Optimal Taxation in OLG-ABHI Economies

Quantitative Macro

Raül Santaeulàlia-Llopis

MOVE-UAB and Barcelona GSE

Fall 2017

1. "Taxing Capital? Not a Bad Idea After All!," Conesa, Kitao and Krueger (2009)

2. "On the Optimal Provision of Social Insurance: Progressive Taxation versus Education Subsidies in General Equilibrum," Krueger and Ludwig (2016)

“Taxing Capital? Not a Bad Idea After All!,”
Conesa, Kitao and Krueger (2009)

Taxing Capital? Not a Bad Idea After All! Conesa, Kitao and Krueger (2009)

- Quantitatively characterize the optimal capital and labor income tax in an overlapping generations model (OLG) standard incomplete market (SIM) model.
- Agents face idiosyncratic (uninsurable) labor income shocks and permanent labor productivity differences.
- The optimal capital income tax rate is significantly positive at 36%.
- The optimal progressive labor income tax is, roughly, a flat tax of 23% witha deduction of $7,200 (relative to average household income of $42,000).

What do we know about capital income taxes?

- A classic result in optimal taxation (Judd 1985 and Chamley 1986) is that capital income tax should be zero in the long-run.
- (At least) Two modeling choices can invalidate this results:
  (1) With tight borrowing constraints and/or uninsurable idiosyncratic income risk, in general capital income taxes should not be zero (Hubbard and Judd, 1986, Aiyagari, 1995, and Imrohoroglu, 1998).
  (2) Lifecycle models of optimal capital income taxes show optimal capital income tax is different from zero, at least if the tax code cannot be conditioned on the age of the household (Alvarez et al. 1992, Erosa and Gervais, 2002, and Garriga, 2003).

Goal

Quantitatively characterize the optimal capital and labor income tax in a reallistaclly calibrated model with both elements, an OLG structure and SIM.

In addition:

(1) Agents are heterogenous in their innate ability to generate income (a fixed effect in labor productivity). If society values a equitable distribution of welfare, then this model induces a positive redistributive role for taxes (not only because of idiosyncratic shocks but also because of heterogeneity in innate ability)

(2) The authors alllow progressive taxation to achieve this redistribution. They show that government's desire to tax capital may depend on whether it has access to progressive labor income taxes.

What is social welfare?

To find optimal taxes, we need to define the social welfare function to evaluate alternative policies:

- Welfare criterion: Ex-ante (before ability is realized) expected (with respect to idiosyncratic shocks) lifetime utility of a newborn in a stationary EQ.
- This criterion incorporates the policy maker's concern for (a) insurance against idiosyncratic shocks and (b) redistribution across households with different ability.

  To see the latter note that transferring an extra dollar from the highly able to the less able, ceteris paribus, increases social welfare since the value function is strictly concave in the ability to generate income.

Such insurance and redistribution can be achieved by progressive labor income taxes or taxation of capital income (or both).

The policy maker trades off this concern against the standard distortions these taxes impose on labor supply and capital accumulation.

- Time is discrete, t.
- There are J overlapping generations at any t.
- A continuum of new households is born with mass growing at rate n each t.
- Each HH faces a probability of death by age groups. The probability of survival from age j to $j+1$ is ψ_j, with $\psi_J = 0$.
- No annuity markets, and a fraction of HHs. leaves unintended bequets Tr_t that are redistributed in a lump-sum manner across individuals that are alive.
- After age j_r, agents retire and receive social security payments SS_t, which are financed by proportional payroll taxes $\tau_{ss,t}$, paid up to an income threshold $\overline{y}$.

- HH endowed with 1 unit of time. They decide how much to work in competitive labor markets or consume leisure.
- $a_0 = 0$, besides transfers from accidental bequests.
- Heterogeneity in three dimensions"
  - Age, hence average labor productivity ε_j. Retired agents $\varepsilon_{j \geq j_r} = 0$.
  - Permanent differences in ability (productivity) reflecting education and innate abilities, $i \in I$. Born with it and unchangeable. The probability of begin born with ability α_i is denoted by $p_i > 0$. This feature of the model, together with a SWF that values equity, gives a welfare-enhancing role to redistributive fiscal policies.
  - Workers also face idiosyncratic transitory risk. Let $\eta \in E$ denote a generic realization of this shock un the current period.

- The stochastic process of labor productivity is identical and independent (*iid*) across agents and follows a finite-state Markov chain with stationary transitions over time, i.e.,

$$Q_t(\eta, E) = Prob(\eta' \in E|\eta) = Q(\eta, E) \tag{1}$$

We assume that Q has only strictly positive entries which implies a unique, strictly postiive, invariant distribution associatedwith Q which we denote Π.

All agents start their life with average productivity $\overline{\eta} = \sum_{\eta} \eta \Pi(\eta)$ where $\Pi(\eta)$ is the probability of η under the stationary distribution.

This way, different realizations of the stochastic process then give rise to cross-sectional productivity distributions that become more dispersed as a cohort ages. Hence, in the absense of explicit insurance markets for this risk a progressive tax system is an effective policy to share this risk across agents.

- Individual states: At any t, households are characterized by (a, η, i, j), where a are current holdings of one period risk-free bonds.

  A HH of type (a, η, i, j), working l_j hours commands pre-tax income $\varepsilon_j, \alpha_i, \eta l_j w_t$ where w_t is a common aggregate wage per efficiency unti of labor at time t.

- Aggregate state: $\Phi_t(a, \eta, i, j)$ denote a meaure of households of agents of type (a, η, i, j) at date t. The aggregate state of the economy at time t is completely described by the joint meaure Φ_t over asset positions, labor producitivity status, ability and age.

- Preferences over consumption and leisure $\{c_j, (1-l_j)\}_{j=1}^{J}$ are assumed of the form,

$$E\left\{\sum_{j=1}^{J}\beta^{j-1}u(c_j, 1-l_j)\right\}$$

where β is the time discount factor. Expectation are taken over η and mortality.

The Model
Technology

- The aggregate resource constriant is,

$$C_t + K_{t+1} - (1-\delta)K_t + G_t \leq AK_t^{1-\theta} N_t^{\theta}$$

where K_t, N_t and C_t represent aggregate capital, labor (measured in efficiency units) and consupmtion. Labor share is θ. The constant A normalizes model units. The depreciation rate of capital is δ.

The government engages in three activities:

- It levies taxes
- It spends resources G_t
- It runs a balanced budget SS system.

- The SS system is defined by benefits SS_t for each retired HH, independent of HH earnings history.
- SS taxes are levied up to a maximum lbor income level $\overline{y}$ as in the actual US system.
- The payroll tax rate $\tau_{ss,t}$ is set to assure period-by-period budget balance of the system.
- We take the *SS* system as exogenously given and not as subject of optimization of the policy maker.

- Government faces a sequence of exogenously given government consumption $\{G_t\}_{t=1}^{\infty}$ and has three fiscal instruments to finance expenditure.

(1) A proportional tax $\tau_{c,t}$ on consumption expenditures, which we take as exogenously given.

(2) A capital income tax according to a constant marginal tax rate $\tau_{K,t}$. That is, capital income tax is proportional. Note that capital income is $r_t(a + Tr_t)$, where r_t is the risk free interest rate, a asets and Tr_t the accidental bequests.

(3) Government can tax each individuals taxable labor income according to a progressive schedule. Define HH's pre-tax labor income as $yp_t = w_t \alpha_i \varepsilon_j \eta l_t$. A part of this pre-tax income is accounted by SS contributions paid by the employer $ess_t = .5\tau_{ss,t} \min\{yp_t, \overline{y}\}$ which is not part of taxable income under the U.S. tax lay. Thus, taxable labor income is

$$y_t = \begin{cases} yp_t - ess_t & \text{if} \quad j < j_r \\ 0 & \text{if} \quad j \geq j_r \end{cases}$$

That is, pension income is not taxable.

- Labor income taxes can be made an arbitrary function of individual taxable labor income. We denote the taxe code by $T(y)$.

  Our investigation of the optimal tax code involves finding the labor income tax function T and the capital tax rate τ_K that maximize a SWF.

- Workers cannot insure by trading explicit insurance contracts (SIM).
- Annuity markets insuring idiosyncratic mortality risk are assumed absent.
- Agents trade 1-period risk free bonds to self-insure against labor productivity risk.
- This self-insurance is limited by preventing agents from borrowing altogether, in line with Aiyagari 1994 and Krusell and Smith 1998.[1]

[1] Instead, if agents were allowed to borrow, it may be optimal for an agent with a low ψ_j to borrow up to the limit, since with high probability she would not have to pay back this debt. Such strategic behavior would be avoided if lenders could provide loans with interest rates that depend on survival probabiliies.

- Let $a \in \mathbf{R}_+$, $\eta \in \mathbf{E} = \{\eta_1, ..., \eta_n\}$, $i \in \mathbf{I} = \{1, ..., M\}$, $j \in \mathbf{J} = \{1, ...J\}$ and let the $\mathbf{U} = \mathbf{R}_+ \times \mathbf{E} \times \mathbf{I} \times \mathbf{J}$.

- Let $\mathbf{B}(\mathbf{R}_+)$ be the Borel σ-algebra of $\mathbf{R}_+$ and $\mathbf{P}(\mathbf{E})$, $\mathbf{P}(\mathbf{I})$, and $\mathbf{P}(\mathbf{J})$ the power sets of $\mathbf{E}, \mathbf{I}$ and $\mathbf{J}$, respectively.

- Let $\mathbf{M}$ be the set of all finite measures over the measurable space $\mathbf{U}$,

$$\mathbf{B}(\mathbf{R}_+) \times \mathbf{P}(\mathbf{E}) \times \mathbf{P}(\mathbf{I}) \times \mathbf{P}(\mathbf{J})$$

Definition of a Recursive Competitive Equilibrium

Given a sequence of government expenditures $\{G_t\}_{t=1}^{\infty}$, consumption tax rates $\{\tau_{c,t}\}_{t=1}^{\infty}$ and initial conditions K_1 and Ψ_1, a competitive equilibrium is a sequence of

- functions for the household, $\{v_t, c_t, a_t', l_t\}_{t=1}^{\infty}$,
- production plans for the firm, $\{K_t, N_t\}_{t=1}^{\infty}$,
- prices $\{r_t, w_t\}_{t=1}^{\infty}$,
- government labor income tax functions $\{T_t : \mathbf{R}_+ \to \mathbf{R}_+\}_{t=1}^{\infty}$ and capital income taxes $\{\tau_{K,t}\}_{t=1}^{\infty}$,
- social security taxes $\{\tau_{ss,t}\}_{t=1}^{\infty}$ and benefits $\{SS_t\}_{t=1}^{\infty}$,
- transfers $\{Tr_t\}_{t=1}^{\infty}$, and
- measures $\{\Psi_t\}_{t=1}^{\infty}$, with $\Psi_t \in \mathbf{M}$ such that:

(1) Given prices, policies, transfers, and initial conditions, for each t, v_t solves the functional equations (with c_t, a'_t, and l_t as associated policy functions):

$$v_t(a,\eta,i,j) = \max_{c\geq 0, a'\geq 0, 0\leq l\leq 1} u(c, 1-l) + \beta\psi_j \int v_{t+1}(a',\eta',i,j+1)Q(\eta, d\eta')$$

subject to,

- For $j < j_r$,

$$(1+\tau_{c,t})c + a' = y_t + (1+(1-\tau_{K,t}r_t)(a+Tr_t) - T_t(y_t),$$

  with taxable income $y_t = w_t\varepsilon_j\alpha_i\eta l - ess_t$ and $ess_t = \tau_{ss,t}\min\{w_t\varepsilon_j\alpha_i\eta l, \overline{y}\}$.
- For $j \geq j_r$,

$$(1+\tau_{c,t})c + a' = SS_t + (1+(1-\tau_{K,t}r_t)(a+Tr_t).$$

(2) Prices w_t and r_t satisfy:

$$r_t = (1-\theta)A\left(\frac{K_t}{N_t}\right)^{-\theta} - \delta$$
$$w_t = \theta A\left(\frac{K_t}{N_t}\right)^{1-\theta}$$

(3) The social security policies satisfy,

$$\tau_{ss,t}\int ess_t\Phi_t(da\times d\eta\times di\times dj) = SS_t\int \Phi_t(da\times d\eta\times di\times dj)$$

(4) Transfers (bequests) are given by:

$$Tr_{t+1}\int \Phi_t(da\times d\eta\times di\times dj) = \int (1-\psi_j)a_t'(a,\eta,i,j)\Phi_t(da\times d\eta\times di\times dj)$$

(5) Government budget balance,

$$
\begin{aligned}
G_t = \int \tau_{K,t} r_t (a + Tr_t) \Phi_t (da \times d\eta \times di \times dj) \\
+ \int T_t(y_t) \Phi_t (da \times d\eta \times di \times dj) \\
+ \tau_{c,t} \int c_t(a, \eta, i, j) \Phi_t (da \times d\eta \times di \times dj)
\end{aligned}
$$

(6) Market clearing,

$$K_t = \int a\Phi_t(da \times d\eta \times di \times dj)$$

$$N_t = \int \eta_j \alpha_i \eta l_t \Phi_t(da \times d\eta \times di \times dj)$$

and

$$\int c_t(a,\eta,i,j)\Phi_t(da \times d\eta \times di \times dj) + K_{t+1} + G_t = AK_t^{1-\theta}N_t^{\theta} + (1-\delta)K_t$$

(7) Law of motion,

$$\Phi_{t+1} = H_t(\Phi_t)$$

where the function $H_t : \mathbf{M} \to \mathbf{M}$ can be written explicitly as follows.

