agents supply zero labor and 12.38% of agents hold zero assets at the optimum if the social welfare is evaluated in terms of the utilitarian criterion (the corresponding numbers are 8.46% and 12.85% if evaluated in terms of the Rawlsian criterion). The fractions become somewhat higher in

the case of the two-bracket income tax; they are 8.66% and 13.39%, respectively (the corresponding numbers for the Rawlsian are 12.48% and 14.64%).[22] The higher fractions could be due to the regressive structure of the marginal tax rates in the optimal two-bracket income tax.

7.1. Increasing risk aversion vs. shock volatility

Table 12 reports the results if agents become more risk averse, and Table 13 reports the results if agents face larger volatility of idiosyncratic shocks.[23] These results are qualitatively similar to those in the linear-tax case (compared with Tables 3 and 4). For example, if shock volatility becomes larger, CEV increases substantially, and aggregate labor hours decrease significantly while aggregate asset holdings increase significantly. The most interesting finding is perhaps that the optimal two-bracket tax schedule changes from regressive to progressive in the marginal tax rate as σ_ε in the baseline economy is increased to $1.5\sigma_\epsilon$ and $2\sigma_\epsilon$. We explain this result below.

Kaplow (2008, p. 64) provided an intuition for why the optimal two-bracket tax schedule derived by Slemrod et al. (1994) is regressive in the marginal tax rate. A high first-bracket tax rate is inframarginal rather than marginal to high-income agents and thus it collects a high amount of tax revenue from high-income agents without distorting their labor supply. On the other hand, a low second-bracket tax rate induces high labor effort from high-income agents and thus it collects a high amount of tax revenue from high-income agents without affecting low-income agents.

The above intuition fails to apply once incomes are driven by luck. First, a high first-bracket tax rate may not be inframarginal to high-income agents' effort since their high incomes may simply stem from good luck, not from labor effort. Second, a low second-bracket tax rate may not induce high labor effort from high-income agents since once again their high incomes may simply stem from good luck, not from labor effort. In fact, as demonstrated by Strawczynski (1998), the second-bracket tax rate should be set at 100% if high incomes are completely due to pure luck. This 100% tax-rate result is in stark contrast to the celebrated result of the "zero distortion at the top" established by Sadka (1976) and Seade (1977) in the Mirrlees (1971) framework.

The first row of Table 13 shows that the resulting two-bracket tax schedule is regressive rather than progressive in the marginal tax rate. It implies that the regressive ability-driven force underlying Slemrod et al. (1994) dominates the progressive luck-driven force underlying Strawczynski (1998) in the baseline economy. However, it seems logical to conjecture that if future uncertainty facing agents becomes sufficiently high so that the progressive luck-driven force underlying Strawczynski (1998) is strong enough, then the opposite result with progressive marginal tax rates will arise at the optimum. This

Table 8
Optimal linear tax under various debt levels.

Debt/output D/Y	Optimal tax τ	Transfer/income g/y	Welfare gain CEV
0%	16%	16.0%	0.39%
10%	15%	14.6%	0.37%
20%	15%	14.3%	0.36%
60%	15%	12.8%	0.27%

---

[22] Similar to the optimal linear tax, it is interesting to note that none supply zero labor and hold zero assets at the same time in the case of the optimal two-bracket tax.
[23] Many comparative-static results for the two-bracket are qualitatively similar to those for the linear. To avoid repetition, we have not reported them.

Table 9
Optimal labor taxes under various capital taxes.

Capital tax τ_k	Optimal labor tax τ_l	Hours n	Assets a	Welfare gain CEV
0%	18%	0.28	2.75	0.47%
10%	17%	0.28	2.68	0.41%
20%	13%	0.29	2.69	0.33%
30%	13%	0.29	2.55	0.23%
40%	12%	0.29	2.41	0.10%

conjecture is indeed borne out as shown in the second and the third rows of Table 13 when volatility σ_ϵ in the baseline economy is increased to $1.5\sigma_\epsilon$ and $2\sigma_\epsilon$.

