

Age-dependent taxes with endogenous human capital formation*

Carlos E. da Costa

FGV-EPGE

carlos.eugenio@fgv.br

Marcelo R. Santos

INSPER

marcelors2@insper.edu.br

February 14, 2015

Abstract

We calculate optimal age-dependent labor income taxes in an environment for which the age efficiency profile is endogenously determined by human capital investment. Heterogeneous individuals are exposed to idiosyncratic shocks to their human capital investments, a key element, along with the endogeneity of human capital itself in the determination of optimal age-dependent taxes. Our results highlight the complementary role of capital income taxation and the quantitative relevance of accounting for general equilibrium effects. *Keywords:* Age-dependent taxes; Human Capital Accumulation *J.E.L. codes:* E6; H3; J2.

1 Introduction

[Akerlof \[1978\]](#) first used the term 'tag' to refer to characteristics that groups of people may possess which should be taken into account in order to alleviate the distortions caused by taxation. Despite the fact that the knowledge of its importance dates back to as early as [Akerlof's \[1978\]](#) work, age, a natural tag for tax policy, has received little attention until recently. This is unfortunate since age is free from most of the drawbacks that might have precluded the practical uses of other tags.¹ A series of recent contributions in the realm of the New Dynamic Public Finance literature have

*Carlos da Costa gratefully acknowledges financial support from CNPq. We thank Luis Araujo, Felipe Iachan, seminar participants at IPEA, FEA-RP, and the 2013 REAP meeting their invaluable comments. All errors are our sole responsibility.

¹See [Weinzierl \[2012\]](#) for a discussion of possible reasons for the sub-optimal use of tags in current policies.

however risen the interests on the power of age-dependent taxes — [Weinzierl \[2011\]](#), [Farhi and Werning \[2013\]](#), [Findeisen and Sachs \[2014\]](#).²

A common characteristic of all aforementioned works is the assumption that the age-efficiency profile is invariant to policy. The main goal of the current work is to assess whether this simplifying assumption may lead to relevant errors in optimal tax prescriptions by explicitly taking into account the endogenous nature of life-cycle productivity profiles.

The exogeneity of an agent's productivity along his or her life-cycle is clearly an inaccurate description of how one's productivity evolves. Yet, the simplifications allowed for by this assumption is thought to outweigh its costs. This fact is often explicitly recognized in the literature. [Weinzierl \[2011\]](#), for example, referring to the use of an exogenous path for wages along the life-cycle, argues that “The specific results of this paper therefore require that a substantial portion of variation of wages with age is inelastic to taxes. A few considerations suggest that this requirement's effects on the paper's results may be limited.” The underlying view is, therefore, that assuming the age-efficiency profile to be exogenous does not generate important quantitative departures from what is believed to be a better description of the evolution of productivity along the life-cycle.

It is our view, however, that one should not take for granted that the costs of such simplification are really small. Recent works by [Kapicka \[2006\]](#), [Kapička \[2011\]](#), [Best and Kleven \[2013\]](#) endogenizes human capital formation in a [Mirrlees's \[1971\]](#) framework and derives policy prescriptions that are the opposite of that found in [Weinzierl \[2011\]](#) and [Farhi and Werning \[2013\]](#). The latter find that taxes ought to increase with age whereas the former suggest that they should decrease.³ It is important to emphasize that a key element driving [Weinzierl's \[2011\]](#), [Farhi and Werning's \[2013\]](#) and [Findeisen and Sachs's \[2014\]](#) results are absent in [Erosa and Gervais \[2002\]](#), [Garriga \[2003\]](#), [Kapicka \[2006\]](#), [Kapička \[2011\]](#): an increased cross-sectional dispersion of skills due to idiosyncratic labor productivity shocks.

We write a model of human capital investment where insurance concerns are key in policy design. The economy we study incorporates endogenous human capital for-

²It is not to say that the idea of age dependent taxes is new. [Erosa and Gervais \[2002\]](#), [Garriga \[2003\]](#), for example, connected the usefulness of taxing capital income to the optimality of age-dependent taxes. They have shown that capital income taxes should be used in an overlapping generations economy for which it is optimal but not possible to make labor income taxes vary with age.

³Decreasing labor income taxes are also optimal in [Erosa and Gervais \[2002\]](#) and [Garriga \[2003\]](#) in an environment without idiosyncratic uncertainty. They do not have endogenous human capital. It is the way the chosen functional forms of preferences induce the elasticity of labor supply to behave along the life-cycle that drives their results.

mation to a setting which is essentially that of [Conesa et al. \[2009\]](#): an overlapping generations economy calibrated to account for the age-efficiency profiles found in the data. We consider two extreme versions of human capital investments: a Ben-Porath type of human capital of investment, which we refer to as learning-or-doing, and a learning-by-doing specification, for which it is the very act of working that increases one's future productivity.

We find that allowing for endogenous human capital formation preserves the overall findings of [Weinzierl \[2011\]](#), [Farhi and Werning \[2013\]](#): marginal tax rates ought to increase with age. Because we rule out negative marginal tax rates in our exercises, the optimal tax system is characterized by the yearly increase in the marginal tax rate and the age at which an individual starts paying taxes.⁴ In the learning-by-doing model, the elasticity of labor supply is lower for young individuals, thus leading to higher taxes at early ages than for the model with exogenous age-efficiency profile. We find the opposite to be true for the learning-or-doing model. Individuals should start paying taxes later and the marginal tax rates should increase faster than in the model with exogenous human capital. The quantitative effects may be substantial. At age 35, for example the difference between optimal marginal tax rates for the learning-by-doing and the learning of doing model is around fifteen percentage points.⁵

Endogeneity also changes the sensitivity of age-dependence to the level of capital income taxation. Optimal labor income taxes respond more strongly to variations of capital income taxes when human capital is endogenous. The same is true, and in a more important sense, when we consider the general equilibrium effects. Optimal labor income taxes display greater response to changes in capital income taxes for all models but they are much more pronounced when human capital is endogenous.

As for welfare gains, we find them to be substantial for all models. Key to these welfare gains is the relaxation of credit constraints. By delaying the moment at which one starts paying taxes, the government relaxes the restriction imposed by the imperfection in capital markets.

Accumulation of physical capital increases substantially as we move taxes from the early to the late stages of the life-cycle. As earnings become more concentrated earlier in life, agents increase their savings, which in our model takes the form of investments in physical capital. As a result we find substantial variation in the capital-income ratio when we move from the current system to the optimal age-dependent one.

⁴We also allow for an exemption level, but, as we shall discuss, we find it to be optimally set to zero at all ages.

⁵We also calculate the costs of using the policy derived under the assumption that productivity is exogenous when it is in fact endogenous. The welfare costs are not negligible but are small when compared to the gains from introducing age-dependence in the first place.

To relate our results with the recent discussion in [Krueger and Ludwig \[2013\]](#), we explore the role of progressivity in the tax system by allowing lower incomes to be exempt from taxes. We find that the optimal exemption level is *zero* for all specifications of human capital accumulation when taxes are allowed to depend on age. It is important to bear in mind that, for all but the learning-by-doing model, the exemption levels are non-zero if we rule out age-dependence. [Krueger and Ludwig \[2013\]](#) mention progressivity as a substitute for age-dependent taxes. Our findings regarding the welfare gains suggest it to be a very imperfect one. Note also that the degree of progressivity we find to be optimal for an age-independent tax system is lower than the current one.

Finally, the role of general equilibrium effects in the form of producer prices adjustments is also explored. When we rule out these adjustments, the degree of age dependence falls substantially and all but disappears when the marginal tax on capital income is set to zero and human capital is endogenous. In contrast, if the age-efficiency profile is exogenous age-dependence remain important. Welfare gains with respect to the benchmark become negative for all models.

The rest of the paper is organized as follows. After a brief literature review, Section 2 explains the environment as well as the policy instruments available to the Government, with the definition of a recursive equilibrium for the economy presented in 2.6. Details of the calibration and the planner's maximization problem are presented in Section ???. Our main findings are found in Section 4 and we use sections 4.2 and 4.3 to discuss some variations of the model. Section 5 concludes the paper.

Literature Review

The idea that labor income taxes ought to vary with time is certainly not new. A direct application of Ramsey formulae found in [Lucas and Stokey \[1983\]](#), [Chari and Kehoe \[1999\]](#), [Judd \[1985\]](#), for example, leads to the conclusion that, barring very specific functional forms for preferences, taxes ought to respond to productivity changes. If productivity evolves, then taxes ought to vary with time.

In dynastic settings time and age are not distinguished so time-varying taxes were not initially associated with the need for age-dependent taxes.⁶ These however become important when we focus on life cycle aspects of human behavior.

The combination of insurance as a central concern of public policy, and the non restrictive choice of policy instruments led those working in the New Dynamic Public Finance (NDPF) literature to start exploring the value of conditioning taxes on age.⁷

⁶Early models of optimal taxation in a dynamic setting as [Lucas and Stokey \[1983\]](#) emphasized the role of spreading distortions across periods, but never related their findings to age-dependent taxes.

⁷This literature has evolved by extending [Mirrlees's \[1971\]](#) methodology to a dynamic setting. That

Indeed, under this modelling tradition taxes should optimally depend on one's history, which, in particular leads to different tax schedules for agents of different ages. Soon some began to explicitly take into account tax systems for which the history of earnings is ruled out but age-dependence is still allowed. Best known amongst these are [Farhi and Werning \[2013\]](#), [Weinzierl \[2011\]](#), which asked how much of the gains attained from fully optimal tax systems obtains from using linear age-dependent taxes. [Bastani et al. \[2010\]](#), [Findeisen and Sachs \[2013\]](#) have also ruled out history dependence but allowed non-linearity in the age-dependent taxes.

All these papers find that it is optimal for taxes to increase with age. This is mostly driven by the insurance aspect of taxation. None of the papers discussed so far deal with the endogeneity of human capital. Human capital endogeneity is, however, central for [Kapička \[2006\]](#), [Kapička \[2011\]](#), [Kapička and Neira \[2013\]](#), [?](#), all still in the realm of the NDPF.

[Kapička \[2006\]](#), [Kapička \[2011\]](#) rule out idiosyncratic risk and find that marginal tax rates ought to *decrease* with age. Human capital formation takes the form of a learning-or-doing technology meaning that human capital acquisition requires one to forgo current labor income to increase future productivity. This type of technology leads labor supply elasticity to be higher when one is young since schooling provides another alternative to working besides leisure. Although this creates a motive for one to increase taxes at early ages, another force related to distortions in portfolio formation may lead to optimal taxes in the opposite direction whenever the government does not want to bias investment against physical capital — see [Peterman \[2011\]](#).

[Kapička and Neira \[2013\]](#) combines uncertainty and human capital accumulation in a dynamic Mirrlees' economy. In a two period model with additively separable preferences, they find that the marginal tax rate faced by each individual increases on average from the first to the second period. Because the government has full control of everyone's consumption, very sophisticated capital income taxes may be needed, e.g., [Kocherlakota \[2004\]](#).⁸ Endogeneity of human capital is also investigated by [?](#), which allow for arbitrary non-linear taxes. As all other works in this tradition there are not different cohorts of individuals that inhabit the economy at the same time.

Following a very different tradition, [Keane \[2011\]](#) investigates optimal labor income taxation along one's life cycle. [Keane's \[2011\]](#) main concern is to understand the role of human capital accumulation in determining the elasticity of labor supply to labor income taxes. He uses the findings in [Imai and Keane \[2004\]](#) of a very large omis-

is, policy instruments are not defined a priori and are only restricted by the informational structure of the economy.

⁸See, however, [Werning \[2010\]](#).

sion bias in the most commonly estimations of labor supply elasticities to highlight the potentially substantial underestimation of welfare costs of taxation. Age-dependent taxes are not explicitly discussed, even though it is related to his discussion of temporary versus permanent tax changes. [Keane \[2011\]](#) focuses on a finite economy with exogenous prices, which we believe, misses some important consequences of adopting age-dependent taxes. Overlapping generations settings we view as ideal for this task.