- For all J such that $1 \notin J$

$$\Phi_{t+1}(A \times E \times I \times J) = \int P_t((a, \eta, i, j); A \times E \times I \times J)\Phi_t(da \times d\eta \times di \times dj)$$

where

$$P_t((a, \eta, i, j); A \times E \times I \times J) \begin{cases} Q(e, E)\phi_j & \text{if } a'_t(a, \eta, i, j) \in A, i \in I, j+1 \in J \\ 0 & \text{else} \end{cases}$$

- For $J = \{1\}$,

$$\Phi_{t+1}(A \times E \times I \times \{1\}) = (1+n)^t \begin{cases} \sum_{i \in I} p_i & \text{if } 0 \in A, \overline{\eta} \in E \\ 0 & \text{else} \end{cases}$$

Definition of Recursive Stationary Equilibrium

A stationary EQ is a competitive EQ in which per capita variables and functions as well as prices and policies are constant, and aggregate variables grow at the constant rate of the population n.

- Agents are born at age 20 (model age 1). They retire at age 65 (46 in the model) and die with probability one at age 100 (81 in the model).
- Population grows at an annual rate of $n = 0.011$, the long-run average in the US.
- Survival probabilities are taking from a previous study, Bell and Miller 2002.

Calibration
Preferences

- Time-separable preferences over consupmtion and leisure. Benchmark assumes a standard Cobb-Douglas specifiation,

$$u(c, 1-l) = \frac{\left(c^{\gamma}(1-l)^{1-\gamma}\right)^{1-\sigma}}{1-\sigma}$$

where γ is a share parameter determining the relative importance of consumption, σ determines the risk aversion of the HH.[2]

Set $\sigma = 4$ and choose $\beta = 1.001$ and $\gamma = 0.377$ such that the stationary EQ of the economy with benchmark tax system features a capital-output of 2.7 and an average share of time worked of one-third of the time endowment. The calibrated σ and γ implies an intertemporal elasticity of substitution is 0.5.

Also, with these shape of preferences the elasticity of labor supply is 1.

[2] The CRRA in consumption is given by $-\frac{cu_{cc}}{u_c} = \sigma\gamma + 1 - \gamma$.

Three components, a deterministic age-dependent component ε_j, a type-dependent fixed effect α_i, and a transitory idiosyncratic shock η. The natural logarithm of wages of a HH is,

$$\log(w_t) + \log(\varepsilon_j) + \log(\alpha_i) + \log(\eta)$$

- The age-productivity profile $\{\varepsilon_j\}_{j=1}^{j_r-1}$ is taken from Hansen (1993).
- We consider two ability types with equal population mass $p_i = 0.5$ and fixed effects $\alpha_1 = e^{-\sigma_\alpha}$ and $\alpha_2 = e^{\sigma_\alpha}$ so that $E(\log(\alpha_i)) = 0$ and $Var(\log(\alpha_i)) = \sigma_\alpha^2$.
- The transitory income shock has some persitence following an AR(1) with paramter ρ and unconditional variance σ_η^2.

This implies three free parameters to choose $(\sigma_\alpha^2, \rho, \sigma_\eta^2)$. We choose them targeting how cross-sectional labor income dispersion over the life cycle.

Storesletten, Telmer and Yaron (2004) document that (i) at cohort of age 22 the cross-sectional variance of household labor income is 0.2735 and (ii) at age 60 is about 0.90, and (iii) it increases roughly linearly in between.

In the current model labor supply (and hence labor earnings) are endogenous, responding optimally to the labor productivity process. We choose the three parameters so that in the benchmark parameterization the model displays a cross-sectional household age-earnings variance profile consistent with these facts. The implied parameters are $(\sigma_\alpha^2 = 0.14, \rho = 0.98, \sigma_\eta^2 = 0.0289)$.

- The capital share is set to $1-\theta=0.36$.
- The depreciation rate is set to $\delta=0.08$ to match the investment-output ratio of 0.255.[3]

[3] Note that our parameter choices yield a benchmark real interest rate of 5 percent, and with population growth of 1.1 percent the economy is deep in the dynamically effficient region.

The government stands in for all levels (federal, state and local) and consumres resources, collects tax revenues and operates a social security system.

- Government spending G accounts for 17% of GDP in the stationary equilibrium. G is kept constant across the quantitative experiments, hence if an income tax system delivers higher Y, the corresponding ratio G/Y declines.
- The proportional consumption tax is set to $\tau_c = 0.05$, following Mendoza, Razin and Tesar (1994).
- The pay-as-you-go SS system is defined by a payroll tax of $\tau_{ss} = 0.124$ of labor income up toa limit of 2.5 times average income, with benefits determined by the budget balance of the system.

- We want to determine the optimal capital and labor tax functions. Ideally one would impose no restriction on the set of tax functions the government can choose from. Maximization over such an unrestricted set is computationally infeasible, however.

  Therefore we restrict the set of tax functions to a flexible three parameter family proposed by Gouveia and Strauss (1994). If y is taxable income, total taxes are given by:

$$T^{GS}(y;\kappa_0,\kappa_1,\kappa_2) = \kappa_0\left(y - \left(y^{-\kappa_1}+\kappa_2\right)^{-\frac{1}{\kappa_1}}\right)$$

  where $(\kappa_0,\kappa_1,\kappa_2)$ are parameters.

  Roughly speaking, κ_0 controls the level of the average tax rate, κ_1 determines the progressivity of the tax code. For $\kappa_1 \to 0$ the tax system reduces to a pure flatt tax, while other values encompass a wide range of progressive and regressive tax functions.

  Without discriminating between capital and labor GS estimate for the US that $\kappa_0 = 0.258$ and $\kappa_1 = 0.768$. The authors in the current paper use these estimated tax system as benchmark. The parameter κ_2 adjusts to ensure government budget balance.

- Define y_l and y_k as taxable labor and capital income, respectively.
- The set of tax functions the government optimizes over is given by:

$$\mathcal{T} = \{\, T_l(y_l), T_k(y_k) : T_l(y_l) = T^{GS}(y_l; \kappa_0, \kappa_1, \kappa_2) \text{ and } T_k(y_k) = \tau_k y_k \}$$

  and thus by the four parameters $(\kappa_0, \kappa_1, \kappa_2, \tau_K)$ one of which (we take κ_2) is determined by budget balance.
- Note that the the choices of $(\kappa_0, \kappa_1, \tau_K)$ are restricted by the requirement that there has to exist a corresponding κ_2 that balances the budget.

- We now need to define the SWF that ranks different tax functions. We assume that the government wanst to maximize the ex-ante lifetime utility of an agent born into the stationary EQ implied by the chosen tax function.

  That is, the government's objetive is to maximize

$$SWF(\kappa_0, \kappa_1, \tau_k) = \int v_{(\kappa_0,\kappa_1,\tau_k)}(a = 0, \eta = \overline{\eta}, i, j = 1) d\Phi_{(\kappa_0,\kappa_1,\tau_k)}$$

  Given that all newborn HH start with zero assets and average labor productivity, social welfare is simply equal to average expected lifetime utility across the two ability groups.

  Here $v_{(\kappa_0,\kappa_1,\tau_k)}$ and $\Phi_{(\kappa_0,\kappa_1,\tau_k)}$ are the value function and invariant cross-sectional distribution associated with tax system characterized by $(\kappa_0, \kappa_1, \tau_k)$.

- The optimal capital income tax is $\tau_k = 0.36$.
- The optimal labor income tax is characterized by $\kappa_0 = 0.23$ and $\kappa_1 \approx 7.0$. Therefore, the labor income tax code is basically a flat tax with marginal rate of 23% and a deduction of about $7,200 (relative to average income of $42,000)

- Capital drops substantiall (by 6.64%) and hence also output and consumption. This is a consequence of the heavy capital income tax 0.36 relative to the benchmark (where the highest marginal rate is 0.26).
- The change in taxes also induces adjustments in labor supply. While the average hours worked drop by 0.56%, labor efficiency units drop by only 0.11%; thsu labor supply shifts from less to more productive households.

- The optimal tax implies higher welfare, equivalent to 1.33% increase in consumption in all ages and states of the world, keeping labor supply allocations unchanged. Note that the decline in capital (and consumption) reduces welfare, but the increase in leisure increases it.
- Given the form of the utility function, the welfare consequences of switching from a steady state consumption-labor allocation (c_0, l_0) to (c_*, l_*) are given by,

$$CEV = \left[\frac{W(c_*, l_*)}{W(c_0, l_0)} \right]^{\frac{1}{\gamma(1-\sigma)}} - 1$$

where $W(c, l)$ is the expected lifetime utility at birth of HH, given the tax system.

- We can decompose CEV into the change from c_0 to c_* and the change from l_0 to l_*. Let CEV_c and CEV_l be defined as,

$$W(c_*, l_0) = W(c_0(1 + \text{CEV}_c), l_0)$$
$$W(c_*, l_*) = W(c_*(1 + \text{CEV}_l), l_0)$$

Then, it is easy to verify that $1+\text{CEV} = (1+\text{CEV}_c)\ (1+\text{CEV}_l)$ or $\text{CEV} \approx \text{CEV}_c + \text{CEV}_l$.

- We further decompose CEV_c into a level effect CEV_{cL} and a distribution effect CEV_{cD}:

$$W(\widehat{c}_0, l_0) = W(c_0(1 + \text{CEV}_{cL}), l_0)$$
$$W(c_*, l_0) = W(\widehat{c}_0(1 + \text{CEV}_{lD}), l_0)$$

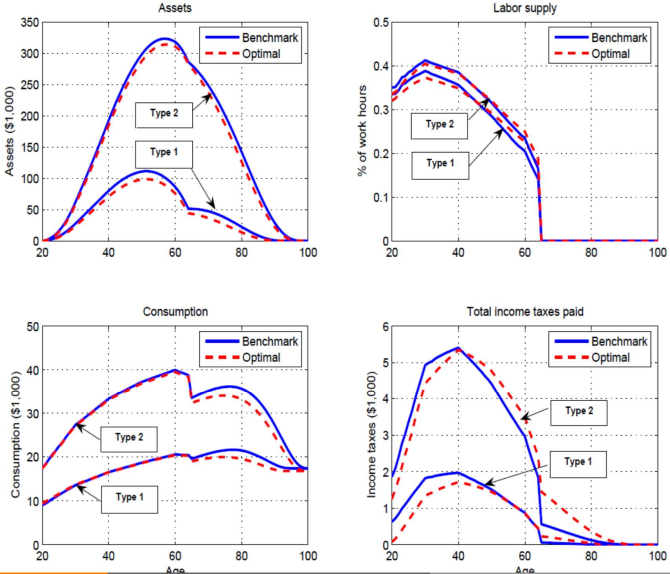
where $\widehat{c}_0 = (1 + g_C)c_0 = \frac{C_*}{C_0}c_0$ is the consumption alloation resulting from scaling the allocation c_0 by the change in aggregate consumption $\frac{C_*}{C_0}$. A simple calculation shows that the level effect equals the growth rate of consumption $\text{CEV}_c L = \frac{C_*}{C_0} - 1$. A similar decomposition applies to leisure.

Table: Decompostion of Welfare

CEV (in percent)	1.33
Consumption:	
Total	1.29
Level	-1.63
Distribution	2.97
Leisure:	
Total	0.04
Level	0.41
Distribution	-0.37

- The welfare gains stem from a better allocation of consumption across types and states of the world, and from a reduction of the average time spent working.
- This more than offsets the lower average level of consumption and the less favorable, in utility terms, distribution of leisure over the life cycle.

Optimal Lifecycle Profiles by Permanent Types



(Fig1) The asset holdings (relevant base for the capital income tax). Hump-shape behavior typical of any life cycle model. That is, the main burden of the capital income tax is borne by the households aged 40 to 70.

The negative impact on asset accumuation of higher capital income taxes due to optimal taxes (compared with the benchmark) is visible across all ages.

(Fig2) Labor supply (relevant base for the labor income tax).

- First, independent of the tax, labor supply tends to decline (despite the fact that labor productiivity peaks at age 50). This decline is mainly driven by $\beta > 1$ and a substantially positive (after tax) return that makes beneficial for HH to postpone leisure to older ages.