Kaplow's provided intuition takes the shape of the wage rate or productivity type distribution as given. As such, it may fail to apply too if the underlying wage rate distribution changes substantially. This possibility is explored in a recent paper by Apps et al. (forthcoming), which emphasizes the important driver of the wage rate distribution in shaping the optimal tax structure. Similar to Slemrod et al. (1994), they considered a piecewise linear income tax in the framework of Mirrlees (1971), but argued that the parameters of the lognormal wage distributions employed by Slemrod et al. (taken from Stern, 1976) are based on data in the late 1960's/early 1970's; however, wage inequality has increased dramatically since then. Employing Pareto rather than lognormal wage distributions to match recent real-world data, they showed that marginal tax rate progressivity is optimal when the wage distribution is at first relatively flat and then rises steeply in the higher deciles, a pattern currently observed in many OECD countries.

It is known that the second moment of a lognormal distribution is finite, whereas the second moment of a Pareto distribution is not finite under plausible parameter values. This may explain why as far as modeling the recent dramatic increase in wage inequality is concerned, the upper tail of wage rates is much better modeled by a Pareto distribution than a lognormal distribution. Sticking to the lognormal assumption, a possible way of capturing the recent dramatic rise in earnings inequality is to let the variance of a lognormal productivity-type distribution increase substantially. This increase leads to a substantial increase in earnings inequality (in terms of Gini coefficients) in stationary equilibrium. We find that the Gini coefficients of earnings for σ_ϵ, $1.5\sigma_\epsilon$ and $2\sigma_\epsilon$ in the absence of tax equal 0.44, 0.58 and 0.67, respectively. This result is not surprising in that an increased shock volatility implies not only ex ante an increase in earnings risk, but also ex post an increase in earnings inequality once shocks are realized. Table 13 shows that the increased shock volatility turns the optimal two-bracket tax schedule from regressivity to progressivity in the marginal tax rate. Although we adopt a different approach, our result is consistent with the finding of Apps et al. (forthcoming). That is, the two-bracket tax schedule should be progressive rather than regressive in the marginal tax rate once earnings risk/inequality has increased to a large extent.

Table 10
Optimal linear tax with permanent component of labor productivity.

Economy	Tax τ	Incentives		Gini coefficient			Output Y	CEV
		Hours n	Assets a	Wealth	Gross earnings	Net income		
No tax	0%	0.34	3.38	0.58	0.45	0.42	0.93	NA
Opt. tax (Utilitarian)	17%	0.27	2.63	0.61	0.50	0.39	0.74	0.51%
Opt. tax (Rawlsian)	23%	0.25	2.36	0.62	0.52	0.37	0.72	1.04%

Table 11
Optimal two-bracket tax in the baseline economy.

	Opt. tax			Hours	Assets	Output	
	τ_1	τ_2	y_0	n	a	Y	CEV
Utilitarian	0.2	0.1	$1.43y$	0.28	2.71	0.79	0.44%
Rawlsian	0.3	0.2	$1.36y$	0.24	2.28	0.70	0.86%

	Gini coefficient		
	Wealth	Gross earnings	Net income
Utilitarian	0.61	0.51	0.40
Rawlsian	0.63	0.55	0.38

Table 12
Optimal two-bracket tax: various risk aversions.

Economy	Opt. tax			Hours n	Assets a	Output Y	CEV
	τ_1	τ_2	y_0				
$\mu = 2$ (baseline)	0.2	0.1	$1.43y$	0.28	2.71	0.79	0.44%
$\mu = 3$	0.2	0.1	$1.42y$	0.28	2.87	0.81	1.36%
$\mu = 4$	0.3	0.1	$1.47y$	0.25	2.80	0.79	2.73%

8. Upper tail of earnings distribution

This section makes an attempt to account for the upper tail of the earnings distribution in the real world. The exploration will enable us to gain further insights into the issue of marginal tax rate progressivity or regressivity at the optimum.

It is assumed in our model that the labor productivity shock follows an AR(1) process with a stationary lognormal distribution. We adopt an estimated AR(1) process using the PSID data and employ a five-state Markov chain to approximate this process. The first and the second rows of Table 14 provide a comparison of the earnings distribution between the U.S. data and our baseline economy in stationary equilibrium (the income tax rate is set at 35%).[24] The Gini coefficient from our model is lower than, but not far away from, the actual data; however, the model does not capture the earnings share of the top income group well—the top 5% hold 31.2% of total earnings in the data, while only 21.2% in the baseline model.