More closely related to our work are those that exploit the distinction between time and age that characterize overlapping generations economies. The macroeconomic aspects that drove the findings in [Lucas and Stokey \[1983\]](#), [Chamley \[1986\]](#), [Judd \[1985\]](#) can thus be disentangled from the life cycle concerns that drive the usefulness of age-dependent taxes in overlapping generations models. That is, on the one can take seriously the life-cycle wage profile of individuals, which typically display large variations across ages, while at the same time taking into account the fact that time variation in taxes is not the same as age variation.

Exploring these possibilities offered by overlapping generations models, [Erosa and Gervais \[2002\]](#), [Garriga \[2003\]](#) show how the absence of age-dependent taxes may lead to the optimality of non-zero capital income taxation. The idea is that capital income taxation substitute, however imperfectly, for the missing age-dependent labor income taxes. Using a representative agent for each cohort, and assuming away any form of uncertainty, they are able to derive optimal tax prescriptions and conditions under which age-dependent taxes are unnecessary. [Erosa and Gervais \[2002\]](#), [Garriga \[2003\]](#) find positive taxes on capital to be optimal in their setting. They relate this finding to marginal tax rates which ought to decrease with age. As we have seen, this stands in contrast with the findings of [Weinzierl \[2011\]](#), [Farhi and Werning \[2013\]](#) and [Find-eisen and Sachs \[2013\]](#). Neither [Erosa and Gervais \[2002\]](#) nor [Garriga \[2003\]](#) consider redistributive nor insurance motives.

[Conesa et al. \[2009\]](#) incorporates both dimensions of policy concern: redistribution and insurance. Their focus is, again, capital income taxes, not age-dependent labor income taxes. Because the two are closely connected in the OLG setting, it is important to mention their findings.⁹ For their preferred parametrizations of the model, [Conesa et al. \[2009\]](#) find large positive taxes. Relating it to decreasing age dependent taxes along the lines of [Erosa and Gervais \[2002\]](#), [Garriga \[2003\]](#) is not correct however. [Conesa et al. \[2009\]](#) allow for progressive income taxes within the three parameter family of tax functions introduced by [Gouveia and Strauss \[1994\]](#). Because marginal tax rates increase with income, and potential income varies with ages, marginal tax

⁹Only in Section V.B. do they explicitly discuss the role of age-dependent taxes, and relate the two (p. 41): "...a positive capital income tax mimics a labor income tax that is falling in age."

rates will ultimately vary with age as well. Moreover, interpreting their results is made more difficult by the fact that the progressive tax they use is levied on both labor and capital incomes.

We take into account elements from all these literatures in our investigation. The overlapping generations structure separates the time from the age dimension of policies. Idiosyncratic uncertainty is considered, to capture the insurance role of age-varying policy. We allow for two different forms of human capital accumulation: learning-or-doing and learning-by-doing. These two forms of human capital accumulation lead to very different incentives to work at different ages, and may potentially lead to very different policies. Finally, we explicitly investigate the role of general equilibrium effects caused by the impact of policies on both human and physical capital accumulation in generating our results.

It is finally worth mentioning recent work by [Krueger and Ludwig \[2013\]](#). They study optimal progressive income taxation when there is endogenous human capital choices through education. Education is a binary choice variable made in an exogenously specified period of one's life. They do not allow for age dependent taxes.

2 The Environment

At each point in time, the economy is inhabited by multiple cohorts of individuals of different ages. Each cohort is comprised of a continuum of measure one of individuals who live for a finite, albeit random, number of periods.

2.1 Demography

Each period, j , a new generation is born. For an individual born in period j , uncertainty regarding the time of death is captured by the fact that everyone faces a probability ψ_{t+1} of surviving to the age $t + 1$ conditional on being alive at age t . Hence, an individual born in j is alive in $j + t$ with probability $\prod_{k=1}^t \psi_k$. We also assume that there is $T > 0$ such that $\psi_{T+1} = 0$.

Our focus is on one's working life, hence an agent's life starts at the age $t = 16$.

Because all our analysis will be concentrated in the steady-state, we shall drop all time indices from aggregate variables and use t to represent age.

We may map the survival probability straight into the time invariant age profile of the population denoted $\{\mu_t\}_{t=1}^T$. It is modelled by assuming that the fraction of agents

t years old in the population is given by the following law of motion

$$\mu_t = \frac{\psi_t}{(1 + g_n)} \mu_{t-1},$$

where g_n denotes the population growth rate. We normalize $\{\mu_t\}_{t=1}^T$ so that $\sum_{t=1}^T \mu_t = 1$.

2.2 Households

Preferences Individuals derive utility from consumption, c , and leisure, l .

Preferences defined over random paths of (c_t, l_t) are represented by the time-separable von-Neumann Morgenstern utility,

$$\mathbb{E} \left[\sum_{t=1}^T \beta^{t-1} \left(\prod_{k=1}^t \psi_k \right) U(c_t, l_t) \right], \quad (1)$$

where β is the subjective discount factor, and \mathbb{E} is the expectation operator conditional on information at birth.

In most of what follows temporary utility will be of the form

$$U(c_t, l_t) = \frac{(c_t^{1-\rho} l_t^\rho)^{1-\gamma}}{1-\gamma}, \quad (2)$$

for $\rho \in (0, 1)$, $\gamma > 0$, $\gamma \neq 1$.

Note that this specification for preferences implies a Frisch elasticity of labor supply which decreases with hours worked. This property of preferences represented by (2) has played a role in the findings in [Erosa and Gervais \[2002\]](#). Since the data exhibits a pattern of decreasing hours along the life-cycle, optimal taxes should *decrease* with age in their model.

Labor Supply and Retirement Every period, individuals choose labor supply, consumption, human capital investment and asset accumulation to maximize their objective, (1), subject to a budget constraint which we shall explain momentarily.

Each person has a unit time endowment which can be directly consumed in the form of leisure, l , or used in market related activities. An individual of age t who works for n hours supplies to the market a total of $n_t s_t e^{(u+z_t)}$ efficiency units which are paid at a rental rate w . The variable $u \sim \mathcal{N}(0, \sigma_u^2)$ is a permanent component of an individual's skills. It is realized at birth and retained throughout one's life. On the other hand, z

evolves stochastically according to an AR(1) process, $z_t = \varphi_z z_{t-1} + \varepsilon_t$, with innovations $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$.

Whereas u aims at capturing the heterogeneity at birth, everyone's most relevant lottery, z is the main source of uncertainty that affects one's choices. The parameter φ_z accommodates the empirically observed persistence of productivity shocks. s_t is what we call the age-efficiency profile, the term which distinguishes the models of human capital we study.

Labor productivity shocks are independent across agents. As a consequence, there is no uncertainty regarding the aggregate labor endowment even though there is uncertainty at the individual level.

Retirement is mandatory at the age of 65, or $t = 50$.

Human Capital Accumulation At the center of our analysis is the process governing the age-efficiency profile s_t . Absent uncertainty, s_t would be the only term leading individuals to vary their choices along the life-cycle. It is the assumptions that we make about how s_t is determined that will differentiate the three models we present here.

Most of the optimal taxation literature treats s_t as exogenous.¹⁰ We shall consider, besides this, two alternative formulations for human capital accumulation: Learning-by-doing and Learning-or-doing.

Leaning-by-doing In the case of *learning-by-doing*, individuals accumulate human capital by working. That is, the law of motion for s is given by

$$s_{t+1} = \pi s_t^{\phi_s} n_t^{\phi_n} + (1 - \delta_h) s_t, \quad (3)$$

where (ϕ_s, ϕ_n) are parameters that govern both the persistence of the age-efficiency profile and the impact of hours worked on its evolution.

Each agent's time constraint is $l_t + n_t = 1$.

Leaning-or-doing With a Ben-Porath or *leaning-or-doing* technology for human capital accumulation, an agent acquires human capital by spending time, e_t , training in periods in which he or she is also working. The law of motion for human capital is:

$$s_{t+1} = \pi s_t^{\phi_s} e_t^{\phi_e} + (1 - \delta_h) s_t. \quad (4)$$

¹⁰Notable exceptions are [Kapička \[2011\]](#), [Peterman \[2011\]](#), [Keane \[2011\]](#).

The time constraint is now given by $n_t + l_t + e_t = 1$.

These two models of human capital accumulation are likely to have different implications on the design of the optimal tax policy. Under the leaning-by-doing framework, individuals supply labor *less* elastically earlier in their life. This is so because work not only generates income in the current period but also increases their future productivity. Simple inverse elasticity reasoning leads learning-by-doing to be a force toward making taxes higher earlier in life.

Learning-or-doing, on the other hand, means that time spent on training works as a substitute for labor. Because training is more valuable when one is young this entails a greater elasticity of labor supply for the young. This is a force pushing toward higher taxes later on one's life.

This is not the whole story, though. One may view e_t as an alternative (to physical capital) of investments. Marginal tax rates which increase with age thus represent a form of tax on the return to this investment.

Asset Accumulation Besides choosing how much leisure to consume and, in the learning or doing model, how to split the remaining time between work and human capital accumulation, individuals trade a risk free asset which holdings we denoted by a_t .

Asset holdings are subject to an exogenous lower bound. More precisely, we follow [Conesa et al. \[2009\]](#) in assuming that agents are not allowed to contract debt at any age, so that the amount of assets carried over from age t to $t + 1$ is such that $a_{t+1} \geq 0$. Because no agent can hold a negative position in assets at any time, we assume without loss that assets take the form of capital, $a_t = k_t$, as in [Aiyagari \[1994\]](#).

The rather extreme assumption about borrowing limits we adopt guarantees that nobody dies in debt. Without this assumption lending would be subject to default risk. Incentives to borrow would vary with age simply due to the time-varying probability of never paying the debt. This could lead to either age dependent borrowing constraints, age-dependent lending rates or both. In any case, this would substantially increase the problem. A consequence of this tight borrowing limit is that lifespan uncertainty, lead to a fraction of the population leaving accidental bequests. We shall discuss what happens to accidental bequests momentarily.

Asset accumulation is, of course, an important aspect of life-cycle choices which we aim at capturing here. As we shall make clear, there is exogenous (as well as endogenous) variation in productivity along the life-cycle. Consumption smoothing thus provides a reason for one to accumulate assets. Another aspect of choices is that individuals may resort to self-insurance to protect themselves against the uncertainty on

labor income. Savings will be, to some extent, motivated by precautionary reasons.

Budget Constraints To write each agent’s flow budget constraint we need to specify the fiscal policy that is being used by the Government. In our case, it is important to distinguish the current fiscal policy, needed to calibrate the model, from the ones we evaluate. The current tax system will be the benchmark for our studies.

The Benchmark Tax System We approximate the benchmark labor income tax with a tax schedule of the form $T(y) = \tau_w \min \{y - \bar{y}; 0\}$, where \bar{y} is an exemption level calibrated as a fraction of the economy’s average income. We assume that consumption is taxed at a rate τ_c and capital income at a rate τ_k .

Beyond that, in the US economy, Social Security introduces other wedges on agent’s choices. Note, in particular, that the wedges in the consumption-leisure introduced Social Security vary with distance from retirement. The current fiscal policy is, in this sense, *not* age-independent.¹¹

To isolate the desirable properties of an optimal tax system driven by the environment from those required to handle existing distortions induced by other aspects of policy, we eliminate social security in our definition of a benchmark. Our candidate optimal tax systems, therefore, require a complete overhaul of the current one. As hinted before, the main reason why we do this is to focus on age-dependence which requires our disentangling the dependence introduced in our exercises from those which are already present due to the characteristics of the social security ‘contract’.

The problem with eliminating Social Security is that the capital stock will experience a large increase due to increased savings for retirement. Recall that all savings must be done through capital accumulation since agents are not allowed to borrow. In contrast, transfers associated with social security create a distinction between the sum of (implicit) individual savings and aggregate savings, thus reducing the economy’s capital stock. To keep the capital-output ratio, K/Y , stays at the level currently found in the data we must change our calibration, in particular, choosing a different value for β .¹²

Optimal Systems We search for the optimal systems within a restricted class. That is our candidate optimal systems are comprised of an age-dependent labor income

¹¹The point here is that benefits are tied to contributions. Yet, the farther away benefits are from contributions, the greater the distortion induced by the opportunity cost of social security ‘investments’.

¹²Other parameters will also have to be changed to satisfy our other targets. See Section 3.

tax of the form

$$T_t(y) = \tau_{w,t} \min \{y - \bar{y}; 0\}.$$

rate $\tau_{w,t}$, an age-independent consumption tax τ_c , and an age-independent capital income tax τ_k .