- Second, optimal tax code induces HH to work more at ages at which they are more productive. This is due to lower income taxes and a sizeable deduction that makes the labor supply follow closer the age-efficiency productivity profile, as it allivates the severity of the borrowing constraint early in life. Especially for the low-skilled group the increase in labor supply at age 50 to 60 is substantial, indicating a high elasticity of hours with respect to marginal labor income taxes for this group. Overall, the optimal tax system induces a flatter life cycleprofiel of labor supply thus leisure.

(Fig3) Consumption shows a plausible hump over the life cycle and a discrete fall at the time of retirement, due to the nonseparability between consumption and leisure. Relative to benchmark, a lager optimal capital income tax makes future consumption mor expensive and thus flattens its profile.

(Fig4) The optimal tax code leads to more redistribution across types, by taxing more the high-skilled, high labor income-earners who also hold a larger fration of financial assets. Alo, since under the optimal tax code HH aged 40 to 60 work more than under the benchmark, they pay higher labor income taxes (despite the fact that their marginal tax rates have been reduced). Finally, the higher capital income taxes of the optima tax code explain why retired capital holders pay a larger tax bill under this system.

- The crucial elements of the model include
  - (i) Endogenous labor choice, formally represented by $\gamma < 1$.
  - (ii) Tight borrowing constraint $a' \geq 0$.
  - (iii) Ex-ante heterogeneity in ability $\sigma_\alpha^2 > 0$.
  - (iv) Heterogeneity in productivity shocks $\sigma_\eta^2 > 0$.
  - (v) Model elements that undergo a meaningful life cycle (mortality risk, age-efficiency profile, PAYGO SS system).
- Next table summarizes the optimal tax code under vatrios versions of the model wher various combinations of these elements are shut down. In all models paramters have always been re-calibrated to match (with the GS tax function) the same targets as in the benchmark.

Summary of Results

Table 4: Summary of Quantitative Results

Model	End. Lab.	BC	Type	Idio.	Life Cyc.	β	r	τ_k	τ_l	Prog.
M1	No	No	No	No	No	0.983	4.5	10	19	No
M2	No	No	No	No	Yes	1.001	3.2	-24	100	Yes
M3	No	Yes	Yes	Yes	Yes	1.001	4.3	-34	100	Yes
M4	Yes	No	No	No	No	0.979	4.7	20	17	No
M5	Yes	No	No	No	Yes	1.009	5.6	34	14	No
M6	Yes	No	Yes	No	Yes	1.009	5.2	32	18	Yes
M7	Yes	No	Yes	Yes	Yes	1.005	5.6	35	23	Yes
Bench	Yes	Yes	Yes	Yes	Yes	1.001	5.6	36	23	Yes

- Positive capital income taxes requires endogenous labor supply
- The size of the optimal capital income tax (given endogenous labor supply) depends crucially on the presence of realistic life cycle elements.
- While heterogeneity in innate ability and productivity shocks are key determinants of the progressivity of labor income, they do not affect the capital income result.
- Borrowing constraints are not key to having the capital income results if labor income is allowed to be progressive.

“On the Optimal Provision of Social Insurance: Progressive Taxation versus Education Subsidies in General Equilibrum,” Krueger and Ludwig (2016)

(1) Goal: Quantitatively characterize the optimal tax and education policy transition in an economy where progressive taxes provide social insurance against idiosyncratic wage risk, but distort the education decision of households.[4]

(2) Optimally chosen college education subsidies mitigate these distortions.

(3) Highlight the general equilibrium feedback effects from policies to relative wages of skilled and unskilled workers: subsidizing higher education increases the share of workers with a college degree thereby reducing the college wage premium which has important redistributive benefits.

(4) Optimal policy results:

   (a) Education subsidies are always high ($\theta =$150% the tuition cost).
   (b) The optimal progressivity is optimally moderate. Importantly, optimal income tax progressivity crucially depends on whether the transitional costs of policies are explicitly taken into account and how strongly the college wage premium responds to policy changes in GE.

[4] Use a large-scale OLG model with endogenous human capital formation, borrowing constraints, income risk, intergenerational transmission of wealth and ability, and incomplete financial markets.

- In the presence of uninsurable idiosyncratic earnings risk, progressive taxation provides valuable social insurance among ex ante identical HHs.

  In addition it might enhance equity among ex ante heterogenous HHs, which is beneficial if the SWF values such equity.

- However, if high-earnings HHs face higher average tax rates than low-earnings HHs, this might discourage the incentives of these HHs to become earnings-rich through making conscious human capital accumulation decisions. That is, in addition to standard distortions of labor supply, progressive labor income taxes also reduce th eincentivs to acquire higher education, generating a non-trivial quantitative trade-off for the benevolent uilitarian government.

  The resulting skill distribution in the economy worsens, and aggregate economic activity might be depressed through this channel, which compounds the potentially adverse impact of progressive taxes on production through the classic labor suply channel. The distortion on human capital acquisition can be potentially mitigaged by an education subsidy which then becomes part of the optimal fiscal constitution.

This paper

- Quantitatively characterize the optimal mix of progressive income tax and education policy transition, within a simple parametric class, in an economy where progressive taxes provde social insurance against idiosyncratic wage risk, but distort the education decision of HHs.
- Optimally chosen tertiary education subsidies mitigate these distortions, making both policies potentially complementary.
- Academic talent is transmitted across generations through two channels: persistence of innate ability vs. the impact of parental education.
- Different forms of labor are imperfect substitutes generate GE feedback effects from policies to relative wages of skilled and unskilled workers.

The authors show that:

(1) Subsidizing higher education has important redistributive benefits, by shrinking the college wage premium in GE. This Stiglitz (1982) effect of fiscal policy on relative factor prices may make progressive taxes and education subsidies potential *policy substitutes* for providing SI.

(2) A full characterization of the transition path is crucial for policy evaluation.

(3) Optimal policy result (next page)

(a) Education subsidies are optimally high ($\theta =$150% the tuition cost).

(b) The optimal labor income tax progressivity is moderate.

The authors follow the Ramsey tradition and restrict the choices of the government policy to simple, easily implementable tax policies. The labor income tax function is:

$$T_t(y_t) = \max\left\{0, \tau_{l,t}\left(y_t - d_t\frac{Y_t}{N_t}\right)\right\} = \max\left\{0, \tau_{l,t}\left(y_t - Z_t\right)\right\} \tag{2}$$

where y_t is taxable labor income, $\frac{Y_t}{N_t}$ is the per capita income in the economy and $Z_t = d_t\frac{Y_t}{N_t}$ measures the size of the labor income tax deduction. Therefore for every period there are two policy tax parameters $(\tau_{l,t}, d_t)$. Note that the tax system is potentially progressive (if $d_t > 0$) or regressive (if $d_t < 0$).

The optimal tax deduction is $d = 6\%$ of average income and the contant marginal tax rate stands at approximately $\tau_l = 22\%$. This intertemporally optimal tax reform generates welfare gains equivalent to more than 3% of permanent consumption, relative to the status quo policy, which is calibrate to broadly approximate the U.S. policy ($\theta = 39\%, \tau_l = 28\%, d = 27\%$).

- Population grows at an exogenous rate.
- Parents give birth to children at age j_f and fertility rate of household is f, constant across education groups. Note f is also the number of children per household.
- Let ψ_j be the age-specific survival rate. Assume $\psi_j = 1$ for all $j = 0, .., j_r - 1$, and $0 < \psi_j < 1$ forall $j = j_r, ..., J - 1$, where j_r is the fixed retirement age and J denote the maximum age (hence $\psi_J = 0$).
- Population dynamics are:

$$N_{t+1,0} = fN_{t,j_f}$$
$$N_{t+1,j+1} = \psi_j N_{t,j}, \quad \text{for} \;\; j = 0, ..., J$$

The population growth rate is $n = f^{\frac{1}{1+j_f}} - 1$.

The Model

Technology

- Worker skills s are defined by college attainement with $s = \{c, n\}$, where $s = c$ denotes college and $s = n$ noncollege.
- Output is produced according to a standard Cobb-Douglas production function,

$$Y_t = K_t^{1-\theta} L_t^{\theta} = K_t^{1-\theta} \left[\left(L_{t,n}^{\frac{\rho-1}{\rho}} + L_{t,c}^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}} \right]^{\theta}$$

  where $L_{t,s}$ denotes the aggregate labor of skill s, measured in efficiency units andlet K_t denote the capital stock. Note that skilled and unskilled labor are imperfectly substitutable (but within skill groups labor is perfectly substitutable across different ages) with an elasticity of substitution ρ. As long as $\rho > 1$, skilled and unskilled work are imperfect substitutes. Also θ is labor share.

- Perfect competition among firms and CRS production function implies zero profits for all firms and an indeterminate size distribution of firms, so there is no need to specify the ownership structure of firms in the household sector, and without loss of generality we can assume the existence of a single representative firm.

- Profit maximization of firms implies the standard FOCs

$$r_t = (1-\theta)\left(\frac{K}{L}\right)^{-\theta} - \delta$$

and

$$w_{n,t} = \theta\left(\frac{K}{L}\right)^{1-\theta}\left(\frac{L_{n,t}}{L_t}\right)^{-\rho}$$
$$w_{c,t} = \theta\left(\frac{K}{L}\right)^{1-\theta}\left(\frac{L_{c,t}}{L_t}\right)^{-\rho}$$

- Hence, the skill premium is,

$$\frac{w_{c,t}}{w_{n,t}} = \left(\frac{L_{c,t}}{L_{n,t}}\right)^{-\rho}$$

and depends on the relative supplies of non-college and college labor (unless $\rho = 1$).

- HHs are born at age $j = 0$ and form an independent HH at age $j_a = 18$.
- HHs give birth at age j_f and children live with adult HH until they reach 18. Hence for parent ages $j = \{j_f \ldots, j_f + j_a - 1\}$ children are present in the parental HH.
- Parents face a standard time-separable expected lifetime utility function,

$$E_{j_a} \sum_{j=j_a}^{J} \beta^{j-j_a} u \left(\frac{c_j}{1 + \mathbf{1}_{\mathcal{J}_s} \zeta f}, l_j \right)$$

  where c_j is total HH consumption, l_j is leisure, and $\mathbf{1}_{\mathcal{J}_s}$ is an takes value one for the periods when children are living with the HH, that is, for $j \in \mathcal{J}_s = [j_f, j_f + j_a - 1]$, and zero otherewise. ζ is an adult equivalence parameter.
- Expectations are taken w.r.t. mortality risk and labor productivity risk.
- There is altruism: At parental age j_f (or $j_f + j_a$?), when children leave the HH, the children's expected lifetime utility enters the parental lifetime utility function with a weight $\nu\beta^{j_f}$, where ν measures the strength of parental altruism.

- At age $j = j_a$, before any decision is made, HHs draw their innate ability to go to college, $e \in \{e_1, ..., e_N\}$ according to a distribution $\phi(e|s_p, \gamma_p)$ that depends on parental education s_p and parental productivity γ_p defined below.
- Innate ability e also affects future wages diretly and independently of educaiton (in an stochastic way), described below.

- A HH at age j_a must decide whether going to college or not (it takes some periods to get the degree), and college implies a per-period resource cost of $\kappa w_{t,c}$ that is proportional to aggregate college wages, $w_{t,c}$. In the quantitative implementation of the model a period will be four years, so HH that decide to go to college go for one period only.
- Education subsidies imply that a fraction θ_t of the resource cost is paid by the government.
- In addition, a constant fraction θ_{pr} of the education costs is borne by private subsidies paid from accidental bequests, described below (think of alumni donations and support from private foundations).
- Going to college also rquires a fraction $\xi(e) \in [0, 1]$ of time at age j_a in the period in which the housheold attends school. This fraction is a function of e to reflect that more able people rquire less time to elarn and thsu can enjoy more leisure time or work longer hours while attending college.
- A HH that completed college (all that go, complete) has skill $s = c$, and $s = n$ otherwise.

- HHs start their economic life at age j_a with an initial endowment of financial wealth $b \geq 0$ received as inter-vivo transfer from their parents. This is a one-time payment only and captures the idea that parents finance part of the children's college.
- These inter-vivo transfer, after having observed their child's ability e, but not conditioning on the child going to college or not.
- All HHs also receive transfers from accidental bequests. We assume that asses of HH that die at age j are redistributed uniformly across all HHs of age $j - j_f$, that is, among the age cohort of their children (parents do not know their chldren, but they know their age so they distribute for the entire children's age cohort). These age dependent transfers are denote by $Tr_{t,j}$.

- In each period HH are endowed with one unit of time. A HH of age j with skill $s \in \{n, c\}$ earns a wage

$$w_{t,s}\varepsilon_{j,s}\gamma\eta$$

per unit of time worked.

That is, wages depend on

(i) An skills-specific average wage $w_{t,s}$,

(ii) A deterministic age profile $\varepsilon_{j,s}$ that differs across education groups $s = \{c, n\}$,

(iii) A fixed effect $\gamma \in \Gamma_s = \{\gamma_{H,s}, \gamma_{L,s}\}$ that generates permanent differences within each skill group, and

(iv) An idiosyncratic stochastic shock η.