Castaneda et al. (2003) argued that reported earnings data such as the PSID are not specifically concerned with a careful measurement of the earnings in the top tail of the earnings distribution and, moreover, they are often subject to a significant amount of top-coding; as a result, econometric studies based on these data may not be reliable to addressing the upper tail of earnings inequality. To get around the problem, they instead used the U.S. Lorenz curve of earnings reported in Table 14 to calibrate the process of stochastic productivity shocks. What they did in the end was to allow for the incidence of some abnormally high productivity shock with a low probability. This approach enables them to generate a highly skewed earnings distribution in a dynamic framework and, in particular, match the earnings share of the very rich in the U.S. data.

To account for the upper tail of the earnings distribution in the U.S. economy, we follow an approach similar to Castaneda et al. (2003). Specifically, we modify the original Markov chain of the productivity shock, $z \in \{z_1, z_2, ..., z_5\}$, by in addition incorporating an abnormally high productivity shock z_h. To construct a new shock process with the addition of z_h, denoted by $\hat{z} \in \{z_1, z_2, ..., z_5, z_h\}$, we assume that only those agents who receive z_5 have a chance to reach z_h with a

---

[24] The U.S. data are compiled by Castaneda et al. (2003) using the 1992 Survey of Consumer Finances, and they are confirmed by many other empirical studies; see Castaneda et al. (2003).

Table 13
Optimal two-bracket tax: various shock volatilities.

Economy	Opt. tax			Hours	Assets	Output	
	τ_1	τ_2	y_0	n	a	Y	CEV
σ_ε (baseline)	0.2	0.1	$1.43y$	0.28	2.71	0.79	0.44%
$1.5\sigma_\varepsilon$	0.1	0.3	$0.18y$	0.20	2.95	0.88	3.35%
$2\sigma_\varepsilon$	0.2	0.4	$0.29y$	0.14	3.57	1.10	7.16%

probability 0.1. Agents who currently receive z_h are assumed to either stay at z_h or move down to z_5 in the next period. The probability of staying at z_h, $Pr\left(\hat{z}' = z_h | \hat{z} = z_h\right)$, is assumed to equal that of staying at z_5, $Pr(z' = z_5 | z = z_5)$.

We calibrate the value of z_h such that the model Gini coefficient of earnings can be close to the data, 0.63. As a result, $\hat{z}$ takes six possible values:

$$\hat{z} \in \{0.3424, 0.5852, 1.0000, 1.7089, 2.9202, 5.5547\},$$

with the transition matrix:

$$\Pr\left(\hat{z}' | \hat{z}\right) = \begin{pmatrix} 0.8072 & 0.1925 & 0.0003 & 0.0000 & 0.0000 & 0.0000 \\ 0.0694 & 0.7886 & 0.1418 & 0.0001 & 0.0000 & 0.0000 \\ 0.0001 & 0.1010 & 0.7980 & 0.1010 & 0.0001 & 0.0000 \\ 0.0000 & 0.0001 & 0.1418 & 0.7886 & 0.0625 & 0.0069 \\ 0.0000 & 0.0000 & 0.0003 & 0.1925 & 0.7265 & 0.0807 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.1928 & 0.8072 \end{pmatrix}.$$

The implied stationary distribution of $\hat{z}$ is:

$$\Pi(\hat{z}) = (0.0836, 0.2319, 0.3257, 0.2319, 0.0836, 0.0434).$$

The distribution of earnings in the model economy with the incorporation of the new shock process $\hat{z}$ and the 35% income tax rate is presented in the third row of Table 14. Note that the earnings of the top 5% in the data are captured much better with the modified shock process in the model (31.2% in the data vs. 31.7% in the model).

Table 15 reports the optimal linear and two-bracket income taxes with the new shock process under the title of "High-shock economy" along with "Baseline economy." The optimal linear tax rate becomes 25% and the resulting CEV is 1.19%. Both figures are significantly higher than those in the baseline economy.