The flow budget constraint that individuals face in our model economy is, therefore,

$$k_{t+1} = [1 + r(1 - \tau_k)]k_t + (1 - \tau_{w,t})y_t + \tau_{w,t} \min \{\bar{y}; y_t\} + \epsilon - (1 + \tau_c)c_t \forall t, \quad (5)$$

for $t \leq T$.

By assumption $a_1 = 0$. Moreover, given that there is no altruistic bequest motive and death is certain at the age $T + 1$, agents who survive until age T consume all their available resources. That is, $a_{T+1} = 0$, and

$$c_T = \frac{1 + r(1 - \tau_k)]k_{T-1} + \epsilon}{1 + \tau_c}. \quad (6)$$

In both (5) and (6), ϵ is a lump sum transfer related to the involuntary bequests left by those who die before reaching age $T + 1$. Note that ϵ is not age-dependent, i.e., we assume that the lump sum transfer is identical across cohorts. Moreover, since in a steady-state time and age can be treated identically, ϵ need not be indexed.

Recursive Formulation of Households' Problem Let $V_t(\omega_t)$ denote the value function of an individual aged $t < T + 1$, where $\omega_t = (a_t, u, z_t, s_t) \in \Omega$ is the individual state. In addition, considering that agents die for sure at age T and that there is no altruistic link across generations, we have that $V_{T+1}(\omega_{T+1}) = 0$. Thus, the optimization problem of individuals aged t under the exogenous productivity path problem and the leaning-by-doing economies can be recursively represented as follows. Let $\omega' = (a', u, z', s')$, then,

$$V_t(\omega) = \max_{n, a' \geq 0} : [U(c, 1 - n) + \beta \psi_{t+1} E_{z'} V_{t+1}(\omega')], \quad (7)$$

subject to (5), in the case of the exogenous productivity path economy, and to (5) and (3), in the case of the learning by doing economy.

The same problem under the learning-or-doing approach is given by:

$$V_t(\omega) = \max_{n, e, a' \geq 0} : [U(c, 1 - n - e) + \beta \psi_{t+1} E_{z'} V_{t+1}(\omega')] \quad (8)$$

subject to (5) and (4) where $\omega' = (a', u, z', s')$.

It should be stressed that we have imposed non-negativity constraints on asset holdings. We have thus taken an extreme (though plausible) position with regards to capital markets. Relaxing a little the assumption by allowing some exogenous limit is likely to have little effect on our conclusions, at the cost of introducing a whole new set of issues that would have to be dealt with to maintain the internal consistency of the model.

Also important is the fact that we have only used individual state variables in ω . It is apparent that prices do enter the value function. Indeed, in solving the model we will need to find the equilibrium prices by explicitly taking into account how they enter the policy functions associated with (7).

2.3 Technology

Technology is standard. The production side of the economy aggregates and the technology for producing the consumption good is summarized by a Cobb-Dougllass production function with constant returns to scale,

$$Y = BK^\alpha N^{1-\alpha},$$

where K is aggregate capital, N is aggregate efficient units of labor, and B is a scale parameter.

Every period, the standing representative firm solves the static optimization problem

$$\max_{K,N} \{BK^\alpha N^{1-\alpha} - \delta K - wN - rK\},$$

where r is the rental rate of physical capital and w is the rental rate of human capital, i.e. the wage rate. Note that we assume that the rental rate of capital is net of depreciation costs which are born directly by the firm.

The first order conditions for the firm's profit maximization problem are,

$$(1 - \alpha)BK^\alpha N^{-\alpha} = w, \tag{9}$$

and

$$\alpha BK^{\alpha-1} N^{-\alpha} - \delta = r. \tag{10}$$

2.4 Government

As we have already seen, the government levies taxes τ_k on capital income and τ_c , on consumption. It also taxes labor income, according with the potentially age-dependent

function, $T_t(y)$. We have, in this sense, restricted the government's choice of τ_k and τ_c to be age independent but allowed taxes on labor income to vary with age.

The restriction on consumption taxes is natural if we accept that taxes on consumption are anonymous. This is a reasonable assumption for the majority of goods due to negligible transaction costs. As for capital income taxation, one may argue that most savings are not anonymous since they require the existence of institutions that guarantee the enforcement of contracts. Therefore, we think of age independence as a true arbitrary restriction on taxes which we impose to focus on our main question.

Tax revenues are raised to finance an exogenous flow of expenditures, G . As previously discussed, we also assume that the government collects the accidental bequests and transfers to all agents in the economy on a lump-sum basis.

The Government budget constraint is, therefore,

$$G \leq \tau_k r K + \sum_t \tau_{w,t} Y_t + \tau_c C - \epsilon,$$

where C is aggregate consumption, and Y_t is total labor income earned by t years old individuals.

2.5 Discussion of the Environment

Before we move on to the definition of an equilibrium for our economy, it is worth pointing out some of the features in our model that are bound to play a role in policy formulation.

First, notice that our assumption about preferences leads to Frisch elasticities of labor supply that decrease with hours worked. Indeed, let ϵ^f denote the Frisch elasticity of labor supply, then $\epsilon^f = [1 + \rho(\gamma - 1)]^{-1}(1/n - 1)$. Elasticities are, of course, crucial in the determination of optimal taxes. So, understanding how hours vary along the life-cycle will be important in understanding policy prescriptions. ¹³

An alternative assumption used by the literature is separable iso-elastic preferences of the form $u(c, l) = c^\alpha/\alpha + (1 - l)^\gamma/\gamma$. This latter assumption has the advantage of eliminating one channel for the optimality of age-dependent taxes, but the drawback of providing a poorer fit of the data.

Next, our choice of eliminating social security and recalibrating the parameters to guarantee that the stock of capital is unchanged means that we need to make agents more impatient. Eventual welfare gains found from moving to a age-dependent system

¹³The Frisch elasticity is calculated for the model without human capital accumulation. With endogenous human capital the expressions for Frisch elasticity become much more involved. See [Keane \[2011\]](#)

should be interpreted baring this in mind.

2.6 Recursive competitive equilibrium

At each point in time, agents differ from one another with respect to age t and to state $\omega = (a, u, z, s) \in \Omega$. Agents of age t identified by their individual states ω , are distributed according to a probability measure λ_t defined on Ω , as follows. Let $(\Omega, F(\Omega), \lambda_t)$ be a space of probability, where $F(\Omega)$ is the Borel σ -algebra on Ω : for each $\eta \subset F(\Omega)$, $\lambda_t(\eta)$ denotes the fraction of agents aged t that are in η .

Given the age t distribution, λ_t , $Q_t(\omega, \eta)$ induces the age $t + 1$ distribution λ_{t+1} as follows. The function $Q_t(\omega, \eta)$ determines the probability of an agent at age t and state ω to transit to the set η at age $t + 1$. $Q_t(\omega, \eta)$, in turn, depends on the policy functions in (7), and on the exogenous stochastic process for z .

A recursive competitive equilibrium for the economy with human capital accumulation based on learning-by-doing is as follows.¹⁴

Definition 1. Given the policy parameters, a **recursive competitive equilibrium** for the exogenous path and the learning-by-by doing economies are a collection of value functions $\{V_t(\omega)\}$, policy functions for individual asset holdings $d_{a,t}(\omega)$, for consumption $d_{c,t}(\omega)$, for labor supply $d_{n_w,t}(\omega)$, prices $\{w, r\}$, age dependent but time-invariant measures of agents $\lambda_t(\omega)$, transfers ϵ and a tax on consumption τ_c such that:

- (i) $\{d_{a,t}(\omega), d_{n_w,t}(\omega), d_{c,t}(\omega)\}$ solve the dynamic problems in (7);
- (ii) individual and aggregate behaviors are consistent, that is:

$$K = \sum_{t=1}^T \mu_t \int_{\Omega} d_{a,t}(\omega) d\lambda_t$$

$$N = \sum_{t=1}^T \mu_t \int_{\Omega} d_{n_w,t}(\omega) s_t(\omega) \exp(u + z_t) d\lambda_t$$

$$C = \sum_{t=1}^T \mu_t \int_{\Omega} \{d_{c,t}(\omega)\} d\lambda_t;$$

- (iii) $\{w, r\}$ are such that they satisfy the optimum conditions (10) and (9);

¹⁴In the case of the learning-or-doing approach, there is a small change in the definition in which we take into account the policy function for the time spent on human capital accumulation.

(iv) The final good market clears:

$$C + G + \delta K = K^\alpha N^{1-\alpha};$$

(v) given the decision rules, $\lambda_t(\omega)$ follows the law of motion:

$$\lambda_{t+1}(\eta) = \int_{\Omega} Q_t(\omega, \eta) d\lambda_t \quad \forall \eta \subset F(\Omega);$$

(vi) the distribution of accidental bequests is:

$$\epsilon = \sum_{t=1}^T \mu_t \int_{\Omega} (1 - \psi_{t+1}) d_{a,t}(\omega) d\lambda_t$$

(vii) taxes are such that the government's budget constraint,

$$\tau_c C = G - \tau_k r K - \sum_{t=1}^T \mu_t \int_{\Omega} \tau_{w,t} \min\{d_{n_w,t}(\omega) s_t(\omega) \exp(u + z_t) - \bar{y}; 0\} d\lambda_t.$$

is satisfied every period.

For the learning-or-doing economy there is an additional policy function $d_{e,t}(\omega)$ mapping the state $\omega = (a, u, z, s)$ into a human capital investment, e . Note also that item (vii) is redundant if conditions (i)–(vi) hold.

2.7 The Planner's Program

The planner's objective requires some discussion. For any Paretian objective, the planner must maximize a non-decreasing function of agents' expected utilities. There are two relevant questions to be answered. First is how we weight different agents of the same cohort. Second, how we weight the different cohorts.

For the first question, we assume that the Planner chooses policy parameters in order to maximize a Utilitarian social welfare function. That is, the government weights equally all individuals of the same cohort. As for the second, what we do in practice is to follow [Conesa et al. \[2009\]](#) in assuming that the government maximizes the ex-ante lifetime utility of an agent born into the stationary equilibrium implied by the optimal policy. Therefore, we implicitly assume that there is no discounting.

The instruments available for the Government to pursue its objective are: age-dependent labor income tax schedules, $T_t(\cdot)$; age-independent capital income, τ_k , and consumption taxes, τ_c .

Given the computational costs involved we shall restrict our search to a tax system comprised of a fixed, age-independent exemption level, \bar{y} , and a single potentially age-dependent marginal tax rate, $\tau_{w,t}$. Moreover, we assume the age-dependence to be of a linear form by adopting the following parametrization,

$$\tau_{w,t} = \xi_0 + \xi_1 t, \quad (11)$$

where t is age. We restrict marginal tax rates to be non-negative.

To make our findings more easily comparable across settings, instead of choosing \bar{y} , we choose a parameter ϱ that gives the exemption level as a percentage of average income. As we vary both ϱ and ξ_1 the parameter ξ_0 adjusts to hold the government revenue at 18 percent of GDP. These exercise is done in such a way as to maximize a utilitarian social welfare function, which is how we have defined the planner's problem.

Although we do not optimize with respect to τ_k , we explore different values for this parameter to better understand how it interacts with age-dependent taxes, a theme which we have shown to be frequently discussed in the literature.

3 Calibration

To carry out our quantitative analysis, we need first to find values for all the parameters of the model. We accomplish this by calibrating the model for the U.S. economy.

The population age profile $\{\mu_t\}_{t=1}^T$ depends on the population growth rate g_n , the survival probabilities ψ_t and the maximum age T that an agent can live. In this economy, a period corresponds to one year and an agent can live 75 years, so $T = 75$. Additionally, we assumed that an individual is born at age 16, so that the real maximum age is 90 years.

Data on survival probability by age were extracted from [Bell and Miller \[2005\]](#) and are shown in [Figure 1](#). Given the survival probabilities, the population growth rate is chosen so that the age distribution in the model replicates the dependency ratio observed in the data. By setting $g_n = 0.0105$, the model generates a dependency ratio of 17.27%, which is close to the dependency ratio observed in the data for 2000.