- The probabilities of drawing fixed effects and idiosyncratic shocks differ across skill groups.
    - The probability of drawing fixed effects (prior to labor market entry) is a function of the ability of the HH, $\pi_s(\gamma|e)$.
    - The idiosyncratic shock η is mean-reverting and follows an skills-specific Markov chiang with states $\mathcal{E}_s = \{\eta_{s,1}, ..., \eta_{s,M}\}$ and transitions $\pi_s(\eta'|\eta) \geq 0$. Let Π_s denote the invariant distribution associated with π_s. Note that prior to making the education decision a HH's η is drawn from Π_n, because work done during the period agents can opt to go to college agents also work (as noncollege HHs).

- Thus, at the beginning of every period in working life the individual state variables are
$$(j, s, \gamma, \eta, a)$$
HH's age j, education s, fixed effect γ, productivity shock η, and assets a.

Note that e is not an individual state variable. In order to permit the share of HHs that go to college to vary smoothly with economic policy it is important that the set $\{e_1, ..., e_M\}$ is sufficiently large. However, given the large state space for HHs of working age, additionally keeping track of e is costly. Instead, stochastically mapptin e into a fixed effect γ after the eduation decision and restricting γ to take only four values (two value per skill group) reduces this burden signficantly. Assumption to achieve a balance between computational feasibility and accuracy.

- HHs can only insure against mortality and labor productivity shocks by accumulating a risk-free one-period bond that pays a real interst rate of r_t. In equilibrium the total net supply of this bond equals the capital stock K_t in the economy, plust the oustanding government debt B_t.
- The is a borrowing limit set by the financing of college education. HHs that borrow to pay for college tuition and consumption while in college face age-dependent borrowing limits $\underline{A}_{j,t}$ and also face the constraint that their balance of outstanding student loans cannot increase after college comletion. This assumption rules out that student loans are used for general consumption smoothing.
- The constraints $\underline{A}_{j,t}$ are set such that student loans need to be fully repaid by age j_τ at which early mortality sets in. This way HH never die in debt (personal bankruptcy is not an issue).
- Borrowing for purposes othen than student loans is rule out. This implies that HHs without a college degree can never borrow.
- The constraints $\underline{A}_{j,t}$ as being determined by public student loan programs, and one may interpret the borrowing limits as government policy parameters that are being held fixed in our analysis.

- The government finances an exogenous stream G_t of non-education expenditures and an endogenous streatm E_t of education expenditures.
- It does so by issuing government debt B_t (with a given initially given B_0), by levying linear consumption taxes $\tau_{c,t}$, linear capital income taxes $\tau_{k,t}$ (where capital income is $r_t a_t$, and labor income taxes $T_t(y_t)$ that are not restricted to be linear. The labor income tax follows the specification described above in equation 2.
- Consumption and capital income taxes $(\tau_{c,t}, \tau_{k,t})$, are exogenously given, but we optimize over labor income tax schedules.

- The government uses tax revenues to finance education subsidies and exogenous government spending,

$$G_t = gY_t$$

where the share of output $g = \frac{G_t}{Y_t}$ is a parameter calibrated from the data.

- The government administers a PAYGO SS system htat collectes payrooll taxes $\tau_{ss,t}$ and pays benefits $p_{t,j}(\gamma, s)$ which depends on the wages a household has earned during her working years, an dthus on her characteristics (γ, s) a well as on how many periods the household is retired (which, given today's date t can be inferred from the current age j of the HH). The current SS system will be calibrated to incorporate its progressive beneift through the function $p_{t,j}(\gamma, s)$. The introduction of this progressivity adds a lifecycle saving motive and helps generate a more plausible wealth distributios.

  Since part of the labor income that is paid by the employer as SS contribution is not subject to income taxes, taxable labor income equals $(1 - 0.5\tau_{ss,t})$ per dollar of labor income earned, that is,

$$y_t = (1 - 0.5\tau_{ss,t})w_{t,s}\varepsilon_{j,t}\gamma\eta l$$

1. Newborn individuals are economically inactive but affect parental utility until they form a new HH at age j_a.
2. Initial state variables when a new HH forms are (a) age $j = j_a$, (b) parental education s_p, (c) parental productivity γ_p, and (d) own education $s = n$.

   Then, an ability level $e \sim \pi(e|s_p, \gamma_p)$ is drawn.

   Then, parents decide on the inter-vivos transfer b, which are transferred within the period are constitute the intial endowment of assets a.

   Then, initial idiosyncratic labor productivity η is drawn according to Π_n.

   Thus the state of a HH prior to the college decision is

$$z = \left(j_a, e, s = n, \eta, a = \frac{b}{1 + (1 - \tau_k r}\right)$$

3. Given state z, at age j_a the college decision is made. If a HH goes to collete at age j_a, her education switches to $s = c$ at that age.

   Then HHs draw their labor productivity fixed effect γ from the education and ability-contingent distribution $\pi(\gamma|e, s)$.

4. At age j_a, but after the college decision has been made, the HH problem differs between non-college and college HHs since the latter need to spend time and resources on college. A HH that goes to college but works part time does so for non-college wages:

$$w_{t,n}\varepsilon_{j,n}\gamma\eta$$

   where η is drawn from the non-college distribution. Note that γ is fixed. At the end of the college period j_a the idiosyncratic shock η of college-bound HHs is re-drawn from the college distributionΠ_c and now evolves according to $\pi_c(\eta'|\eta)$ for those with $s = c$.

   Furthermore, college-educated HHs draw their fixed effect from the distribuion $\pi(\gamma|c, e)$ prior to entering the labor market.

5. Ages $j_a+1,...,j_f-1$: Between age of j_f-1 and j_f the decision problem changes because children now enter the utility function and households maximize over per capita consumption $\frac{c_j}{1+\xi f}$.
6. Ages $j_a+j_f,...,j_a+j_f-1$: Between age of j_a+j_f-1 and j_a+j_f the decision problem changes again since at atge j_a+j_f children leave the household and the decision about the inter-vivos transfer b is made and lifetime utilities of children enter the continuation utility of parents.
7. Age j_f. HHs make transfers b to their children conditional on observing the skils e of their children.
8. Age $j_a+j_f+1,...,j_\tau-1$. Only utility from own consupmtion and leisure entres the lifetime utility at these ages.
9. Age $j=j_\tau,...,J$. HHs are now in retirement and only earn income from capital and from SS benefits $p_{t,j}(e,s)$.

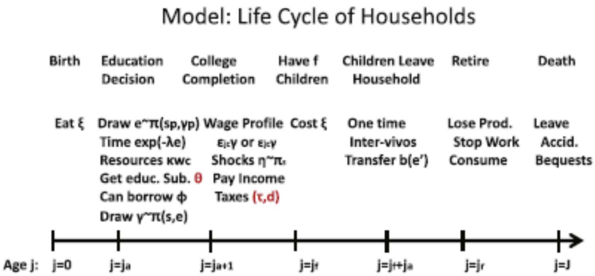


Fig. 1. Time line in the model. *Note*: This figure summarize the key life cycle events and decisions.

(1) *Child at* $j = 0, ..., j_a - 1$: No economic decisions. They live with their parents and consume resources.

(2) *Education decision at* j_a: Before HHs make the education decision HHs draw ability e, their initial η and receive inter-vivos transfers b. The education decision solves,

$$\mathbf{1}_{s,t}(e,\eta,b) = \begin{cases} 1 & \textit{if} \quad V_t\left(j_a, s=c, \eta, \frac{b}{1+(1-\tau_k)r}\right) > V_t\left(j_a, e, s=n, \eta, \frac{b}{(1+(1-\tau_k)r)}\right) \\ 0 & \textit{otherwise}, \end{cases}$$

where the lifetime utility at age $j = j_a$ (conditional on having chosen but not completed education $s \in \{n, c\}$) is,

$$V_t\left(j_a, e, s, \eta, \frac{b}{1+(1-\tau_k)r}\right) = \sum_{\gamma \in \Gamma_s} \pi_s(\gamma|e) V_t\left(j_a, \gamma, e, s, \eta, \frac{b}{1+(1-\tau_k)r}\right)$$

where

$$V_t\left(j_a, \gamma, e, s, \eta, \frac{b}{1+(1-\tau_k)r}\right)$$

is the value of function at age j_a after the fixed effect γ has been drawn from $\pi_s(\gamma|e)$ is defined below.

(3) Problem at $j = j_a$: After having made the education decision at age $j = j_a$ and having drawn the fixed effect γ HHs choose how much to work, how much to consume and save.

The problem is different between college-bound and non-college bound HHs. HHs first draw the fixed effect γ from distribution $\pi_s(\gamma|e)$ and then solve

$$V_t(j, \gamma, e, s, \eta, a) =$$

$$\max_{c, l \in [0, 1 - \mathbf{1}_S \xi(e)], a' \geq -\mathbf{1}_S \underline{A}_{j,t}} \left\{ u(c, 1 - l - \mathbf{1}_S \xi(e)) + \beta \varphi_j \sum_{\eta'} \pi_s(\eta'|\eta) V_{t+1}(j+1, \gamma, s, \eta', a') \right\}$$

subject to the budget constraint,

$$(1 + \tau_{c,t})c + a' + \mathbf{1}_S(1 - \theta_t - \theta_{pr})\kappa w_{t,c} + T_t(y_t) = (1 + (1 - \tau_{K,t})r_t)(a + Tr_{t,j}) + (1 - \tau_{ss,t})w_{t,n}\varepsilon_{j,n}\gamma\eta l$$

where $y_t = (1 - 0.5\tau_{ss,t})w_{t,n}\varepsilon_{j,n}\gamma\eta l$.

Note that e is a redundant state for non-college bound HHs at age j_a, but not for HHs going to college as it affects the time loss through $\xi(e)$. It does become a redundant state at age $j_a + 1$ as its effects are embedded in γ in both the non-college and college cases and that is why it does not appaer in the value function on the right hand side of the Bellman equation above.

(4) Problem at $j = j_a + 1, ..., j_f - 1$: This is the period after education has been completed and before children arrive. The problem reads as,

$$V_t(j, \gamma, s, \eta, a) =$$

$$\max_{c, l \in [0,1], a' \geq -\mathbf{1}_S \underline{A}_{j,t}} \left\{ u(c, 1-l) + \beta \varphi_j \sum_{\eta'} \pi_s(\eta'|\eta) V_{t+1}(j+1, \gamma, s, \eta', a') \right\}$$

subject to the budget constraint,

$$(1 + \tau_{c,t})c + a' + T_t(y_t) = (1 + (1 - \tau_{K,t})r_t)(a + Tr_{t,j}) + (1 - \tau_{ss,t})w_{t,n}\varepsilon_{j,n}\gamma\eta l$$

where $y_t = (1 - 0.5\tau_{ss,t})w_{t,n}\varepsilon_{j,n}\gamma\eta l$.

Note again that e is not a state variable. At this point all the effect of e occurs through γ.

(5) Problem at $j = j_f, ..., j_f + (j_a - 1)$: This is the period where children live with adults and consume resources. The problem reads as,

$$V_t(j, \gamma, s, \eta, a) =$$

$$\max_{c,l \in [0,1], a' \geq -\mathbf{1}_S \underline{A}_{j,t}} \left\{ u\left(\frac{c}{1+\zeta f}, 1-l\right) + \beta \varphi_j \sum_{\eta'} \pi_s(\eta'|\eta) V_{t+1}(j+1, \gamma, s, \eta', a') \right\}$$

subject to the budget constraint,

$$(1+\tau_{c,t})c + a' + T_t(y_t) = (1 + (1-\tau_{K,t})r_t)(a + Tr_{t,j}) + (1-\tau_{ss,t})w_{t,n}\varepsilon_{j,n}\gamma\eta l$$

where $y_t = (1 - 0.5\tau_{ss,t})w_{t,n}\varepsilon_{j,n}\gamma\eta l$.

Note that the way children consume resource is through ζ, the adult equivalence parameter.

(6) Problem at $j = j_f + j_a$: This is the period at which children leave home, parents give them an inter-vivos transfer b and the children's lifetime utility entres that of their parents. The dynamic problem becomes,

$$V_t(j, \gamma, s, \eta, a) =$$

$$\max_{\substack{c(e'), l(e') \in [0,1], \\ b(e') \geq 0, a'(e') \geq -\mathbf{1}_S \underline{A}_{j,t}}} \sum_{e'} \pi_s(e'|\gamma) \Bigg\{ u\left(c(e'), 1 - l(e')\right) + \beta \varphi_j \sum_{\eta'} \pi_s(\eta'|\eta) V_{t+1}(j+1, \gamma, s, \eta', a'(e')) +$$

$$\nu \sum_{\eta'} \Pi_n(\eta') \max \left[V_t\left(j_a, e', n, \eta', \frac{b(e')}{1 + (1-\tau_K)r}\right), V_t\left(j_a, e', c, \eta', \frac{b(e')}{1 + (1-\tau_K)r}\right)\right] \Bigg\}$$

subject to the budget constraint,

$$(1 + \tau_{c,t})c(e') + a'(e') + b(e')f + T_t(y_t) = (1 + (1 - \tau_{K,t})r_t)(a + Tr_{t,j}) + (1 - \tau_{ss,t})w_{t,n}\varepsilon_{j,n}\gamma\eta l(e')$$

where $y_t = (1 - 0.5\tau_{ss,t})w_{t,n}\varepsilon_{j,n}\gamma\eta l(e')$.