Most importantly, we find that the optimal two-bracket tax structure becomes progressive in the marginal tax rate with $\tau_1 = 0.2$ and $\tau_1 = 0.4$ rather than $\tau_1 = 0.2$ and $\tau_2 = 0.1$ in the baseline economy. The intuition for this result is arguably similar to that of increasing shock volatility before. Thus, this result is also consistent with the finding of Apps et al. (forthcoming).

9. Conclusion

Following Stern (1976), Slemrod et al. (1994), Strawczynski (1998), and Apps et al. (forthcoming), this paper quantitatively characterizes optimal linear and two-bracket income taxes but synthesizes these previous studies in a framework where both ability and luck matter for the determination of individual income. Substantive findings include: (i) a significant fraction of agents supply zero labor or hold zero assets at the optimum; (ii) neglecting tax distortion imposed on either of

Table 14
Distribution of earnings.

Economy	Gini	Bottom 40%	Top 20%	Top 10%	Top 5%
US data	0.63	3.2	61.4	43.5	31.2
Baseline $\tau = 0.35$	0.566	5.6	56.9	37.4	21.2
High-shock $\tau = 0.35$	0.638	3.9	64.2	47.3	31.7

Table 15
Optimal tax with high shock z_h.

Tax system	Opt. tax			Gini coefficient			
	τ_1	τ_2	y_0	Wealth	Earnings	Net income	CEV
Baseline economy							
Linear	0.00	0.16	0.00	0.60	0.49	0.38	0.39%
Two-bracket	0.2	0.1	$1.43y$	0.61	0.51	0.40	0.44%
High-shock economy							
Linear	0.00	0.25	0.00	0.66	0.59	0.40	1.19%
Two-bracket	0.2	0.4	$1.39y$	0.65	0.56	0.37	1.37%

labor–leisure and consumption–saving decisions will lead to the prescription of tax codes that deviate substantially from the optimum; and (iii) the optimal two-bracket tax schedule will turn from regressive to progressive in the marginal tax rate once the volatility of idiosyncratic shocks becomes sufficiently large. The last finding is consistent with the results in Apps et al. (forthcoming), and it also reconciles the contradictory results regarding the optimal two-bracket tax schedule between Slemrod et al. (1994) and Strawczynski (1998).

Several abstractions of our model from real-world complications are worth noting. First, we assume homogeneous preferences and exclude preference heterogeneity across agents. Second, we confine our analysis to individual-specific shocks and disallow for aggregate shocks. Third, we consider the two simplest forms of the piecewise linear income tax (linear and two-bracket) and do not go beyond them.[25] Finally, channels of insurance against shocks are limited to agents' self-insurance and the government's provided taxes and transfers. Relaxing these abstractions should be worth investigating.

References

Aiyagari, S.R., 1994. Uninsured idiosyncratic risk and aggregate saving. Quarterly Journal of Economics 109, 659–684.
Aiyagari, S.R., 1995. Optimal capital income taxation with incomplete markets, borrowing constraints, and constant discounting. Journal of Political Economy 103, 1158–1175.
Aiyagari, S.R., McGrattan, E.R., 1998. The optimum quantity of debt. Journal of Monetary Economics 42, 447–469.
Apps, P., Long, N.V., Rees, R., 2013. Optimal piecewise linear income taxation. Journal of Public Economic Theory (forthcoming).
Atkinson, A.B., Stiglitz, J.E., 1976. The design of tax structure: direct versus indirect taxation. Journal of Public Economics 6, 55–75.
Barillas, F., Fernandez-Villaverde, J., 2007. A generalization of the endogenous grid method. Journal of Economic Dynamics and Control 31, 2698–2712.
Barro, R.J., 1991. A cross-country study of growth, saving, and government. In: Bernheim, D.G., Shoven, J.B. (Eds.), National Saving and Economic Performance. University of Chicago Press, Chicago.
Blomquist, S., Micheletto, L., 2008. Age-related optimal income taxation. The Scandinavian Journal of Economics 110, 45–71.
Boadway, R., Sato, M., 2011. Optimal income taxation with uncertain earnings: a synthesis. CESifo Working Paper No. 3654.
Brewer, M., Saez, E., Shephard, A., 2010. Mean-testing and tax rates on earnings. In: Besley, T., Blundell, R., Gammie, M., Poterba, J. (Eds.), Dimensions of Tax Design: The Mirrlees Review. Oxford University Press, Oxford.
Carroll, C.D., 2006. The method of endogenous gridpoints for solving dynamic stochastic optimization problems. Economics Letters 91, 312–320.
Castaneda, A., Diaz-Gimenez, J., Rios-Rull, J., 2003. Accounting for the U.S. earnings and wealth inequality. Journal of Political Economy 111, 818–857.
Chamley, C., 1986. Optimal taxation of capital income in general equilibrium with infinite lives. Econometrica 54, 607–622.
Chetty, R., 2012. Bounds on elasticities with optimization frictions: a synthesis of micro and macro evidence on labor supply. Econometrica 80, 969–1018.
Chetty, R., Saez, E., 2010. Optimal taxation and social insurance with endogenous private insurance. American Economic Journal: Economic Policy 2, 85–116.
Chetty, R., Guren, A., Manoli, D., Weber, A., 2011. Are micro and macro labor supply elasticities consistent? A review of evidence on the intensive and extensive margins. American Economic Review Papers and Proceedings 101, 1–6.
Conesa, J.C., Krueger, D., 2006. On the optimal progressivity of the income tax code. Journal of Monetary Economics 53, 1425–1450.
Conesa, J.C., Kitao, S., Krueger, D., 2009. Taxing capital? Not a bad idea after all! American Economic Review 99, 25–48.