In order to calibrate the preference parameters we proceed as follows. First, we choose the discount factor β in such a way that the equilibrium of our benchmark economy implies a capital-output ratio of 3.0, which is the value observed in the data.

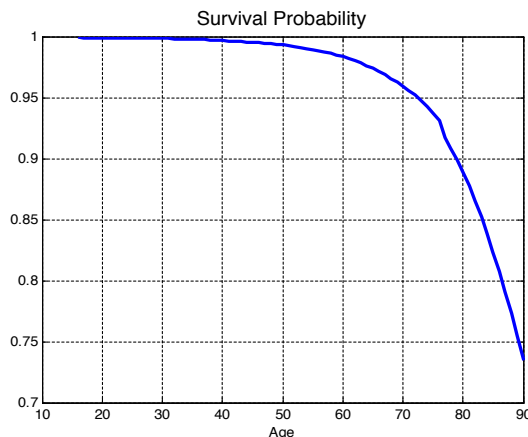


Figure 1: The figure displays the survival probability by age, ψ_t , for the year 2000. Source: [Bell and Miller \[2005\]](#).

Then we fix the parameter γ to 4.0 and choose the share of consumption in the utility function so that working-aged individuals work on average one third of their time endowment. The values of β and ρ obtained for each type of model (ie., Exogenous, Learning-by-doing and Learning-or-doing) are presented in Table 1.

For Cobb-Douglas preferences, the coefficient of relative risk aversion is given by $1 - \rho + \rho\gamma$, while the Frisch elasticity for leisure is given by $-\frac{1-\rho+\rho\gamma}{\gamma}$. Thus, the values reported in Table 1 entail a value of 2.98 for the coefficient of relative risk aversion parameter and of 0.74 for the Frisch Elasticity for leisure, which is consistent with the empirical evidence in [Auerbach and Kotlikoff \[1987\]](#), [Rust and Phelan \[1997\]](#) and [Domeij and Flodén \[2006\]](#).

The parameters that characterized the stochastic component of individuals productivity are $(\sigma_u^2, \varphi_z, \sigma_\epsilon^2)$. Several authors have estimated similar stochastic process for labor productivity. Controlling for the presence of measurement errors and/or effects of some observable characteristics such as education and age, the literature provides a range of $[0.88, 0.96]$ for φ_z and of $[0.10, 0.25]$ for σ_ϵ^2 . In this article, we rely on the estimates of [Kaplan \[2012\]](#), setting $\varphi_z = 0.94$ and $\sigma_\epsilon^2 = 0.016$. Then, the parameter σ_u^2 was chosen in order for the Gini index for labor income in the model to match its counterpart in the data, which is nearly 0.43. The value obtained for σ_u^2 is in line with the estimates in [Kaplan \[2012\]](#) who provides a point estimate of 0.056 for this parameter.

The values of technological parameters (α, δ) are also summarized in [1](#). We chose a value for α based on U.S. time series data from the National Income and Product Accounts (NIPA). The depreciation rate, in turn, is obtained by $\delta = \frac{I/Y}{K/Y} - g$. We set the

Table 1: Parameter Values - Baseline Calibration

Parameter	Value	Source/Target
$\beta_{EXO}, \beta_{LBD}, \beta_{LOD}$	1.002, 0.998, 1.005	$K/Y = 3$
γ	4.00	Micro evidence
$\rho_{EXO}, \rho_{LBD}, \rho_{LOD}$	0.66, 0.71, 0.65	Average $l = 0.32$
σ_u^2	0.07	Gini index of 0.43
φ_z	0.94	Kaplan [2012]
σ_ϵ^2	0.016	Kaplan [2012]
δ	0.05	see text
α	0.36	NIPA
$\phi_{s,LBD}, \phi_n$	0.40, 0.35	see text
$\phi_{s,LOD}, \phi_e$	0.60, 0.13	see text
δ_h	0.05	Heckman et al. [2002]
π	0.15	see text
B	0.90	$w = 1$
τ_c	0.05	Fuster et al. [2007]
τ_k	0.30	Fuster et al. [2007]

investment-product ratio I/Y equal to 0.25 and the capital-product ratio K/Y equal to 3.0. The economic growth rate, g , is constant and consistent with the average growth rate of GDP over the second half of the last century. Based on data from Penn-World Table, we set g equal to 2.7%, which yields a depreciation rate of 5.4%.

The values for the actual age-efficiency profile are constructed similarly to Huggett [1996] and McGrattan and Rogerson [1998]. We use annual earnings and annual hours worked for the age groups 15-24, 25-34, ..., 75-84 from IPUMS (U.S. Department of Commerce, Bureau of the Census 2005). First, we construct hourly wages by dividing annual earnings by annual hours for each age group. Afterwards, we use a second order polynomial to interpolate the points to obtain the age-efficiency profile by exact age.

In order to calibrate the parameters of the skill accumulation functions, we first set $\delta_h = 0.05$, which is consistent with the evidence presented in Heckman et al. [2002] who suggest a range of [0.0016, 0.089] for this parameter. In the LBD case, we follow Chang et al. [2002] who use PSID data set to estimate this equation. In particular, we use their posterior point estimates of $\phi_{s,LBD} = 0.40$ and $\phi_n = 0.35$. In the case of LOD parameters, Heckman et al. [1998] show that the ratio of time spent on training to market hours starts at about 40% at ages 20 – 22 and then declines to near zero by age 45. In addition, the ratio between the average time spent on training over the life-

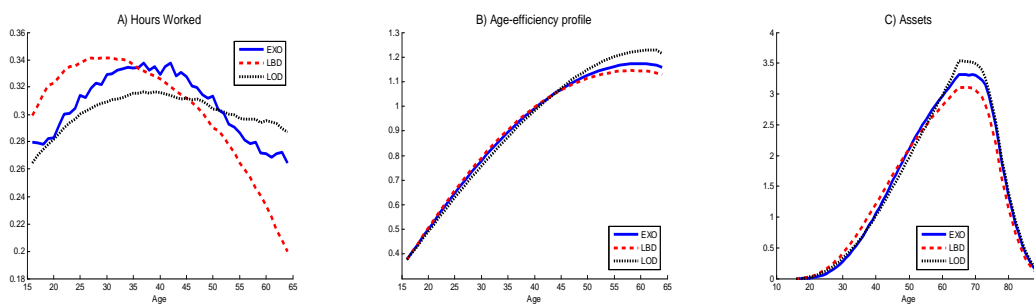


Figure 2: **Life Cycle Profiles - Benchmark Economy.** The first figure in the top displays how the cross-sectional average hours of work varies with age for each of the three assumptions about the age-efficiency profile: exogenous (EXO); learning or doing (LOD) and learning by doing (LBD). The figure in the right displays the age efficiency profile for the three models and the figure in the bottom displays the average cross-sectional holdings of assets by age.

cycle and market hours is about 6%. Thus, by choosing $\phi_{s,LBD} = 0.60$ and $\phi_e = 0.10$, our model is able to reproduce these calibration targets in the benchmark economy. We then calibrate the scale parameter π in order to match the average growth rate of the age-efficiency profile observed in the data, which is nearly 2.2%.

Finally, we specify the others parameters related to government activity. First, we set government consumption, G , to 18% of the output of the economy under the baseline calibration. Following the literature, we assume a consumption tax of 5% and a capital income tax rate of 30%.¹⁵ We assume that anyone with earnings of up to thirty percent of mean income is exempt from labor income taxes, i.e., $\varrho = 0.3$. Marginal tax rates are age-independent, $\xi_1 = 0$, and chosen to raise enough revenue to finance government consumption. The value we find for ξ_0 vary slightly across models, but stands between 23% and 23.5%.

Figure 2 displays the age efficiency profile and corresponding average hours worked and asset accumulation patterns for the benchmark economy.

Averages may, of course, hide a rich diversity in life-cycle patterns. We split the individuals in our economy in three different ability groups. We group the agents in the top 16 percentiles of the distribution of innate ability, u , and label them the high ability group. The agents on the bottom 16 percentiles are labelled low ability. In Figure 3, we plot the same variables considered in Figure 2 for each of these groups along with the overall average to get a sense of how heterogeneity plays in our model.

Life-cycle patterns are qualitatively similar for all groups and all models. High ability individuals do, however, work more hours and accumulate more assets than

¹⁵See, for example, [Fuster et al. \[2007\]](#)

lower ability individuals for all different specifications of human capital dynamics. The age-efficiency profile exhibits some differences across the models. For the exogenous model, we assume that they do not vary across abilities. For the other two models the high ability individuals display a more pronounced increase in productivity, even though this is barely noticeable for the Learning-or-doing model.

4 Results

Our principal results are presented in Table 2. For each model of human capital accumulation, we compare the benchmark and optimal tax numbers for: GDP, capital-output ratio, average hours, wages, real interest rates, policy parameters and welfare. We consider three different levels for the marginal tax rate on capital income, $\tau_k = 30\%$, which is the benchmark value, $\tau_k = 15\%$, and $\tau_k = 0\%$.

Starting with the optimal policy, note that the marginal tax rate faced by a t years old individual is $\tau_{w,t} = \max\{\xi_0 + \xi_1 t; 0\}$. We also consider progressivity by allowing for an exemption level \bar{y} which we calculate using ϱ , where ϱ times the mean income of the economy is his or her exemption level.

For the first set of results, in the upper part of the table, we hold the marginal tax rate on capital fixed at 30 percent. The first thing to note is that ξ_1 is positive: marginal tax rates increase with age. Our results therefore replicate the findings in [Conesa et al. \[2009\]](#), thus suggesting that the insurance aspects dominate. In fact, marginal tax rates increase substantially from the early stages in life, when it is optimally set to zero to the periods close to retirement. The LOD model is the one for which age-dependency is strongest, $\xi_1 = 1.96$ and the LBD model is the one for which it is weakest, $\xi_1 = 1.41$. The model with exogenous age-efficiency profile is somewhere between the two with $\xi_1 = 1.85$. Since we find a negative value for ξ_0 for all models, there are periods of an individual's working life for which he or she faces a zero marginal tax rate on labor income.

To get a full grasp of our findings it is therefore useful to report the age at which individuals start facing positive marginal tax rates and the marginal tax rate they face right before retirement. Individuals start paying taxes earlier for the LBD model, at age 32, and later for the LOD, at age 38. For the model with exogenous age-efficiency profile marginal tax rates become positive when agents reach the age of 35. The year before retiring individuals are facing marginal tax rates of 46% in the LBD model, and 52% for both the LOD and exogenous models.