Note that all parental choices, including the transfer amount b, are a function of e' which is the innate ability of children that affects the children value of γ and their ability to complete college. Also note that all children in the HH are identical.

Since parents do not observe the initial labor productivity η of their children (which is drawn from the invariant non-college distribution $\Pi_n(\eta)$), parental choices are not contingent on it and expectations over η are taken in the Bellman equation of paretns over the lifetime utility of their children. In particular, note that we dont let the transfers b be a function of the initial labor productivity η of their children, since making b contingent on η permits parents to implicitly provide better insurance against η-risk.

Note that children will decide to go to college or not and parents know the chiildren problem (identical to the one parents faced when they were children). In any case, parents can influence (through b) the education decision of their children.

(7) Problem at $j = j_f + j_a + 1, ..., j_r - 1$: This is the period where children have left the HHs and before retirement. The dynamic problem is identical as that in item (4) for ages $ja + 1, ..., j_f$. Note that there is a discontinuity in the value function along the age dimension from age $j_f + j_a$ to age $j_f + j_a + 1$ because the lifetime utility of the child does no longer enter parental utility after age $j_f + j_a + 1$.

(8) Problem at $j = j_r, ..., J$: This is the last stage over the life cycle, retirement. In retirement HHs have no labor income (and consequently no labor income risk).

Thus the maximization problem is given by,

$$V_t(j, \gamma, s, a) = \max_{c, a' \geq 0} \left\{ u(c, 1) + \beta \varphi_j V_{t+1}(j+1, \gamma, s, a') \right\}$$

subject to the budget constraint,

$$(1 + \tau_{c,t})c + a' = (1 + (1 - \tau_{K,t})r_t)(a + Tr_{t,j}) + p_{t,j}(\gamma, s)$$

where the SS benefits $p_{t,j}(\gamma, s)$ are a function of lifetime earnings and hence of (γ, s) as well as on how many periods the HH is retired (which, given today's date t can beinferred from the current age j).

Note also that retired HHs are not allowed to be indebted.

- Let $\Phi_{t,j}(\gamma, s, \eta, a)$ denote the share of a,gents, at time t of age j with characteristics (γ, s, η, a). For each t and j we have $\int d\Phi_{t,j} = 1$.

  Recall that for age j_a and $s = c$ the state space also includes the ability e of the HH, but not the fixed effect γ. To simplify notation we keep distinction implicit whenever there is no room for confusion.

- Let $a \in \mathbf{R}_+$, $\eta \in \mathbf{E} = \{\eta_1, ..., \eta_n\}$, $\gamma \in \mathbf{I} = \{1, ..., M\}$, $j \in \mathbf{J} = \{1, ...J\}$, $s \in \mathbf{S} = \{n, c\}$ and let the $\mathbf{U} = \mathbf{R}_+ \times \mathbf{E} \times \mathbf{I} \times \mathbf{J} \times \mathbf{S}$.

- Let $\mathbf{B}(\mathbf{R}_+)$ be the Borel σ-algebra of $\mathbf{R}_+$ and $\mathbf{P}(\mathbf{E})$, $\mathbf{P}(\mathbf{I})$, $\mathbf{P}(\mathbf{J})$ and $\mathbf{P}(\mathbf{S})$ the power sets of $\mathbf{E}, \mathbf{I}$, $\mathbf{J}$ and $\mathbf{S}$, respectively.

- Let $\mathbf{M}$ be the set of all finite measures over the measurable space $\mathbf{U}$,

$$\mathbf{B}(\mathbf{R}_+) \times \mathbf{P}(\mathbf{E}) \times \mathbf{P}(\mathbf{I}) \times \mathbf{P}(\mathbf{J}) \times \mathbf{P}(\mathbf{S})$$

Definition of Recursive Competitive Equilibrium (RCE)

Given a stream of government spending $\{G_t\}_{t=0}^{\infty}$, consumption taxes $\{\tau_{c,t}\}_{t=0}^{\infty}$, and initial aggregate captial K_0, initial government debt level B_0 and initial measures $\{\Phi_{0,j}\}_{j=0}^{J}$ of HHs, a RCE is a sequences of

- value and policy function $\{V_t, a'_t, c_t, l_t, \mathbf{1}_{s,t}, b_t\}_{t=0}^{\infty}$,
- production plans $\{Y_t, K_t, L_{n,t}, L_{c,t}\}_{t=0}^{\infty}$,
- prices $\{w_{t,n}, w_{t,c}, r_t\}_{t=0}^{\infty}$,
- tax policies, $\{T_t, \tau_{c,t}, \tau_{K,t}\}_{t=0}^{\infty}$,
- education policies, $\{\theta_t\}_{t=0}^{\infty}$,
- SS policies, $\{\tau_{ss,t}, p_{t,j}\}_{t=0}^{\infty}$,
- government debt levels $\{B_t\}_{t=0}^{\infty}$,
- transfers $\{Tr_{t,j}\}_{t=0,j}^{\infty}$ and
- measures $\{\Phi_{t,j}\}_{t=0}^{\infty} \in \mathbf{M}$ such that:

(1) Given prices, policies, transfers, $\{V_t\}$ solves the Bellman equations in the "Recursive Problems of HHs" and $\{a'_t, c_t, l_t, \mathbf{1}_{s,t}, b_t\}$ are the associated policy functions.

(2) Wages and interest rates are competitive and identical to respective marginal products.

(3) Bequests/transfers are given for all $j \geq j_f$ by

$$N_{t+1,j-j_f+1} Tr_{t+1,j-j_f+1} + \frac{N_{t+1,j-j_f+1}}{\sum_{\iota=j_f}^{J} N_{t+1,j-j_f+1}} \frac{PE_{t+1}}{1+(1-\tau_{K,t+1})r_{t+1}} = N_{t,j} \int (1-\varphi_j) a'_t(j,\gamma,s,\eta$$

where PE_{t+1} are private aggregate education subsidies (from bequests) given by:

$$PE_{t+1} = \theta_{pr} \kappa w_{t+1,c} N_{t+1,j_a} \int_{\{(e,s,\eta,a):s=c\}} d\Phi_{t+1,j_a}$$

(4) Government policies satisfy the government budget constraints:

(a) PAYGO SS:

$$\tau_{ss,t} \sum_{s} w_{t,s} L_{t,s} = \sum_{j=j_r}^{J} N_{t,j} \int p_{t,j}(\gamma, s) d\Phi_{t,j}$$

(b) Government budget:

$$G_t + E_t + (1 + r_r) B_t = B_{t+1} + \sum_{j} N_{t,j} \int T_t(y_t) d\Phi_{t,j} + \tau_{K,t} r_t (K_t + B_t) + \tau_{c,t} C_t$$

where y_t is taxable income defined above, and aggregate consumption is

$$C_t = \sum_{j} N_{t,j} \int c_t(j, \gamma, s, \eta, a) d\Phi_{t,j}$$

and government education expenditures are given by

$$E_t = \theta_t \kappa w_{t,c} N_{t,j_a} \int_{(e,\gamma,s,\eta,a):s=c} d\Phi_{t,j_a}.$$

(5) Markets clear in all periods t,

$$L_{t,s} = \sum_j N_{t,j} \int \varepsilon_{j,s}\gamma\eta l_t(j,\gamma,s,\eta,a)d\Phi_{t,j} \text{ for } s \in \{n,c\}$$

$$K_{t+1} + B_{t+1} = \sum_j N_{t,j} \int a'_t(j,\gamma,s,\eta,a)d\Phi_{t,j}$$

$$K_{t+1} =_t +(1-\delta)K_t - C_t - CE_t - G_t - E_t$$

where Y_t is given by the production function. Also, the aggregate private spending in education is:

$$CE_t = (1-\theta_k)\kappa w_{t,c} N_{t,j_a} \int_{(e,s,\eta,a):s=c} d\Phi_{t,j_a}$$

(6) The law of moton for the aggregate state (the distribution) $\Phi_{t+1,j+1} = H_{t,j}(\Phi_{t,j})$ where $H_{t,j}$ is the law of moting induced by the exogenous population dynamics, the exogenous Markov process for labor productivity, and the endogenous asset accumulation, education, and transfer decisions, respectively, $a'_t, \mathbf{1}_{s,t}, b_t$.

Define the Markov transition funciton at time t for age j as,

$$Q_{t,j}((\gamma, s, \eta, a), (\Gamma \times \mathcal{S} \times \mathcal{E} \times \mathcal{A})) =$$
$$\begin{cases} \sum_{\eta' \in \mathcal{E}} \pi_s(\eta'|\eta) & \text{if} \gamma \in \Gamma, s \in \mathcal{S}, \text{and } a'_t(j, \gamma, s, \eta, a) \in \mathcal{A} \\ 0 & \textit{else} \end{cases}$$

That is, the probability of going from state (γ, s, η, a) into a set of states $(\Gamma \times \mathcal{S} \times \mathcal{E} \times \mathcal{A})$ tomorrow is zero if that set does not include the current education level and productivity type, and $\mathcal{A}$ does not include the optimal asset choice.[5] If it does, then the transition probability is purely governened by the stochastic shock process for η.

[5] There is one exception at age $j = j_c$ college-educated HHs redraw their income shock η and draw their fixed effect according to $\pi(\gamma'|c, e)$. For this group therefore the transition function at that age reads as

$$Q_{t,j}((e, c, \eta, a), (\Gamma \times \{c\} \times \mathcal{E} \times \mathcal{A})) =$$
$$\begin{cases} \sum_{\gamma' \in \Gamma} \sum_{\eta' \in \mathcal{E}} \pi(\gamma'|c, e)\Pi_c(\eta') & \text{if } a'_t(j, e, c, \eta, a) \in \mathcal{A} \\ 0 & \textit{else} \end{cases}$$

The age-dependent measure are given, for all $j \geq j_a$ by

$$\Phi_{t+1,j+1}(\Gamma \times \mathcal{S} \times \mathcal{E} \times \mathcal{A}) = \int Q_{t,j}(., (\Gamma \times \mathcal{S} \times \mathcal{E} \times \mathcal{A}) d\Phi_{t,j}$$

The initial measure over types at age $j = j_a$ (after the college decision has been made) is more complicated. Part of the complexity is that at age j_a the individual state space includes ability e which then becomes a redundant state variable.[6] HHs start with assets from transfers from their parents determined by the inter-vivos transfer function b_t, draw initial mean reverting productivity according to $\Pi_n(\eta')$, determine educaiton according to the index function $\mathbf{1}_{s,t}$ evaluated at their draw e', η' and the optimal bequests of the parents and draw the fixed effect according to $\pi(\gamma'|s, e')$:

For non-college HHs:

$$\Phi_{t+1,j=j_a}(\{e'\} \times \{\gamma'\} \times \{n\} \times \mathcal{A}) = \Pi_n(\eta') \sum_s \sum_\gamma \pi(\gamma'|n, e)\pi(e'|s, \gamma) \int (1 - \mathbf{1}_{s,t}(e', \eta', b_t(\gamma, s, \eta, a, e'))$$
$$\cdot \mathbf{1}_{\frac{b_t(\gamma,s,\eta,a;e')}{1+(1-\tau_{K,t})r_t} \in \mathcal{A}} \Phi_{t,j_f+j_a}(\{\gamma\} \times \{s\} \times d\eta \times da)$$

For college HHs:

$$\Phi_{t+1,j=j_a}(\{e'\} \times \{\gamma'\} \times \{c\} \times \mathcal{A}) = \Pi_n(\eta') \sum_s \sum_\gamma \pi(\gamma'|c, e)\pi(e'|s, \gamma) \int \mathbf{1}_{s,t}(e', \eta', b_t(\gamma, s, \eta, a, e')$$
$$\cdot \mathbf{1}_{\frac{b_t(\gamma,s,\eta,a;e')}{1+(1-\tau_{K,t})r_t} \in \mathcal{A}} \Phi_{t,j_f+j_a}(\{\gamma\} \times \{s\} \times d\eta \times da)$$

[6] Thus the measures for age j_a will be defined over e as well, and it is understood that the transition function $Q_{t,j}$ from age j_a to age $j_a + 1$ (and only at this age) has as first argument (e, γ, s, η, a).

Definition of Recursive Stationary Equilibrium

A stationary EQ is a competitive EQ in which individual functions and all aggregate variables are constant over time. (grow at a contant rate?)