---

[25] Conesa et al. (2009, p. 34) noted: "Ideally one would impose no restrictions on the set of tax functions the government can choose from. Maximization over such an unrestricted set is computationally infeasible, however." This infeasibility is known as the "curse of dimensionality" in numerical analyses (Heer and Maussner, 2009).

Cooley, T.F., Prescott, E.C., 1995. Economic growth and business cycles. In: Cooley, T.F. (Ed.), Frontiers of Business Cycle Research. Princeton University Press, Princeton, NJ.

Cremer, H., Gahvari, F., 1999. Uncertainty, commitment, and optimal taxation. Journal of Public Economic Theory 1, 51–70.

Diamond, P.A., Helms, L.J., Mirrlees, J.A., 1980. Optimal taxation in a stochastic economy: a Cobb–Douglas example. Journal of Public Economics 14, 1–29.

Eaton, J., Rosen, H., 1980. Optimal redistributive taxation and uncertainty. Quarterly Journal of Economics 95, 357–364.

Floden, M., 2006. Labor supply and saving under uncertainty. The Economic Journal 116, 721–737.

Floden, M., Linde, J., 2001. Idiosyncratic risk in the United States and Sweden: is there a role for government insurance? Review of Economic Dynamics 4, 406–437.

Golosov, M., Tsyvinski, A., Werning, I., 2006. New dynamic public finance: a user's guide. NBER Macroeconomics Annual 21, 317–363.

Guvenen, F., 2011. Macroeconomics with heterogeneity: a practical guide. Economic Quarterly 97, 255–326.

Heathcote, J., 2005. Fiscal policy with heterogeneous agents and incomplete markets. Review of Economic Studies 72, 161–188.

Heathcote, J., Storesletten, K., Violante, G., 2009. Quantitative macroeconomics with heterogeneous households. Annual Review of Economics 1, 319–354.

Heer, B., Maussner, A., 2009. Dynamic General Equilibrium Modeling, 2nd ed. Springer-Verlag, Berlin.

Hubbard, R.G., Skinner, J., Zeldes, S.P., 1995. Precautionary saving and social insurance. Journal of Political Economy 103, 360–399.

Huggett, M., 1997. The one-sector growth model with idiosyncratic shocks: steady states and dynamics. Journal of Monetary Economics 39, 385–403.

Huggett, M., Ospina, S., 2001. Aggregate precautionary savings: when is the third derivative irrelevant? Journal of Monetary Economics 48, 373–396.

Inman, R.P., 1987. Markets, governments, and the "new" political economy. In: Auerbach, A.J., Feldstein, M. (Eds.), Handbook of Public Economics, vol. II. Elsevier, Amsterdam.

Judd, K.L., 1985. Redistributive taxation in a simple perfect foresight model. Journal of Public Economics 28, 59–83.