As for progressivity, for *none* of the models it is optimal to set $\varrho > 0$! Progressivity is, in this sense, unnecessary. Of course, this is but one crude way of introducing non-

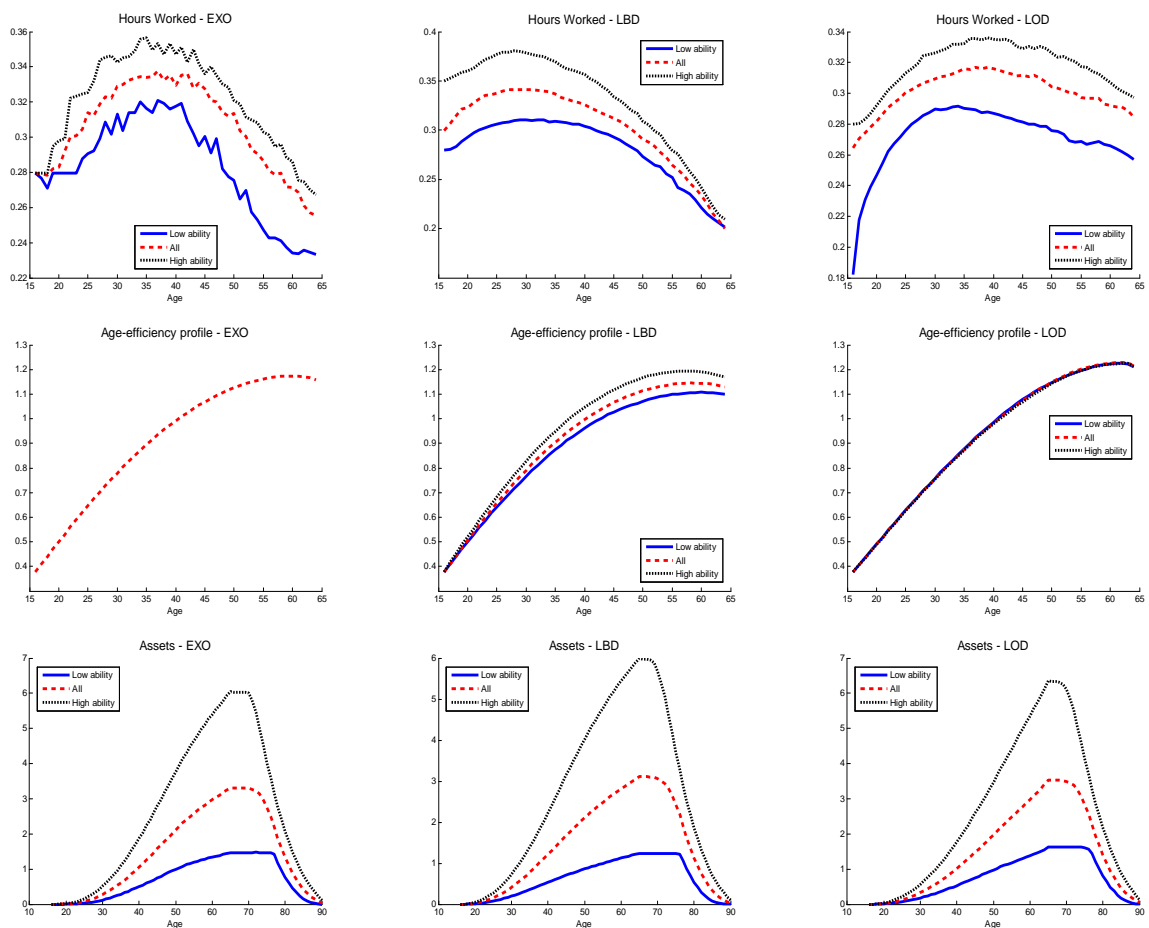


Figure 3: Life Cycle Profiles by Types. The figure displays the average hours (top row), the age efficiency profile (middle row) and asset holdings (bottom row) for the exogenous (left column) the learning-by-doing (middle column) and the learning-or-doing (right column) models at the benchmark tax system.

Table 2: **Optimal with age-dependent taxation.** The table displays the values for the relevant variables – GDP (Y), capital-output ratio (K/Y), average hours (Avg hours), wages (w), real interest rates(r), policy parameters (ξ_0, ξ_1, ϱ) and welfare (Welfare and CEV)– for each of the three models of human capital formation, Exogenous (EXO), Learning-by-doing (LBD) and Learning-or-doing (LOD) for the benchmark and the optimum.

$\tau_k = 0.30$						
Variable	EXO		LBD		LOD	
	Benchmark	Optimal	Benchmark	Optimal	Benchmark	Optimal
Y	0.50	0.59	0.49	0.58	0.50	0.59
K/Y	3.00	3.68	3.01	3.51	3.00	3.65
Avg hours	0.31	0.33	0.30	0.33	0.30	0.32
w	1.01	1.13	1.01	1.10	1.01	1.12
r	5.39%	3.15%	5.35%	3.65%	5.40%	3.25%
ξ_0	23.42%	-68.67%	23.37%	-45.33%	23.28%	-75.39%
ξ_1	0%	1.85%	0%	1.41%	0%	1.96%
ϱ	30%	0%	30%	0%	30%	0%
Welfare	-211.86	-198.90	-157.61	-150.31	-272.12	-258.15
CEV	-	6.37%	-	5.99%	-	5.23%
$\tau_k = 0.15$						
Variable	EXO		LBD		LOD	
	Benchmark	Optimal	Benchmark	Optimal	Benchmark	Optimal
Y	0.50	0.59	0.49	0.57	0.50	0.58
K/Y	3.00	3.82	3.01	3.63	3.00	3.72
Avg hours	0.31	0.33	0.30	0.32	0.30	0.32
w	1.01	1.15	1.01	1.12	1.01	1.14
r	5.39%	2.81%	5.35%	3.30%	5.40%	3.07%
ξ_0	23.42%	-62.31%	23.37%	-34.05%	23.28%	-49.34%
ξ_1	0%	1.83%	0%	1.28%	0%	1.53%
ϱ	30%	0%	30%	0%	30%	0%
Welfare	-211.86	-198.37	-157.61	-150.88	-272.12	-257.92
CEV	-	6.52%	-	5.39%	-	5.02%
$\tau_k = 0.00$						
Variable	EXO		LBD		LOD	
	Benchmark	Optimal	Benchmark	Optimal	Benchmark	Optimal
Y	0.50	0.59	0.49	0.56	0.50	0.58
K/Y	3.00	3.91	3.01	3.72	3.00	3.81
Avg hours	0.31	0.32	0.30	0.31	0.30	0.31
w	1.01	1.17	1.01	1.14	1.01	1.15
r	5.39%	2.59%	5.35%	3.07%	5.40%	2.83%
ξ_0	23.42%	-52.34%	23.37%	-25.41%	23.28%	-41.64%
ξ_1	0%	1.71%	0%	1.18%	0%	1.46%
ϱ	30%	0%	30%	0%	30%	0%
Welfare	-211.86	-198.89	-157.61	-152.21	-272.12	-259.41
CEV	-	6.16%	-	4.22%	-	4.03%

linearity. More general non-linear tax systems are likely to improve welfare. What our results suggest is that the bulk of the gains are captured by an age-dependent proportional labor income tax. This finding is in line with those in [Farhi and Werning \[2013\]](#) and [Weinzierl \[2012\]](#).

The main consequence of introducing age dependence in labor income taxes is to bias labor supply toward younger ages. For both the exogenous and the LOD model this means a very large increase in hours for the first 15 to 20 years after joining the labor force and a steady decline until retirement. With LOD, hours start at a very high level, around 7 hours higher than at the benchmark. They slightly increase until its peak around 33 years then slowly decline until retirement. With LBD agents are on average working fewer hours than they were at the benchmark at ages 50. The same is true with exogenous human capital, at 45, and at 55 for the LOD – see [Figure 6](#).

As for welfare, the introduction of age-dependent taxes is necessarily welfare enhancing, since replicating the benchmark is a feasible policy. The magnitude of welfare gains we obtain is, however, larger than most results found elsewhere. The bulk of this gain is due to an indirect effect: the substantial increase in the economy’s capital stock. It is important to emphasize that this is true despite the fact that we have kept the same marginal tax rate on capital income that prevailed for the benchmark economy, $\tau_k = 30\%$. The main reason for this increase in K/Y is apparent from [Figure 6](#). Age-dependent taxes induce individuals to change their pattern of labor supply and concentrate most of their working hours at early ages. They then accumulate more and faster to support a larger span of years working fewer hours or not working at all. This finding highlights the possible understatement of the value of age-dependent taxes for most of the works that take the capital stock as given.

Next, we reduce the tax on capital income to $\tau_k = 15\%$. The dependence of labor income taxes on age is reduced in the sense that both ξ_1 decreases and ξ_0 increases. As a consequence, the earliest age at which individuals face positive marginal tax rates is reduced. Moving from $\tau = 30\%$ to $\tau = 15\%$ is welfare enhancing for the LOD model but not for the LBD model.

The pattern of decreased dependence on age continues as we move from $\tau = 15\%$ to $\tau = 0\%$. In this, case, however, welfare is lower for all models. One interesting pattern that is apparent from [table 2](#) is the higher sensitivity of endogenous human capital models to changes in the tax rate on capital. Indeed, the age for which individuals start being taxed decreases by more than 10 years as we move from $\tau = 30\%$ to $\tau = 0\%$ for the models with endogenous human capital, while the decrease is of ‘only’ 5 years in the case of the model with exogenous age-efficiency profile. As a consequence the ordering of age dependency initially observed is lost. It is now the tax system associated with

the exogenous age-efficiency profile model which displays the strongest dependence on age.

It is also important to note that, although the increase in capital stock is potentially important to explain the welfare gains, the optimal welfare is not always increasing in K/Y . Indeed, although K/Y always increases as we lower τ_k , welfare is lower for $\tau_k = 0\%$ in all three models.

With regards to the differences between the outcomes for the three different assumptions about human capital formation, we first note that optimal policies are similar at $\tau_k = 30\%$. As we reduce τ_k , however, policies respond more intensely when human capital is endogenous. This is related to the findings in [Jacobs and Bovenberg \[2010\]](#) which emphasizes the role played by capital income taxation when human capital is endogenous.

Optimal age-independent taxes As previously discussed, because we start from an arbitrary (real world) tax system and move to an optimal one for which the benchmark is a possibility, we must obtain welfare gains by moving from the benchmark to the optimal system. In a sense, it is more important for our purposes to assess the welfare gains from moving to an optimal age-dependent tax system when compared to an optimal age-independent one. In [Table 3](#) we present our findings for this case. The most important thing one observes is that welfare gains are substantially smaller if we constrain the planner to only using age-independent taxes. Let us start, however, with the description of how the optimal tax system differs from the benchmark one.

Let us start with the case $\tau_k = 30\%$. The first important thing to note is that ρ is no longer zero for all models. In fact, for the LOD model, it is optimal to set $\rho = 15\%$. For the exogenous human capital model we find $\rho = 5\%$ and only for the LBD model we find it to be optimal to set $\rho = 0\%$. Note that for all three possibilities for age-efficiency profile, the optimal tax system is less progressive than in the benchmark: we have both a lower ξ_0 and a lower ρ . The optimal marginal income tax rate is, for the LBD model, only 14.01%, compared with 23.37% in the benchmark. The numbers for the LOD model are 18.35% and 23.28%, and for the exogenous human capital model, 15.95% and 23.42%. For all three models $\rho = 30\%$ in the benchmark.

These findings are suggestive that, for the level of progressivity we find to be optimal, age-dependent taxes proves to be a more efficient instrument. Redistribution across individuals is desirable for two different reasons. First, is the insurance motive that is created by the incompleteness of markets. Second is the redistribution that is induced by the planner's utilitarian objective. Progressive taxes are useful for addressing both issues, yet they may lead to large dead-weight losses. As for linear

Table 3: **Optimal without age-dependent taxation.** The table displays the values for the relevant variables – GDP (Y), capital-output ratio (K/Y), average hours (Avg hours), wages (w), real interest rates (r), policy parameters (ξ_0, ξ_1, ϱ) and welfare (Welfare and CEV)– for each of the three models of human capital formation, Exogenous (EXO), Learning-by-doing (LBD) and Learning-or-doing (LOD) for the benchmark and the optimum.