The SWF is utilitarian for people initially alive,

$$SWF(\mathcal{T}) = \sum_j N_{1,j} \int V_1(j, \gamma, s, \eta, a; \mathcal{T}) d\Phi_{1,j},$$

where $V_1(.;\mathcal{T})$ is the value function in the first period of the transition induced by new tax system $(\mathcal{T})$ and $\Phi_1 = \Phi_0$ is the initial distribution of HHs in the stationary EQ. under the status quo policy.

In our optimal policy analysis we hold constant the capital income and consumption tax rate as well as the pension contribution rate (and the borrowing constraints for student loans $\underline{A}_{j,t}$) and optimize over labor income taxes and education subsidy rates.

Therefore, given initial conditions (K_0, B_0), consumption taxes, capital income taxes, a pension system, borrowing constraints and a cross-section of HHs Φ_0 determined by a stationary (to be calibrated policy $T_0, \theta_0, d_0, b_0 = \frac{B_0}{Y_0}$), we define the optimal tax reform as the sequence $\mathcal{T}^* = \{T_t, \theta_t, d_t, B_t\}_{t=1}^{\infty}$ that maximizes the SWF, i.e., that solves

$$\mathcal{T}^* \in \arg\max_{\mathcal{T} in \mathbf{T}}$$

Here $\mathbf{T}$ is the set of policies for which an associated competitive EQ. exists.

Unfortunately the set $\mathbf{T}$ is too large a policy space to optimize over. Our objective here is to characterize the optimal one-time policy reform, by restricting the sequences that are being optimized over to,

$$T_t = T_1$$
$$\theta_t = \theta_1$$
$$d_t = d_1$$

for all $t \geq 1$.

Since all admissible policies (T_1, θ_1, d_1) must lie in $\mathbf{T}$, from the definiton of EQ. there must be an associated sequence of $\{B_t\}$ such that the government budget constraint is satisfied in every period. This imposes further restrictions on the set of possible triples (T_1, θ_1, d_1) over which the optimization of the SWF is carried out, i.e., the path of government debt has to be consistent with initial conditions and a long-run stationary EQ.

Calibration

Demographics

- Each period is four years.
- Survival probabilities from Social Security Administration life tables.
- Total fertility rate $f = 1.14$ such that a mother on average has $2f = 2.28$ children.
- Children are born with age 0 and form HHs at bioogical age 18.
- We discard the first two years of childhood: $j_a = \frac{18-2}{4} = 4$.
- HHs require 4 actual years (i.e., 1 model period) to complete college, so they exit college at age $j_c = j_a + 1 = 5$.
- They have children at biological age 30, which is model age $j_f = 7$.
- Retirement occurs at biological age 66 (age bin 62-65), hence $j_r = 16$.
- HHs are alive at most until bin 98-101, and accordingly $J = 24$.

- HH of age j with education $s \in \{n, c\}$, fixed effect $\gamma \in \{\gamma_{l,s}, \gamma_{h,s}\}$ and idiosyncratic shock η earns a wage of:

$$w_{t,s}\varepsilon_{j,s}\gamma$$

where w_s is the skill-specific wage.

- First, the authors estimate the deterministic $\varepsilon_{j,s}$ from the PSID data, and normalize $\varepsilon_{j_c,n} = 1$, i.e., noncollege profile at age j_c to one.

  They rescale the college profile at age j_c, i.e, $\varepsilon_{j_c,c}$ such that the average college wage premium in the model is 80% as in the US in the late 2000s.

- Second, choose the stochastic mean reverting component of wages η as two state Markov chain with skill-specific staets for log-wages $\{-\sigma_s, \sigma_s\}$ and transition matrix,

$$\Pi = \begin{pmatrix} \pi_s & 1-\pi_s \\ 1-\pi_s & \pi_s \end{pmatrix}$$

To parameterize this Markov chain we first estimate the following skill-specific income process (separately by skills):

$$\log w_t = \alpha + z_t$$
$$z_t = \rho z_{t-1} + \eta_t$$

where α is an individual-specific fixed effect that is assumed to be normally distributed with cross-sec variance σ_α^2.

The authors find more persistence in college $\rho_c = .969$ than in non-college $\rho_n = .928$, lower risk in college $\sigma_{\eta,c}^2 = .010$ than in noncollege $\sigma_{\eta,n}^2 = .0192$, and lower dispersion of fixed effects within college individuals $\sigma_{\alpha,c}^2 = .0474$ than within noncollege $\sigma_{\alpha,n}^2 = .0644$.

For each skill group we choose two numbers π_s, σ_s such that the 2-state Markov chain for wages has the same persistence and conditional variance as the continuous income process above. This implies the following Markov chain,

	π_s	σ_s	ε_s
College	.941	.191	$\{.811, 1.188\}$
Noncollege	.871	.250	$\{.755, 1.244\}$

- Third, we choose the fixed component of wages $\gamma \in \{\gamma_{l,s}, \gamma_{h,s}\}$ drawn from an ability-dependent distribution $\pi(\gamma|s, e)$.

  We calibrate the parameters governing this γ so that our model matches the selected wage observations form the data. We assume that,

  $$\pi(\gamma = \gamma_h|e, s = c) = e$$
  $$\pi(\gamma = \gamma_h|e, s = n) = ve$$

  where also note that $\pi(\gamma = \gamma_l|e, s) = 1 - \pi(\gamma = \gamma_h|e, s)$. Here v is a parameter. The distribution of e is discussed below.

  The parameters $\gamma_{l,n}, \gamma_{h,n}, \gamma_{l,n}, \gamma_{h,n}, v$ are chosen jointly such that the stationary EQ. of the status quo economy matches these targets:

  - Normalizations: The average γ is equal to one for each $s \in \{n, c\}$. [2 targets]
  - The estimated variance of the fixed effect for both educatin groups σ_α^2 estimated earlier. Note that these variances in the model are determined, for each s, by the spread between $\gamma_{l,s}$ and $\gamma_{h,s}$ (as well as the probability of drawing them in next bullet) [2 targets]
  - The skill wage premimum of marginal households (those close to indifferent between attending and not attending college under the benchmark policy) as measured by Findeisen and Sachs (2014). In this model this statistic is primarily governed by parameter v [1 target].

Calibration

Technology

- The elasticity of substitution between college and noncollege is set to $\frac{1}{1-\rho} = 1.4$, from Katz and Murphy (1992). They will also study versions of th emodel where the skill groups are assumed perfect substitutes in which case a change in the relative supply of college-educated labor will have no impact on its relative price.
- The capital share is set to $\theta = 1/3$.
- They choose depreciation rate δ to match 7.55% and a yearly interest rate of 5.4%. The capital-oputput ratio (or equivalently, the real interest rate) will be attained by appropriate calibration of the preference parameters, i.e., the discount factor β.

Calibration

Government Policy

The parameters to be chosen are $(b, gy, \tau_c, \tau_k, \tau_s s, \tau_l, d)$:

- Pick $b = 0.60$ to match a government debt to output ratio of 60%.
- Pick $gy = 0.17$ to match a government consumption (net of college education expenditure) to output ratio of 17%.
- Consumption taxes can be estimated from NIPA data as in Mendoza et al. (1994) who find $\tau_c \approx 0.05$.
- For capital income tax rate, $\tau_k = 0.283$ from Chari and Kehoe (2006) for the early 2000s.
- The payroll tax $\tau_{ss} = 0.124$ is chosen to match the curren tSS payroll tax (excluding Medicare).
- Model SS benefits $p_{t,j}(e, s)$ as a concave function of average wages earned during a HHs working life, in order to obtain a reasonalby accurate approximation to th ecurrent progressive US benefit formula, but wihtout the need to add a continuous state variable to the model (see Appendix 3.1).

- We calibrate the labor income tax decuciton to match the sum of standard decutions and exemptions form the US income tax code. Both median income and the size of the standard exemption and deduction varies by HH size and type, but their ratio is roughly constant at 35%.

  Thus we calibrate the deduction and exemption in the benchmark economy to 35% of the (endogenous) median income in the model. That is, we choose the policy parameter d such that,

$$d \frac{\frac{Y}{N}}{med(y^{gross})} = 35\%,$$

  where Y/N is output per capita in the model.

  Finally the marginal tax rate on labor income $\tau_l = 0.275$ is chosen to balance the government budget.

In 2009, according to the US Census Bureau Statistical Abstract of the US 2012 (Table 692), median HH money income of a HH of 4 members was \$73,071, relative to a sum of standard deduction (\$11,400) and 4 times the exemption (4*3,650) of \$26,000. The corresponding numbers for a two person HH are \$53,676 and \$18,700 and for a single person of \$26,080 and 10,350. The corresponding ratios are $d = 25.6\%, d = 34.8\%$ and $d = 35.9\%$.

We approximate money income in the model as,

$$y^{gross}(0, j, \gamma, s, \eta) = (a + Tr)r + (1 - 0.5\tau_{ss,1})w_{0,s}\varepsilon_{j,s}\gamma\eta l$$

Comparison with Heathcote, Storesletten and Violante (2014):

- Alternatively we could have chosen (d, τ_l) such that the model matches the relative dispersionmeasured, e.g., by the variance of logsof pre-tax and disposable income.
- Using the tax function estimates reported in Heathcote et al.(2014) gives an empirical target of 0.72, where as our model implies a value of 0.86, suggesting that our simple benchmark tax system misses progressivity at higher income.
- Matching this statistic would have required $(d = 0.73, \tau_l = 0.32)$ and thus a more progressive taxy system. Note that this alternative choice of the status quo does not affect, abstracting from recalibration of the other parameters, the determination of the optimal tax system.

The bequest parameter ν is chosen so that in EQ total transfers-i.e., the sum of inter-vivo transfers and accidnet bequests-in the economy account for 1.7% of wealth as in the 1986 SCF (summarized by Gale and Scholz, 1994).

The period utility function is specified as,

$$u(c, l) = \frac{\left[c^{\mu}(1 - \mathbf{1}_s \xi(e) - l)^{1-\mu}\right]^{1-\sigma}}{1 - \sigma}$$

The authors a priori choose $\sigma = 4$ and then determine the time discount factor β and the weight on leisure μ in the utility function such that in the benchmark model the capital-output ratio is 2.5 and HHs on average work $1/3$ of their time.

- These preferences imply a Frisch elasticity of labor supply of $\frac{1-\mu(1-\sigma)}{\sigma}\frac{1-l}{l}$ and with an average labor supply of $l = 1/3$ one could be worried that the Frisch labor supply elasticity, which, given the parameter estiamtes will be arond 1 for most HHs, is implausibly high. But note that this elasticity of labor supply of entire housholds, not that of white prime age males on which many lower empirical estimates are based. Also, the average Firsch elasticity for HHs in the age bin 24-54 in our mode is at 0.6 which we view as a conservative estimate.
- The CRRA with this formulation equals $\sigma\mu + 1 - \mu \approx 2$.

The parameters:

- The resource cost for college educaiton κ chosen to match the total average yearly cost of going to college, as a fraction of GDP per capita, $\frac{\kappa w_c}{Y/N}$.
- The share of expenses borne by the government and private sources, θ and $\theta_p r$ chosen to match the cost net of subsidies $(1-\theta-\theta_{pr})\frac{\kappa w_c}{Y/N}$.

Ionescu and Simpson (2014) report an average net price (tuition, fees, room and board net of grants and education subsidies) for a four year college from 2003/04 to 2007/08 to be \$58,654 and for a two year college of \$20,535. They also report a college completion success rate of 67% for four year college and 33% for two year college. Thus, the average net cost of tuition and fees for one year of college is:

$$0.67\frac{\$58,654}{4}+0.33\frac{\$20,535}{2}=\$13,213$$

- Then,

$$(1-\theta-\theta_{pr})\frac{\kappa w_c}{Y/N}=\frac{\$13,213}{\$42,684}=0.31.$$

  where $\frac{Y}{N}$ in constant 2005 dollars is \$42,684.

- Education at a glance (OECD 2012, Table B3.2b) reports that the share of college eduction expenditures born by public and private subsidies is $\theta=0.388$ and $\theta_{pr}=0.166$, so that,

$$\frac{\kappa w_c}{Y/N}=\frac{0.31}{1-\theta-\theta_{pr}}=0.694.$$

  And the cost parameter κ is calibrated so that the EQ of the benchmark model has $\frac{\kappa w_c}{Y/N}=0.694$

Newly formed HHS draw their ability from a distribution $\pi(e|s_p, \gamma_p)$ whose mean $\mu(s_p, \gamma_p)$ depends on the education level s_p and permanent labor productivity γ_p of their parents; recall that the distribution of thelatter is in turn determined by the parental ability e_p.

We interpret $e \in [0, 1]$ as a basic ability to succeed in college and in the labor market.