Kanbur, R., Pirttila, J., Tuomala, M., 2008. Moral hazard, income taxation and prospect theory. The Scandinavian Journal of Economics 110, 321–337.

Kaplow, L., 2008. The Theory of Taxation and Public Economics. Princeton University Press, Princeton.

Keane, M.P., Rogerson, R., 2011. Reconciling micro and macro labor supply elasticities: a structural perspective. NBER Working Paper, p. 17430.

Kocherlakota, N.R., 2010. The New Dynamic Public Finance. Princeton University Press, Princeton.

Laffont, J.-J., Martimort, D., 2002. The Theory of Incentives: The Principal-Agent Model. Princeton University Press, Princeton.

Levine, D.K., Zame, W.R., 2002. Does market incompleteness matter? Econometrica 70, 1805–1839.

Ljungqvist, L., Sargent, T.J., 2012. Recursive Macroeconomic Theory, 3rd ed. MIT Press, Cambridge.

Low, H., 2005. Self-insurance in a life-cycle model of labor supply and savings. Review of Economic Dynamics 8, 945–975.

Low, H., Maldoom, D., 2004. Optimal taxation, prudence and risk-sharing. Journal of Public Economics 88, 443–464.

Lucas Jr., R.E., 1987. Models of business cycles. Yrjo Jahnsson Lectures Series.Blackwell, London.

Marcet, A., Obiols-Homs, F., Weil, P., 2007. Incomplete markets, labor supply and capital accumulation. Journal of Monetary Economics 54, 2621–2635.

Mirrlees, J.A., 1971. An exploration in the theory of optimum income taxation. Review of Economic Studies 38, 175–208.

Mirrlees, J.A., 1974. Notes on welfare economics, information and uncertainty. In: Balch, M., McFadden, D., Wu, S. (Eds.), Essays in Equilibrium Behavior under Uncertainty. North-Holland, Amsterdam.

Mirrlees, J.A., 1990. Taxing uncertain incomes. Oxford Economic Papers 42, 34–45.

Pijoan-Mas, J., 2006. Precautionary savings or working longer hours? Review of Economic Dynamics 9, 326–352.

Sadka, E., 1976. On income distribution, incentive effects, and optimal income taxation. Review of Economic Studies 9, 261–268.

Saez, E., 2001. Using elasticities to derive optimal income tax rates. Review of Economic Studies 68, 205–229.

Salanie, B., 2003. Economics of Taxation. MIT Press, Cambridge.

Seade, J., 1977. On the shape of optimal income tax schedules. Journal of Public Economics 7, 203–236.

Sheshinski, E., 1989. Note on the shape of the optimum income tax schedules. Journal of Public Economics 40, 201–215.

Slemrod, J., Yitzhaki, S., Mayshar, J., Lundholm, M., 1994. The optimal two-bracket linear income tax. Journal of Public Economics 53, 269–290.

Sorensen, P.B., 1994. From the global income tax to the dual income tax: recent tax reforms in the Nordic countries. International Tax and Public Finance 1, 57–79.

Stern, N.H., 1976. On the specification of models of optimum income taxation. Journal of Public Economics 6, 123–162.

Storesletten, K., Telmer, C., Yaron, A., 2004. Consumption and risk sharing over the life cycle. Journal of Monetary Economics 51, 609–633.

Strawczynski, M., 1998. Social insurance and the optimum piecewise linear income tax. Journal of Public Economics 69, 371–388.

Tauchen, G., 1986. Finite state Markov-chain approximations to univariate and vector autoregressions. Economics Letters 20, 177–181.

Tuomala, M., 1984. Optimal degree of progressivity under income uncertainty. The Scandinavian Journal of Economics 86, 184–193.

Tuomala, M., 2010. On optimal non-linear income taxation: numerical results revisited. International Tax and Public Finance 17, 259–270.

Varian, H., 1980. Redistributive taxation as social insurance. Journal of Public Economics 14, 49–68.

Weinzierl, M., 2011. The surprising power of age-dependent taxes. Review of Economic Studies 78, 1490–1518.

Young, E.R., 2010. Solving the incomplete markets model with aggregate uncertainty using the Krusell–Smith algorithm and non-stochastic simulations. Journal of Economic Dynamics and Control 34, 36–41.