$\tau_k = 0.30$						
Variable	EXO		LBD		LOD	
	Benchmark	Optimal	Benchmark	Optimal	Benchmark	Optimal
Y	0.50	0.52	0.49	0.54	0.50	0.52
K/Y	3.00	3.04	3.01	3.03	3.00	3.03
Avg hours	0.31	0.32	0.30	0.33	0.30	0.32
w	1.01	1.01	1.01	1.01	1.01	1.01
r	5.39%	5.24%	5.35%	5.27%	5.40%	5.28%
ξ_0	23.42%	15.95%	23.37%	14.01%	23.28%	18.35%
ξ_1	0%	0%	0%	0%	0%	0%
ϱ	30%	5.0%	30%	0%	30%	15%
Welfare	-211.86	-198.90	-157.61	-155.24	-272.12	-270.93
CEV	-	1.03%	-	1.87%	-	0.57%
$\tau_k = 0.15$						
Variable	EXO		LBD		LOD	
	Benchmark	Optimal	Benchmark	Optimal	Benchmark	Optimal
Y	0.50	0.51	0.49	0.53	0.50	0.51
K/Y	3.00	3.12	3.01	3.15	3.00	3.11
Avg hours	0.31	0.31	0.30	0.32	0.30	0.30
w	1.01	1.03	1.01	1.03	1.01	1.03
r	5.39%	4.95%	5.35%	4.83%	5.40%	4.95%
ξ_0	23.42%	20.51%	23.37%	18.43%	23.28%	22.18%
ξ_1	0%	0%	0%	0%	0%	0%
ϱ	30%	5.0%	30%	0%	30%	10%
Welfare	-211.86	-210.95	-157.61	-155.89	-272.12	-274.18
CEV	-	0.47%	-	1.33%	-	-0.70%
$\tau_k = 0.00$						
Variable	EXO		LBD		LOD	
	Benchmark	Optimal	Benchmark	Optimal	Benchmark	Optimal
Y	0.50	0.51	0.49	0.52	0.50	0.49
K/Y	3.00	3.18	3.01	3.24	3.00	3.19
Avg hours	0.31	0.31	0.30	0.31	0.30	0.29
w	1.01	1.04	1.01	1.05	1.01	1.04
r	5.39%	4.70%	5.35%	4.48%	5.40%	4.67%
ξ_0	23.42%	23.45%	23.37%	22.60%	23.28%	27.17%
ξ_1	0%	0%	0%	0%	0%	0%
ϱ	30%	0%	30%	0%	30%	10%
Welfare	-211.86	-213.77	-157.61	-157.61	-272.12	-278.52
CEV	-	-0.77%	-	-0.04%	-	-2.43%

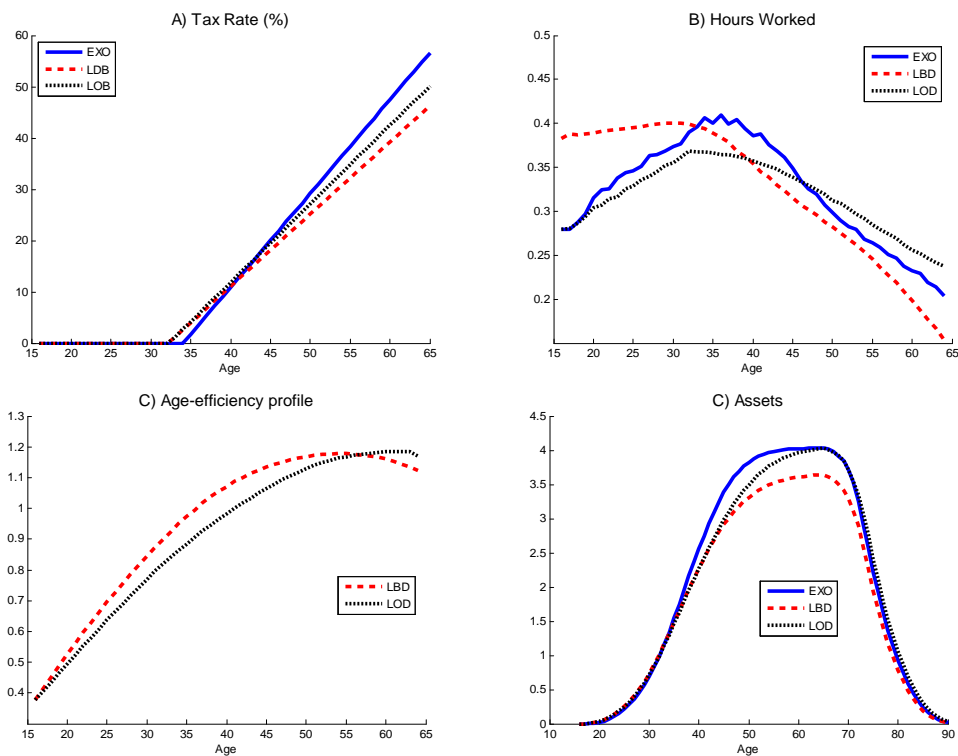


Figure 4: Life Cycle Profiles - Optimal Taxation. The first figure in the top displays how the marginal tax rate on income $\tau_{w,t}$ varies with age for each one of the three assumptions about the age-efficiency profile: exogenous (EXO); learning or doing (LOD) and learning by doing (LBD). The figure in the right displays the average cross-sectional number of hours for each of the three models for the age efficiency profile. The age efficiency profile for the two models with endogenous accumulation of human capital is displayed in the bottom left figure. The figure in the bottom right displays the average cross-sectional holdings of assets by age.

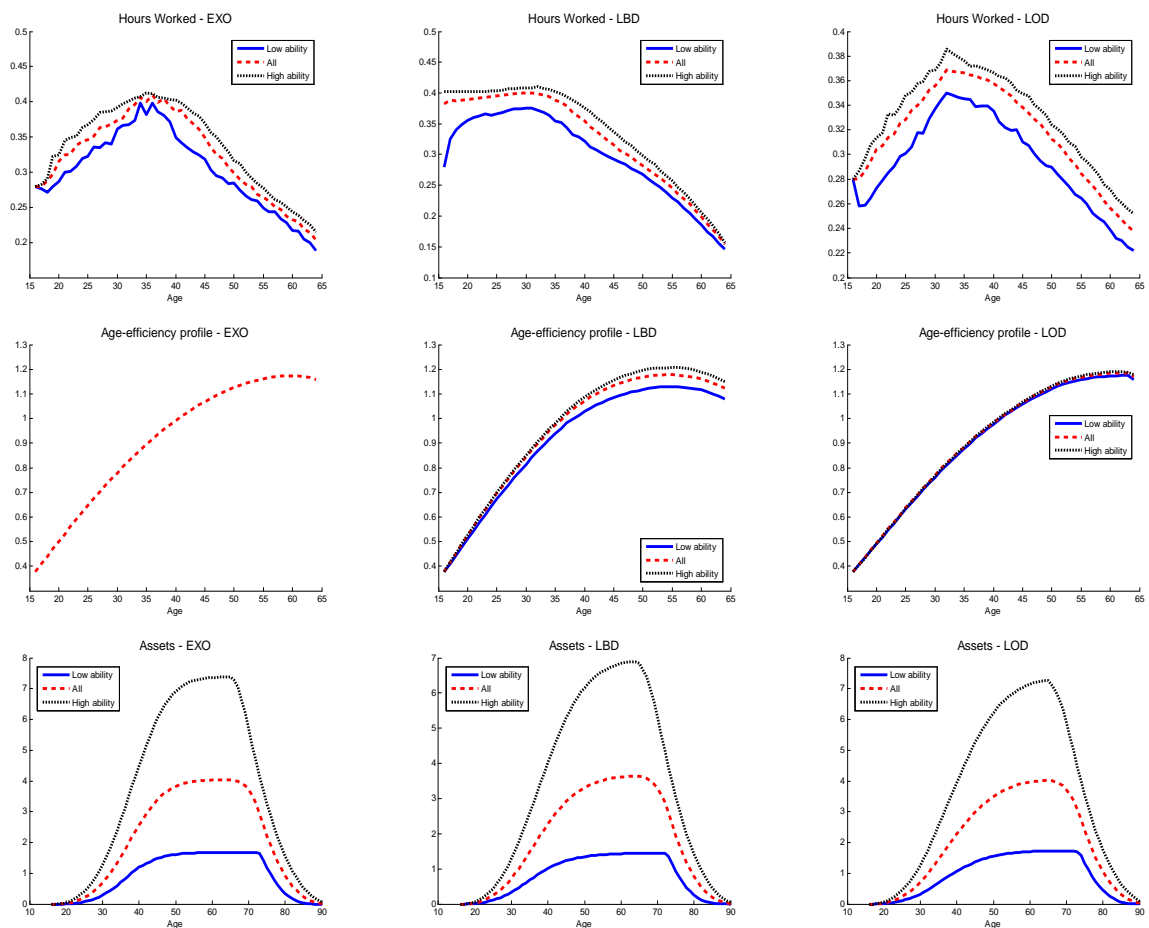


Figure 5: **Life Cycle Profiles by Types.** The figure displays the average hours (top row), the age efficiency profile (middle row) and asset holdings (bottom row) for the exogenous (left column) the learning-by-doing (middle column) and the learning-or-doing (right column) models for agents at the bottom (below the 16-th percentile) and top (above the 83-rd percentile) of the distribution of u at the benchmark tax system.

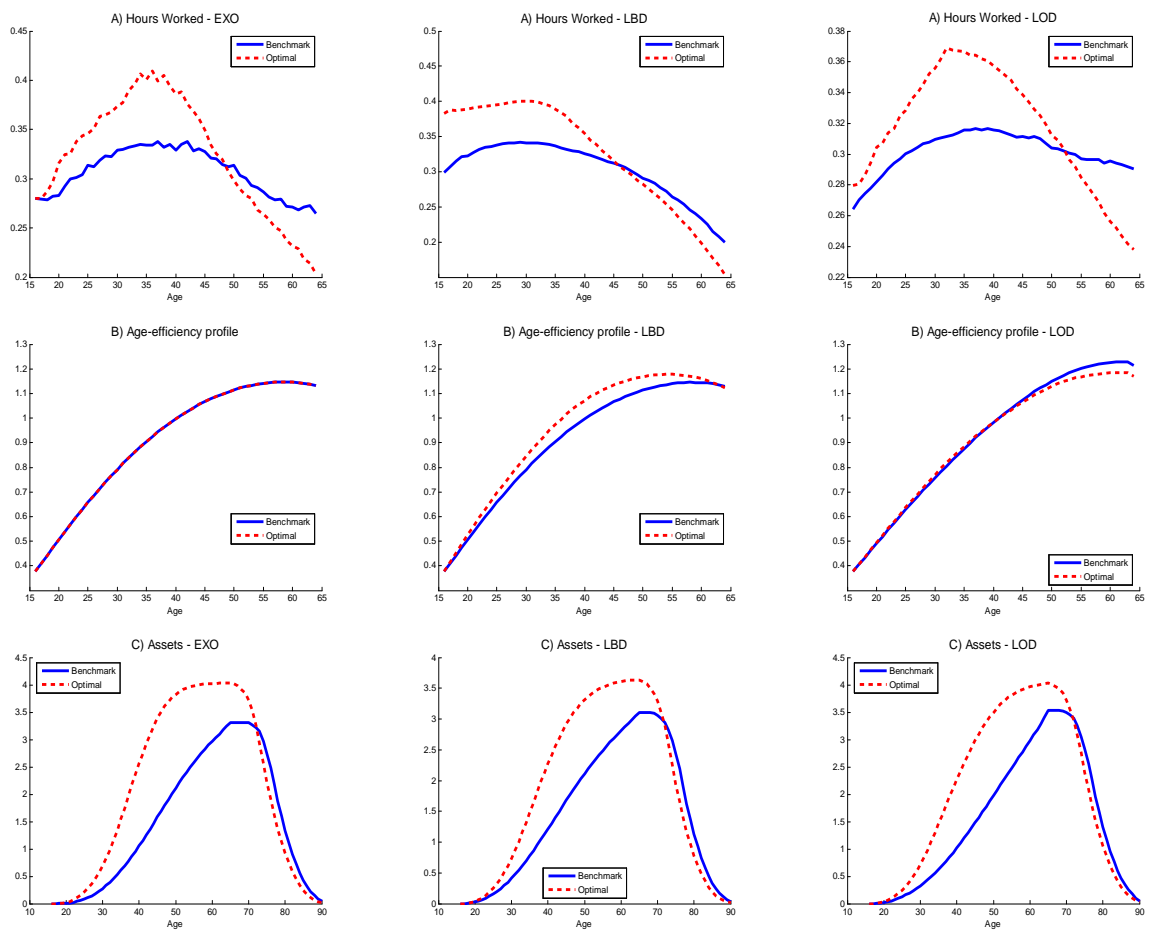


Figure 6: Life Cycle Profiles: Optimal vs Benchmark. The figure displays the average hours (top row), the age efficiency profile (middle row) and asset holdings (bottom row) for the exogenous (left column) the learning-by-doing (middle column) and the learning-or-doing (right column) models at the benchmark and the optimal tax systems.

age-dependent taxes, while they need not be the best instrument for dealing with redistribution for utilitarian reasons, their relative efficiency for dealing with the insurance aspects seems to make the benefits from progressivity unnecessary. Interestingly, as we reduce τ_k it is efficient to decrease the progressivity of the tax system as well. In fact, for $\tau_k = 0$, it is only optimal to set $\varrho > 0$ in the LOD model. Optimal tax prescriptions for the LBD and the exogenous model are similar in this case.

Finally note that as we reduce τ_k from its benchmark level of $\tau_k = 30\%$ welfare gains from tax reform reduce, and become negative for the LOD model at $\tau_k = 15\%$ and for the other two models at $\tau_k = 0\%$.