We assume that e follows a normal distribution with mean $\mu(s_p, \gamma_p)$ and standard deviation σ_e, truncated to the unit interval, that is, for all $e_p \in [0, 1]$,

$$\pi(e|s_p, \gamma_p) = \frac{\psi\left(\frac{e - \mu(s_p, \gamma_p)}{\sigma_e}\right)}{\Psi\left(\frac{1 - \mu(s_p, \gamma_p)}{\sigma_e}\right) - \Psi\left(\frac{-\mu(s_p, \gamma_p)}{\sigma_e}\right)}$$

where ψ is the pdf of a stanard nomal and Ψ is the cdf of a standard normal. Note that both numerator and denominator is dependent on $\mu(s_p, \gamma_p)$.

By assuming that,

$$\mu(s_p, \gamma_p) = \begin{cases} 0.5 - \chi + \zeta \mathbf{1}_{s_p = c} & \textit{for} \quad \gamma_p = \gamma_{l,s_p} \\ 0.5 + \chi + \zeta \mathbf{1}_{s_p = c} & \textit{for} \quad \gamma_p = \gamma_{h,s_p} \end{cases}$$

the distribution of ability is characterized by three parameters χ, ζ, σ_e where χ measures the impact of parental ability on children's ability, whereas ζ captures the importance of parental education.

The authors choose χ to fit the intergenerational persistence of earnings in the data and $\zeta > 0$ to match college completion rates conditoinal on parental education $s = c$ (that is, to match intergeneratinal persistence in education) and σ_e such that 95% of the probability mass of the e-distribution lies in the unit interval $e \in [0, 1]$.

This parameterization of the intergenerational ability transmisison gives both a role to the parental education and to parental ability (through their draws of γ) for shaping children's ability e. They then restrict e to take on a discrete set of $n_e = 31$ values that are evenly spaced in the unit interval.

Based on their ability e the time requirement for attending class and studying in college is given by the function,

$$\zeta(e) = exp(-\lambda e)$$

where $\lambda > 0$ is a parameter that governs the importance of ability e for the time (and htus utility) cost of going to college.

The authors calibrate λ to match the overall share of HHs completing college in the data form the National Educaiton Longitudinal Study (NELS).

- The borrowing constraints faced by agents pursuing a college degree allow such an agent to finance a fraction $\phi \in [0,1]$ all tuition bills with credit.
- We specify a constant (minimum) payment rp such that at the age of retirement all college loans are repaid:

$$\underline{A}_{j_a,t} = \phi(1-\theta-\theta_{pr})\kappa w_{t,c}$$

and for $j = j_a+1, ..., j_r$:

$$\underline{A}_{j,t} = (1+r_t)\underline{A}_{j-1,t-1} - rp$$

and rp is chosen such that the terminal condition $\underline{A}_{j,t} = 0$ is met.

  - The parameter ϕ determines how tight the borrowing constraint for college is.
  - The repayment rp is not a calibration parameter but an endogenously determined repayment amount that insures that HHs do not retire with outstanding student loans.

The maximum amount of publicly provided student loans for four years is given by \$27,000 for dependent undergraduate students and \$45,000 fo independent undergraduate students (the more relavant number given that our students are independent households).[7]

Relative to GDP per capita in 2008 of \$48,000, this given maximum debt constitutes 14% and 23.4% of GDP per capital.

Compare that to the 31% of total costs computed above, this indicates that indpeendent undergrad students can borrow at most approximately 75% of the cost of college, and thus we set $\psi = 0.75$. Again, note that this is an education policy parameter tha tis being held fixed in our opimtal policy analysis.

[7] 66% of students finishing four year colleges have debt, and conditional on havin debt the average amount is \$23,186 and the median amount is \$20,000.

- How HHs make their key economic decisions for a given policy.

  Life cycle profiles of consumption, asset and labor supply are consistent with those reported in the literature (Figure 1, where is it?)

- What is relevant is to explore how the optimal education decision is made, as a function of the initial characteristics of the HH. We want to understand this because policy can have a strong impact on this decision, thus it is important for understanding the optimality of the policy.

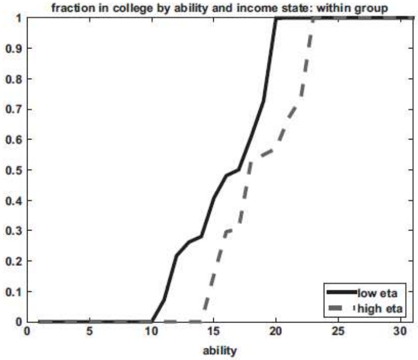


Fig. 2. Fraction of households deciding to go to college. *Note*: This figure plots the share of households in college age going to college, as function of their ability $e \in \{e_1, \ldots, e_{31}\}$. Each line represents an income shock $\eta \in \{\eta_l, \eta_h\}$.

- HHs at the time of college decision differ according to (e, η, b), that is, their ability to go to college e, their wages outside college (determined by the shock η), and their initial asset levels resulting from parental transfers b.

- Fig. 2 shows the share of HHs deciding to go to college, under the status quo policy, as a function of e both for HHs with low and high η realizations.

- All HHs with high abilities ($e \geq 22$) go to college, and none with very low ability ($e \leq 10$) do.

- For HHs in the middle of the ability distribution, their decision depends on the attractiveness of the option of working in the labor market: a larger share of HHs with lower opportunity costs (low η) attends college.

  Note that an interior share strictly between zero and one, conditional on η, indicates that heterogeneity in wealth b among the youngest cohort (which in utrn stems from wealth and thus transfer heterogeniety of their parents) is an important determinant of the college decision for those in the middle of the ability distribution ($e \in [e_{11}, e_{21}]$).

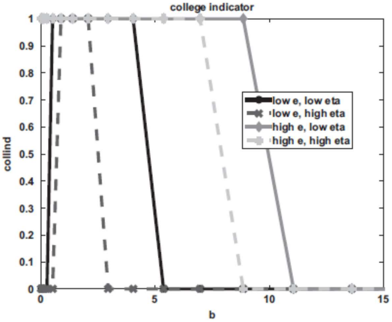


Fig. 3. College decision by initial assets. *Note:* This figure plots the optimal college decision rule against initial wealth b. A value of 0 stands for not attending college, a value of 1 for going to college. Each line represents an income shock $\eta \in \{\eta_l, \eta_h\}$ and ability level e. The low (high) ability group is group 17 (25) out of 31 groups.

- Lower-ability HHs go to college for a smaller range of initial assets than do high ability HHs.
- A higher non-college wage (high η) reduces college attendance.
- The effects of initial wealth on the college decision are non-monotone:
    - For HHs with low wealth (and sufficiently low e) the borrowing constraint is imporant. Although the government subsidizes college (38.8% in status quo) and HHs can borrow 75% of the remaining costs, HHs with very low wealth might still not be able to afford college. That is, either it is impossible for these HHs to keep positive consumption even by working full time while attending college, or the resulting very low level of consupmtion (and/or leisure) makes such a choice suboptimal.
    - As intial wealth increases, the borrowing constraint is relaxed and even theless able HHs decide to go to college.
    - Finally, for very rich HHs that expect to get a dominant share of their lifeitme income from capital income find it suboptimal to invest in college and bear the time and resource cost in exahnge for larger labor earnings after college (this is a income/wealth effect). The stationary asset transfer distribution puts essentially no mass on initial assets $b \geq 5$.

Results

Analysis of optimal policy transitions: The optimal policies

First we summarize the optimal fiscal constitution for our benchmark economy when both transitional dynamics and GE relative wage effects are taken into account.

Second, we aim at providing intuition for the results by

- Analyzing the optimal "steady state policy" in the absence of relative wage movements.
- Documenting the quantitative importance of transitional dynamics.

Table 3
Benchmark results: optimal policies and macroeconomic aggregates in the long run.

Var.	Status quo	Trans. dynamics		Steady state	
		Opt. pol.	Change	Opt. pol.	Change
η	27.55%	21.9%	−5.69% p	26.03%	−1.53% p
d	27.1%	6%	−21.1% p	10%	−17.1% p
θ	38.8%	150%	111.2% p	200%	161.2% p
Z	0.0714	0.0171	−76.05%	0.0281	−60.57%
$\eta \cdot Z$	0.0197	0.0037	−80.98%	0.0073	−62.76%
Y/N	0.2633	0.2847	8.15%	0.2813	6.85%
B/Y	60.16%	79.81%	19.65% p	60.16%	0% p
K/N	0.1731	0.1899	9.73%	0.1866	7.8%
L/N	2.2097	2.3782	7.63%	2.3508	6.39%
K/L	0.1097	0.1118	1.95%	0.1111	1.33%
w	0.1118	0.1123	0.49%	0.1122	0.44%
$\frac{w_c}{w_n}$	0.3801	0.2946	−22.49%	0.2732	−28.12%
r	5.37%	5.26%	−0.11% p	5.27%	−0.09% p
Hours	0.326	0.3392	1.32%	0.3329	0.7%
C/N	0.1628	0.1795	10.28%	0.1731	6.32%
Trans/assets	1.15%	1.01%	−0.15% p	0.78%	−0.37% p
College share	43.89%	54.37%	10.48% p	58.57%	14.68% p
$Gini(c)$	0.235	0.2078	−2.72 p	0.1996	−3.54 p
$Gini(h)$	0.12	0.1266	0.66 p	0.1243	0.43 p
$Gini(a)$	0.5558	0.5369	−1.89 p	0.5493	−0.65 p
CEV		3.4652%		2.7569%	

Note: This table summarizes optimal policies and the long-run change in macroeconomic aggregates, distributions and welfare for $\rho = 0.285$ (an elasticity of substitution between unskilled and skilled labor of 1.4).

- The optimal policy transition is obtained by:
  - A significantly larger education subsidy $\theta = 150\%$ than in status quo 38.8%.
  - A significanlty less progressive tax system: the deduction d falls from 27.1% in status qo to an optimal of 6%. Recall that $Z = d\frac{Y}{N}$ is the size of the labor income tax deduction and d measures the size of the deduction relative to income per capita. Note that τ_l goes down from 27.55% in status quo to an optimal value of 21.9%.
  - The welfare gains are substantial, in the order of 3.5% of lifetime consumption.

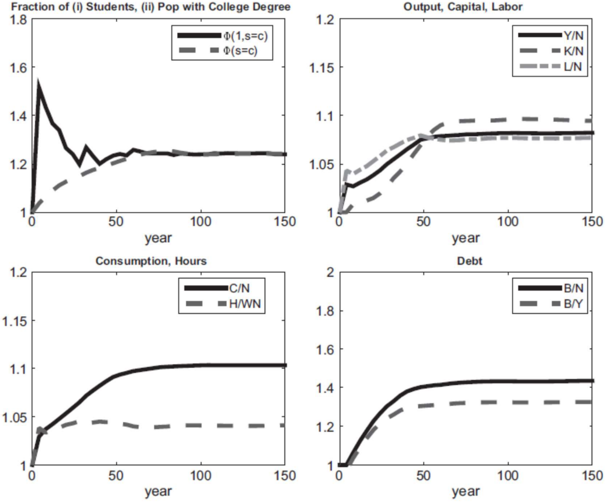


Fig. 4. Evolution of macroeconomic aggregates: imperfect substitutes. *Note:* This figure plots the transition paths of key macroeconomic aggregates against time. The elasticity of substitution between unskilled and skilled labor is 1.4 (ρ=0.285).

- Upper left panel of Fig. 4:
    - First, the increase in the college subsidy and the reduction in the marginal (and average) tax rates for high income earnes induces a college boom on impact.
    - Second, the share of the youngest cohort going to college rises from 44% in the intial steady state to 67% in the first period of the transition. The share of college bound youngsters is falling over time. This is due to the fall in the college wage premium over time (see Figure 5). Althogh at the new steady state the share of young cohorts going to college remains about 23$ (or about 10.5%) above its initial steady state.
    - Third, althoughthe share of the younges cohort going to college increases immediately by close to 60% on policy impact, it takes approximately two generations (roughly 60 years) untilthe overall skill distribution has reached a level close ot is new steadystate value. It is this sluggishy dyanmics of the skill and thus labor productivity distribution that a an anlysis restricted to long-run steady state would miss completely.

- Focusing on the long run steady state, the optimal policy reform increases output and consumptoin per capita (8.15% and 10.28%), see column 2 in Table 3, despite the fact that hours worked only increase by 1.32% (although note that labor efficiency units L/N increase much more by 7.63% on account of a more skilled workforce.

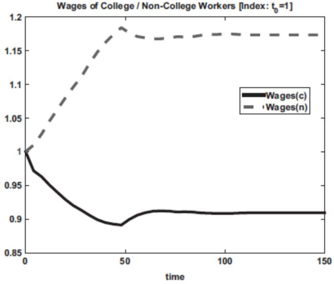


Fig. 5. Wages of college and non-college households: imperfect substitutes. *Note*: This figure plots the transition paths of (average) wages for college-educated and unskilled workers against time. The elasticity of substitution between unskilled and skilled labor is 1.4 (ρ=0.285).

The transition is slow...