4.1 The Consequences of Ignoring Human Capital Endogeneity

The next exercise consists in evaluating the effect of disregarding the endogeneity of human capital in the derivation of optimal taxes. This exercise tries to measure the bias that optimal prescriptions derived in the literature may display.

Our procedure is to use ξ_1 derived from the exogenous human capital model and let ξ_0 adjust to keep the government budget balanced.

Figure 7 displays the effects of using a tax profile derived under the assumption of exogeneity when human capital is, in fact, endogenous. As one can easily see, although the consequences of ignoring human capital accumulation do not seem to be very important in the case of the LOD model, the deviations from the optimum can be substantial for the LBD model. Ignoring endogenous human capital accumulation in this case leads to excessively high taxes for individuals above 40 years old and the opposite for individuals younger than 40. Mirroring this policy sub-optimality is the behavior of hours worked, which is excessive for those under 40 and below the optimum for those above 40 years of age. As a consequence, asset accumulation (and decumulation) starts at an earlier age than it would be optimal.

4.2 Partial Equilibrium

What we have shown previously is that the capital stock increases substantially — capital-output ratios increase by more than 20% as for both the model with exogenous age-efficiency profile and the LOD model as replace the benchmark by an optimal age-dependent tax system. This change in the capital stock leads wages to increase by more than 10 % for all specifications of human capital formation at the same time that the rental rate of capital decreases by around two percentage points.

These large movements in price factors are potentially important for explaining the large welfare changes we find. Just how important they are is what we investigate

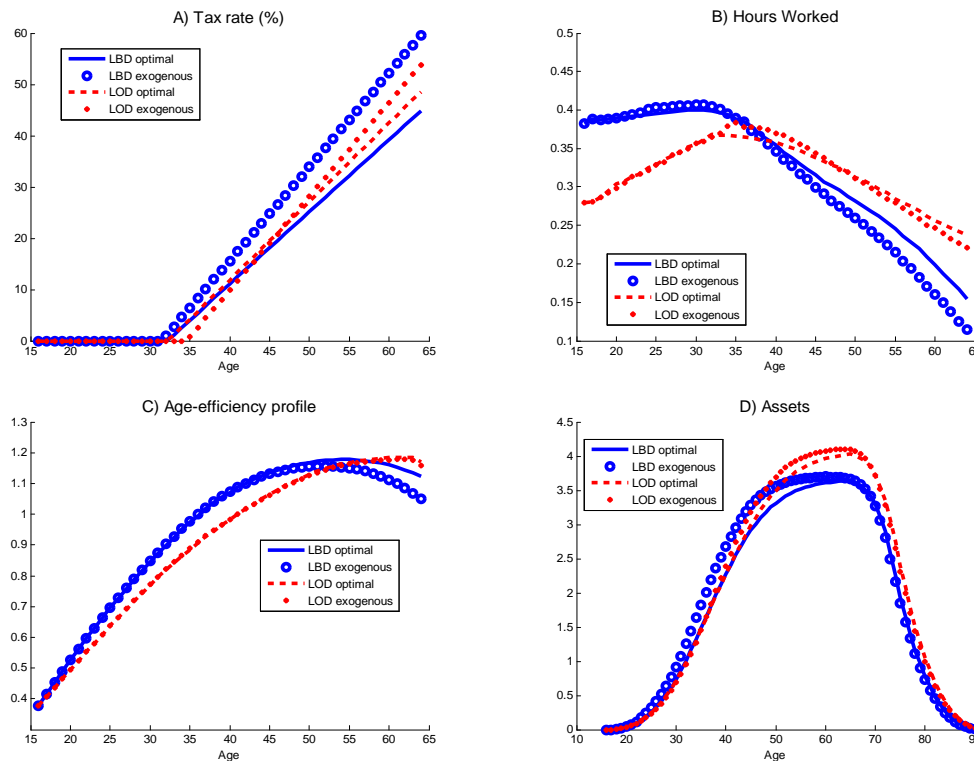


Figure 7: Life Cycle Profiles - Sub-optimal Taxation. The first figure in the top displays how the marginal tax rate on income $\tau_{w,t}$ varies with age for the learning or doing (LOD) and learning by doing (LBD) models using the parameters ξ_1 and ξ_2 in equation (11) calculated for the exogenous model but allowing the parameter ξ_0 to vary to guarantee budget balance. The figure in the right displays the average cross-sectional number of hours for each of the two models for the age efficiency profile. The age efficiency profile for the two models with endogenous accumulation of human capital is displayed in the bottom left figure. The figure in the bottom right displays the average cross-sectional holdings of assets by age. In all figures the life-cycle behavior of the relevant variables for the optimal taxes are also displayed.

next.

Our procedure consists in holding the 'producer prices' — i.e., the rental rates of capital and efficiency hours of work — fixed as we vary the relevant wedges between those producer prices and the consumer prices — net wages and net return on savings.

The first thing to note — see table 5 — is that the growth in capital stock is even more impressive in this case. For the case with exogenous human capital profile, for instance, the capital-output ratio increases from 3 to 4. When we allow prices to adjust in our main exercises, K/Y reaches 'only' 3.68. For the LBD and the LOD models these are 3.64 and 3.84, respectively, without price adjustment and 3.51 and 3.65,

Table 4: **The consequences of disregarding human capital accumulation.** For the LOD and LBD models, the table displays the optimal value for each variable – GDP (Y), capital-output ratio (K/Y), average hours (Avg hours), wages (w), real interest rates(r), policy parameters (ξ_0, ξ_1, ϱ) and welfare (Welfare and CEV)– and the value calculated when ξ_1 is optimal for the exogenous model (Exogenous), with ξ_0 chosen to guarantee budget balance.

Variable	LBD		LOD	
	Optimal	Exogenous	Optimal	Exogenous
Y	0.58	0.58	0.58	0.59
K/Y	3.51	3.79	3.72	3.79
Avg hours	0.33	0.32	0.32	0.32
w	1.10	1.15	1.14	1.15
r	3.65%	2.90%	3.07%	2.90%
ξ_0	-45.33%	-57.59%	-49.34%	-63.31%
ξ_1	1.41%	1.83%	1.53%	1.83%
ϱ	0%	0%	0%	0%
Welfare	-150.31	-151.57	-257.92	-258.27
CEV	-	-1.06%	-	-0.20%

respectively, with price adjustments.

Our first reaction is to think that this is hardly surprising. Since prices are fixed, it is the demand for a savings vehicle which determines alone the economy's capital stock. As we have seen, the move toward age-dependent taxes leads to a large increase in asset accumulation, which we might think would necessarily be made stronger without the price adjustment. Note, however that the optimal tax system will also be different in this case. Because it is the tax reform that leads labor supply to be concentrated in early periods of the life-cycle which induces greater accumulation in the first place, it is paramount to assess how the absence of price adjustments alters the optimal tax system.

Table 5 informs us that without price adjustments the efficient tax system will exhibit less variation with age. For instance, while the age individuals start paying taxes for the model with exogenous human capital formation is 37 when prices adjusts it is only 35 if prices are fixed. For the LOD model it drops from 38 to 32, and for the LBD model from 32 to only 22. The flip side of this pattern is that marginal tax rates which reach as much as 46% near retirement for the LBD model with price adjustments, without them reaches only 34%. The numbers for the LOD are 52% and 43% and, only for the exogenous human capital model we find little variation with a slight increase from 51% to 52% as we shut down the general equilibrium adjustments.

Table 5: **Optimal with age-dependent taxation: Partial Equilibrium.** The table displays the values for the relevant variables – GDP (Y), capital-output ratio (K/Y), average hours (Avg hours), wages (w), real interest rates (r), policy parameters (ξ_0, ξ_1, ϱ) and welfare (Welfare and CEV)– for each of the three models of human capital formation, Exogenous (EXO), Learning-by-doing (LBD) and Learning-or-doing (LOD) for the benchmark and the optimum.

$\tau_k = 0.30$						
Variable	EXO		LBD		LOD	
	Benchmark	Optimal	Benchmark	Optimal	Benchmark	Optimal
Y	0.50	0.55	0.49	0.54	0.50	0.54
K/Y	3.00	4.00	3.01	3.64	3.00	3.84
Avg hours	0.31	0.30	0.30	0.31	0.30	0.30
w	1.01	1.01	1.01	1.01	1.01	1.01
r	5.39%	5.39%	5.35%	5.35%	5.40%	5.40%
ξ_0	23.42%	-62.49%	23.37%	-19.04%	23.28%	-43.26%
ξ_1	0%	1.77%	0%	0.83%	0%	1.33%
ϱ	30%	0%	30%	0%	30%	0%
Welfare	-211.86	-197.66	-157.61	-151.46	-272.12	-260.39
CEV	-	6.73%	-	4.95%	-	3.24%
$\tau_k = 0.15$						
Variable	EXO		LBD		LOD	
	Benchmark	Optimal	Benchmark	Optimal	Benchmark	Optimal
Y	0.50	0.52	0.49	0.52	0.50	0.51
K/Y	3.00	3.83	3.01	3.52	3.00	3.72
Avg hours	0.31	0.29	0.30	0.30	0.30	0.29
w	1.01	1.01	1.01	1.01	1.01	1.01
r	5.39%	5.39%	5.35%	5.35%	5.40%	5.40%
ξ_0	23.42%	-15.34%	23.37%	7.00%	23.28%	-6.84%
ξ_1	0%	0.9%	0%	0.33%	0%	0.68%
ϱ	30%	0%	30%	0%	30%	0%
Welfare	-211.86	-206.27	-157.61	-156.24	-272.12	-272.56
CEV	-	2.34%	-	1.28%	-	-0.93%
$\tau_k = 0.00$						
Variable	EXO		LBD		LOD	
	Benchmark	Optimal	Benchmark	Optimal	Benchmark	Optimal
Y	0.50	0.49	0.49	0.49	0.50	0.49
K/Y	3.00	3.64	3.01	3.44	3.00	3.42
Avg hours	0.31	0.28	0.30	0.29	0.30	0.28
w	1.01	1.01	1.01	1.01	1.01	1.01
r	5.39%	5.39%	5.35%	5.35%	5.40%	5.40%
ξ_0	23.42%	7.05%	23.37%	26.22%	23.28%	21.02%
ξ_1	0%	0.5%	0%	0%	0%	0.14%
ϱ	30%	0%	30%	0%	30%	0%
Welfare	-211.86	-217.43	-157.61	-163.21	-272.12	-288.29
CEV	-	-2.60%	-	-3.92%	-	-5.69%

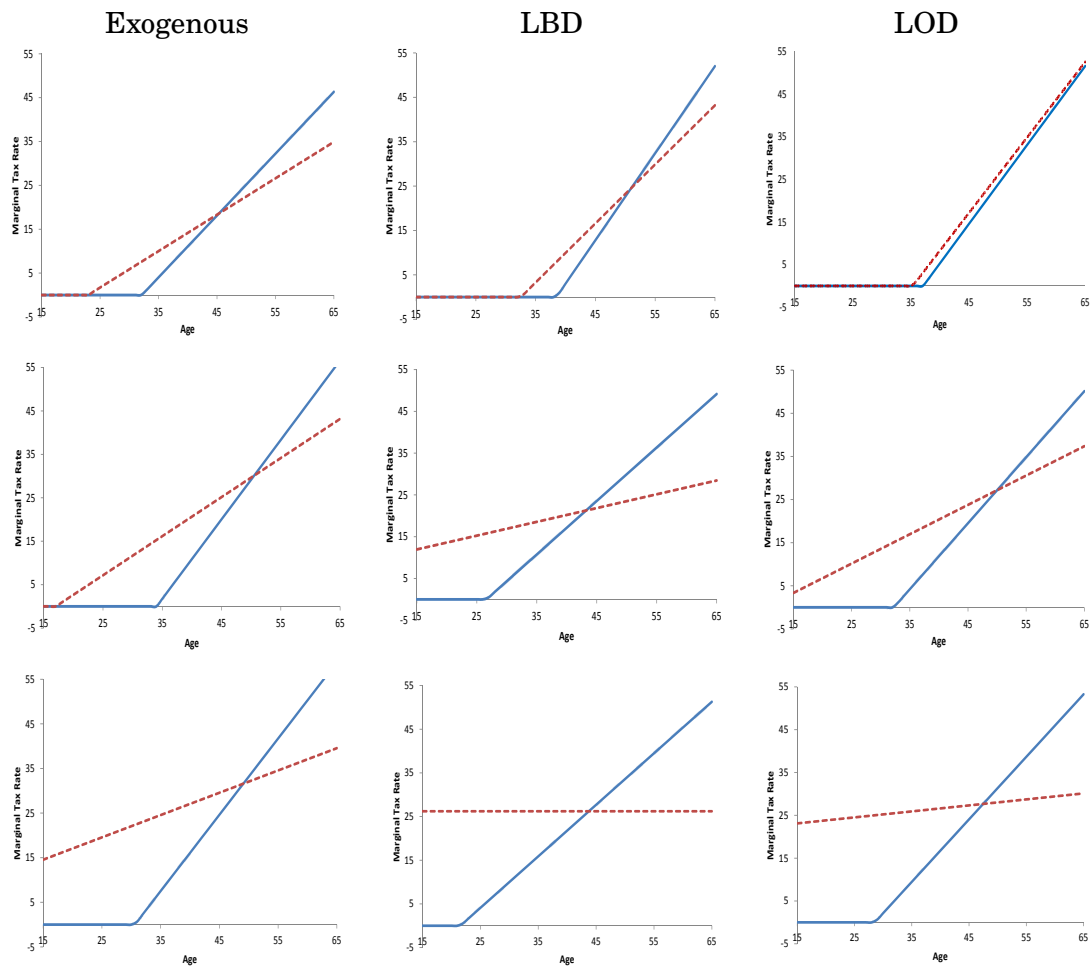


Figure 8: General vs. Partial Equilibrium: Figure displays optimal tax systems for each of the models depending on whether we allow prices to adjust (blue continuous line) or not (red dashed line). For the first row, we keep marginal tax rates on capital at $\tau_k = 30\%$. $\tau_k = 15\%$ for the middle row and $\tau_k = 0\%$ for the bottom one.

Interestingly, the absence of price adjustments substantially increases the role of endogenous human capital in determining optimal policies.

Average hours, do not display much variation with respect to the benchmark. This, again, is in contrast with the case with endogenous price responses, where the increase in capital stock raises the productivity of labor thus reinforcing the importance of labor income tax policy. Finally note that, except for the case of an exogenous age-efficiency profile, the absence of price adjustments leads to lower welfare gains.

As we lower τ_k , the consequences of shutting down the price adjustment become more pronounced. Age dependency is substantially reduced when $\tau_k = 15\%$. Agents start paying taxes at the moment they enter the labor force for both the LBD and the LOD models — starting at 12.28%, in the LBD model, and 4.08% in the LOD model. For the exogenous age-efficiency profile model marginal tax rates become positive at the age of 18. Welfare gains are substantially lower for all models and becomes negative for the LOD model.

Finally, for all three models welfare is lower than at the benchmark (recall that $\tau_k = 30\%$ at the benchmark) when $\tau_k = 0\%$. With regards to the optimal policy, age-dependence all but disappears when human capital is endogenous. Indeed, there is no age-dependence, $\xi_1 = 0$ in the LBD model, whereas marginal taxes increase from 21% at the moment the agent enters the labor force to 28% right before retirement in the LOD model. Only for the exogenous model we see important variation: from 7% when one enters the labor force to 32% right before retirement.

4.3 An Alternative Tax System

We have found that, provided that taxes may depend on age, it is optimal to set $\varrho = 0$. Once we fix ϱ at this value, the reduced computational burden allows us to explore another form of age-dependence. In this case, we use

$$\tau_{w,t} = \xi_0 + \xi_1 t + \xi_2 t^2.$$

Figure 9 displays the optimal marginal tax rates in this case. For all models a concave function obtains, with $\xi_1 > 0$, and $\xi_2 < 0$. Importantly, for all models monotonicity is preserved: marginal tax rates are always increasing in age although at a decreasing speed. Note also that for all models the age at which individuals start paying taxes is reduced relative to the case for which taxes increase linearly with age. What the findings in this section suggest is that it is optimal to keep agents from paying taxes for longer and rapidly increase marginal tax rates as one reaches his or her most productive years.

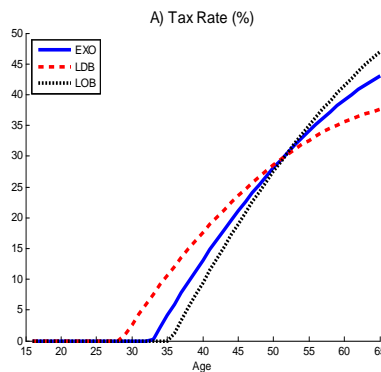


Figure 9: **Alternative Tax System:** The figure displays how the marginal tax rate on income $\tau_{w,t}$ varies with age for each one of the three assumptions about the age-efficiency profile under the alternative tax system. We assume that $\varrho = 0$ and keep τ_k at its benchmark value.

5 Conclusion

We study optimal age-dependent taxes in an overlapping generations where individuals live a meaningful life-cycle and endogenously accumulate human capital.

We find that introducing age dependence leads to substantial gains for all specifications of human capital accumulation. The bulk of this gain is due to the relaxation of credit constraints by lower taxation at younger ages. Progressivity of labor income taxes becomes unnecessary, when marginal tax rates can be conditioned on age, which is in line with Farhi and Werning's [2013] and Weinzierl's [2012] findings.¹⁶ The cost of ignoring the endogeneity of human capital is not large close to the optimum, but increases substantially as we move to sub-optimal systems. This effect is made stronger in the absence of producer price adjustments.

In this paper we only compare steady states. Because the stock of capital is much larger in the new steady state, a full assessment of the benefits of moving to this efficient tax system would require an evaluation of the transition.

References

- S Rao Aiyagari. Uninsured idiosyncratic risk and aggregate saving. *Quarterly Journal of Economics*, 109(3):659–684, 1994. 2.2
- George. Akerlof. The economics of "tagging" as applied to the optimal incometax, welfare programs and manpower planning. *American Economic Review*, 68:8–19, 1978. 1

¹⁶See also [Findeisen and Sachs, 2014]

- Alan Auerbach and Larry Kotlikoff. *Dynamic Fiscal Policy*. Cambridge University Press, 1987. [3](#)
- Spencer Bastani, Soren Blomquist, and Luca Micheletto. The welfare gains of age related optimal income taxation. Technical report, CESifo Group Munich, 2010. [1](#)
- Felicite Bell and Michael Miller. Life tables for the united states social security area 1900-2100. Actuarial Actuarial Study 120, Social Security Administration, 2005. [3](#), [1](#)
- Michael Carlos Best and Henrik Jacobsen Kleven. Optimal income taxation with career effects of work effort. mimeo. LSE, 2013. [1](#)
- Christophe Chamley. Optimal taxation of capital income in general equilibrium with infinite lives. *Econo*, 54:607–622, 1986. [1](#)
- Yongsung Chang, Joao F. Gomes, and Frank Schorfheide. Learning-by-doing as a propagation mechanism. *The American Economic Review*, 92:1498–1520, 2002. [3](#)
- V. V. Chari and Patrick J. Kehoe. Optimal fiscal and monetary policy. In John B. Taylor and Michael Woodford, editors, *Handbook of Macroeconomics*, volume 1, chapter 26, pages 1671–1745. North-Holland: Amsterdam, 1999. [1](#)
- Juan Carlos Conesa, Sigiri Kitao, and Dirk Krueger. Taxing capital? not a bad idea after all. *American Economic Review*, 99(1):25–48, 2009. [1](#), [1](#), [2.2](#), [2.7](#), [4](#)
- David Domeij and Martin Flodén. The labor-supply elasticity and borrowing constraints: Why estimates are biased. *Review of Economic Dynamics*, 9(2):242–1262, 2006. [3](#)
- Andrés Erosa and Martin Gervais. Optimal taxation in life-cycle economies. *Journal of Economic Theory*, 105:338–369, 2002. [1](#), [2](#), [3](#), [1](#), [2.2](#)
- Emmanuel Farhi and Iván Werning. Insurance and taxation over the life cycle. *Review of Economic Studies*, 2013. forthcoming. [1](#), [1](#), [4](#), [5](#)
- S. Findeisen and D. Sachs. Taxes, education policies and college enrollment. mimeo., 2013. [1](#)
- Sebastian Findeisen and Dominik Sachs. Redistribution and insurance with simple tax instruments,. mimeo, 2014. [1](#), [16](#)
- L. Fuster, A. Imrohoroglu, and S. Imrohoroglu. Elimination of social security in a dynastic framework. *Review of Economic Studies*, 74:134–45, 2007. [1](#), [15](#)
- Carlos Garriga. Optimal fiscal policies in overlapping-generation economies. mimeo. Florida State University, 2003. [1](#), [2](#), [3](#), [1](#)
- Miguel Gouveia and R. Strauss. Effect federal individual income tax functions: An exploratory analysis. *National Tax Journal*, 47:317–339, 1994. [1](#)
- J. Heckman, L. Lochner, and C. Taber. Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with hetero-

- geous agents. *Review of Economic Dynamics*, 1:1–58, 1998. **3**
- J. J. Heckman, L. Lochner, and R. Cossa. Learning-by-doing vs. on-the-job training: Using variation induced by the eitc to distinguish between models of skill formation. *NBER WORKING PAPER SERIES*, 2002. **1, 3**
- M. Huggett. Wealth distribution in life-cycle economies. *Journal of Monetary Economics*, 38(3):469–494, 1996. **3**
- Susumu Imai and Micheael P. Keane. Intertemporal labor supply and human capital accumulation. *International Economic Review*, 45:601–641, 2004. **1**
- Bas Jacobs and Lans Bovenberg. Human capitl and optimal positive taxation of capital income. *International Tax and Public Finance*, 17:451–478, 2010. **4**
- Kenneth Judd. Redistributive taxation in a simple perfect foresight model. *Journal of Public Economics*, 28:59–83, 1985. **1**
- Marek Kapicka. Optimal income taxation with human capital accumulation and limited record keeping. *Review of Economic Dynamic*, 9:612–639, 2006. **1, 1**
- Marek Kapička. The dynamics of optimal taxation when human capital is endogenous. mimeo. University of Santa Barbara, 2011. **1, 1, 10**
- Marek Kapička and Julian Neira. Optimal taxation in a life-cycle economy with endogenous human capital formation. mimeo. UC Santa Barbara, 2013. **1**
- Greg Kaplan. Inequality and the lifecycle. *Quantitative Economics*, 3(11-014):471–525, 2012. **3, 1**
- Michael P. Keane. Income taxation in a life cycle model with human capital. Mimeo. Nuffield College, University of Oxford, July 2011. **1, 10, 13**
- Narayana Kocherlakota. Wedges and taxes. *American Economic Review Papers and Proceedings*, 94:109–113, 2004. **1**
- Dirk Krueger and Alexander Ludwig. Optimal progressive taxation and education subsidies in a model of endogenous human capital formation. *American Economic Review, Papers and Proceedings*, 103:496–501, 2013. **1, 1**
- Robert E. Jr. Lucas and Nancy L. Stokey. Optimal fiscal and monetary policy in an economy without capital. *Journal of Monetary Economics*, 12:55–93, 1983. **1, 6**
- E. R. McGrattan and R. Rogerson. Changes in hours worked since 1950. *Federal Reserve Bank of Minneapolis Quarterly Review*, 22(1):2–19, 1998. **3**
- James A. Mirrlees. An exploration in the theory of optimal income taxation. *Review of Economic Studies*, 38:175–208, 1971. **1, 7**
- William B. Peterman. The effect of endogenous human capital accumulation on optimal taxation. Mimeo. University of California - San Diego, July 2011. **1, 10**
- John Rust and Christopher Phelan. How social security and medicare affect retirement behavior in a world of incomplete markets. *Econometrica*, 65(4):781–831, 1997. **3**

Matthew Weinzierl. The surprising power of age-dependent taxes. *Review of Economic Studies*, 78(4):1490–1518, 2011. 1, 1

Matthew Weinzierl. Why do we redistribute so much but tag so little? the principle of equal sacrifice and optimal taxation. Working Paper 18045, NBER, 2012. 1, 4, 5

Iván Werning. Nonlinear capital taxation. Mimeo. MIT, 2010. 8