- As the skill distribution improves only slowly and more and more on average more skilled cohorts enter the labor force over time, effective labor units supplied and output per captia increase slowly along the transition as well (upper right and left panel of Fig. 4)
- The same happens for consumption (lower left panel of Fig. 4).
- Averate hours worked also increase on impact by about 3% because of the decline in the marginal tax rate τ_l. This is true despite the fact that a larger share of the youngest age cohort now goes to college and thus withdraws temporarily from the labor force.
- Finally, given the tax and education policy and the inital debt level the sequence of government debt is determined from the period government budget constraints. The government finds it optimal (through the setting of τ_l, θ, d) to smooth the transitional costs of building up a higher human capital stock (i.e., high education policy costs and relatively lax reveneures due to low—relative to final steady state—econmic activity) by borrowing along the transition. Debt per capita is increasing more strongly than that since GDP is also increasing along the transition, although not at the same speed as deebt itself, which explains the increase in the debt to GDP ratio.

On inequality...

- Consumption and wealth inequality fall despite the reduction in the labor income tax progressivity. This is mainly due to the fact that in the final steady state there are significantly less consumption (and wealth) poor unskilled households.
- Furthermore, the college wage premium shrinks significantly aiding the decline in consumption inequality.

Overall, the policy reform increases educational attainments, raises output and consumption per capita both in the short and the long run on account of a more productive workforce and at the same time reduces consumption inequality on account of narrowin college wage premium, despite the fact that the tax system has become less redistributive.

We reproduce the Table 3 in Table 4 where we assume perfect substitutability across skill groups (which implies killing the GE effects, i.e., there is no effect on the skill wage premium). All other parameters remain constant.

- Comparing the optimal steady state policies reveals the importance of the GE wage effects for optimal policy design. Whereas the optimal education subsidiy is large in both cases, the degree of tax progressivity differs substantially.

  In the benchmark economy where redistribution across different ability e types and thus across education groups can be achieved through the GE wage effects induced by education subsidies. Thus, with imperfect substitution a strongly progressive tax system is not desirable: optimal $d = 10\%$, smaller relative to the benchmark of $d = 27\%$.

  If education policies stimulating college attendance do not affect the college wage premium as in the perfect substitutes case, instead a strongly progressive tax system with a deduction worth 31% of average income (and thus more progressive than the status quo) is optimal.

Results

- Note that with perfect substituion across skill groups the steady state optimal policy reduces hours worked in the steady state substantially (by 8.4%).
- However, absent the reduction in the college wage premium the education boom is so massive that labor input as well as output and consumption per capita rise substantially.
- There is also a reduction in consumption inequality which is achieved with different tools. In particular, without GE effects, registribution has to be achieved via the income tax code, as a deduction of $d = 31\%$ (relative to $d = 10\%$ in the benchmark) shows.
- This in turn has severe consequences for the importance of the transitional dynamics in optimal policy design, as we discuss next.

Table 4
Optimal policies and macroeconomic aggregates in the long run: perfect substitutes.

Var.	Status quo	Trans. dynamics		Steady state	
		Opt. pol.	Change	Opt. pol.	Change
τ_l	27.55%	22.9%	−4.68% p	36.98%	9.43% p
d	27.1%	10%	−17.1% p	31%	3.9% p
θ	38.8%	120%	81.2% p	170%	131.2% p
Z	0.0714	0.0303	−57.56%	0.0866	21.38%
$\tau_l \cdot Z$	0.0197	0.0069	−64.75%	0.032	62.94%
Y/N	0.2633	0.3032	15.15%	0.2794	6.12%
B/Y	60.16%	96.95%	36.79% p	60.16%	0% p
K/N	0.1731	0.1951	12.74%	0.1781	2.92%
L/N	0.4532	0.5259	16.06%	0.4882	7.73%
K/L	0.5346	0.5194	−2.86%	0.5108	−4.47%
w	0.5449	0.5407	−0.78%	0.5368	−1.5%
$\frac{w_c}{w_n}$	1	1	0%	1	0%
r	5.37%	5.54%	0.17% p	5.7%	0.33% p
Hours	0.326	0.3287	0.27%	0.2985	−2.75%
C/N	0.1628	0.1934	18.82%	0.1717	5.47%
Trans/assets	1.15%	1.39%	0.23% p	0.78%	−0.37% p
College share	43.89%	75.79%	31.9% p	82.07%	38.18% p
$Gini(c)$	0.235	0.2186	−1.64 p	0.1987	−3.62 p
$Gini(h)$	0.12	0.1283	0.83 p	0.1176	−0.24p
$Gini(a)$	0.5558	0.5203	−3.55 p	0.5336	−2.22 p
CEV		1.6565%		2.6126%	

Note: This table summarizes optimal policies and the long-run change in macroeconomic aggregates, distributions and welfare for $\rho=1$ (in which case unskilled and skilled labor are perfect substitutes in production).

Results

Analysis of optimal policy transitions: Transitional dynamics

As we discussed, it takes time (and resources) to build up a more skilled workforce, suggesting that an explicit consideration of the transition path is important.

Comparing the optimal policy transitions ($d = 10\%$) and the steady state optimal policy ($d = 31\%$) (see columns 3 and 4 of Table 4) strongly reinforces this point. Explicitly taking into account the transitional costs, the optimal tax code should become less, not more progressive.

This raises two questions:

(a) Why is there such a divergence between steady state optimal and dyanmically optimal policies?

(b) Why does not this divergence show up in the benchmark version of the model (see columns 3 and 5 of Table 3), from $d = 6\%$ to $d = 10\%$ in steady state?

To answer the first question we look at Figure 6 and 7. Recall that in the benchmark, dynamically optimal, and steady state optimal with perfect substitutes the education subsidies are $\theta = 38.8\%, 120\%$ and 170% respectively and the degree of progressivity is given by $d = 27.1\%, 10\%$ and 31% respectively.

Then, we see different the consumption path for policy combinations (see Fig. 7). If the government implements the dynamically optimal increase in the education subsidy (from $\theta = 38.8\%$ to $\theta = 120\%$) but not the dynamically optimal decrease in progressivity from $d = 27.1\%$ to $d = 10\%$, then the economy goes through a recession. As HHs are drawn from the labor market into college, the econmy experiences a tansitional decline in output and per capita consumption.

Following this same argument, if the government implements the optimal steady state policy we find even a more severe recession because the subsidy is substantially higher and so is progressivity.

Note that the economy does not go into a recession if we implement the dynamically optimal policies with $d = 10\%$ and $\theta = 120\%$.

Welfare consequences for different policy combinations: perfect substitutes

	$d = 10\%, \theta = 120\%$	$d = 27\%, \theta = 120\%$	$d = 31\%, \theta = 170\%$
CEV	1.65%	0.09%	-2.80%

Implementing the steady state optimal policy and computing gains that take into account the transitional costs (which are ignored when defining optimality in the steady state) implies a loss of welfare by -2.80%.

That is, an explicit consideration of the transition path when conducting a normative welfare analysis of education and tax policies can change even the qualitattvie prescriptions for optimal fiscal policy dramatically.

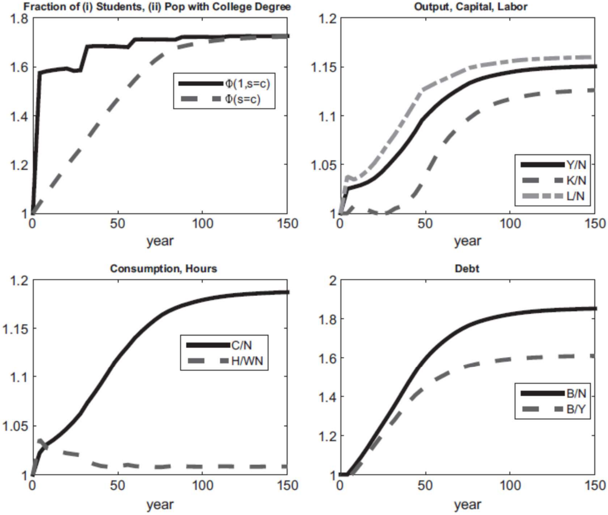


Fig. 6. Evolution of macroeconomic aggregates: perfect substitutes. *Note:* This figure plots the transition paths of key macroeconomic aggregates against time. Unskilled and skilled labor are prefect substitutes ($\rho = 1$).

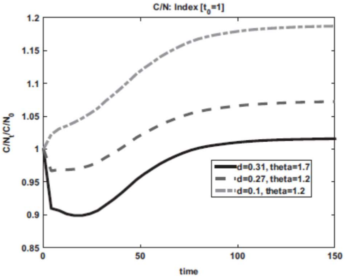


Fig. 7. Evolution of per-capita consumption: perfect substitutes. *Note:* This figure plots the transition paths of per capita consumption under different policy configurations. The elasticity of substitution between unskilled and skilled labor is 1.4 (ρ=0.285).

Why does not the same logic apply to the benchmark case with imperfect skill substitution? There we do not find a strong divergence between steady state and transitional optimal polices.

The answer lies in the GE effect that allow the government to achieve effective redistribution through education subsidies. Therefore a high degree of tax progressivity (a higher d) is not required to achieve a less dispersed consumption distribution even in the steady state.

We discuss this next.

Recall that switching to the optimal policy ($d = 6\%, \theta = 150\%$) in the benchmark economy implied a CEV gain of 3.5% (see column 4, last row in Table 3).

In comparison the welfare gain of this very same policy with perfect substitutes is smaller 1.61%, which is similar to the welfare gain of the optimal transitional policy with perfect substitutes ($d = 10\%, \theta = 120\%$ is 1.65%.).

Thus, the GE response, especially in the reduction fo the college wage premium, is very ipmortant fo the welfare assessmen tof the optimal policy.

We show (appendix) how progressive taxes and education subsidies can be complementary second best policies in a world where private insurance is imperfect, and thus public insurance is potentially beneficial BUT distorts human capital accumulation decisions, with education subsidies mitigating these distortions.

With the two types of labor being imperfect substitutes (i.e., with GE effects) the education policy has indifect beneficail redistributive effects, by reducing the average wage gap between skilled and unskilled workers, and thus might potentially be a sustitute ofr redistributive tax policies in GE.

Constrained-optimal policies.

To explore this complementarity/substitutability across policies we compute optimal policies along one policy dimension, holding the other dimension constant. Restrict to steady states to make the point. Table 6 shows these constrained-optimal policies when one instrument (either the education subsidy θ or the size of the deduction d is held constant).

Table 6
Final steady state: decomposition.

Var.	Perfect substitutes $\rho=1$			Imperfect substitutes $\rho=0.28$		
	Baseline	Const. d	Const. θ	Baseline	Const. d	Const. θ
τ_l	36.9%	35.0%	29.3%	26.0%	32.5%	29.3%
d	31%	27.1%	31%	10%	27.1%	31%
θ	170%	175%	38.8%	200%	175%	38.8%
Z	0.086	0.077	0.08	0.028	0.071	0.080
$\tau_l \cdot Z$	0.032	0.027	0.023	0.007	0.023	0.023
CEV	2.612%	2.508%	0.167%	2.756%	2.413%	0.130%

Note: This table displays the steady state welfare gains for different policy configurations.

We start with the case of perfect substitutes (no GE effects) in the left panel of Table 6:[8]

- When θ is fixed to its status quo value $\theta = 38.8\%$ (which is lower then joint optimal $\theta = 170\%$, the constrained-optimal d is the same as the joint. However, the constrained optimal effective tax deduction defined by Z or $\tau_l Z$ is lower as well. Suggesting a complementary role between θ and d. When d is fixed to its status quo $d = 27.1\%$ (which is lower than the joint optimal $d = 31\%$, the education subsidy is larger.
- The joint reform has (steady state) welfare gains of 2.61% of lifetime consumption, and most of the welfare gains stem from an adjustment of the educ. subsidy:
  - Only adjusting the educ. subsidy generates almost the same gains, 2.5%.
  - Only adjusting the progressivity alone yields welfare gains of 0.17%.

  That is, absent GE wage effects the educ. policy is very effective in stimulating college attendance and long-run labor productivity.

[8] In status quo $d = 27.1\%$, $\theta = 38.8\%$, $Z = 0.0714$, and $\tau_l Z = 0.0197$.

The case of imperfect subsitites is in the right panel of Table 6.[9]

- As in the case of perfect substitutes, the univariate optima favor a higher degree of tax progressivity and education subsdies.
- However, once both poicy instruments are in operation, the two instruments become substitues: in the unconstrained optimum, the degree of education subsidies is higher thanin the univariate optimum and tax progressivity decreases.

The additional effect of a decrase of hte college wage premum alllows the government to achieva a more equal consumption distribution (at least across college- vs. noncollege HHs) eithe rthrough college educ. subsidies or progressive income taxes.

[9] In status quo $d = 27.1\%$, $\theta = 38.8\%$, $Z = 0.0714$, and $\tau_l Z = 0.0197$.

Conclusion. Whether the two policies complement each other or become substitutes in providing redistribution crucially depends on the importance of GE effects. Since we find the empirical evidence for imperfect substitutatility across skills compelling, we conclude that the case for education subsidies and progressive labor income taxes being substitutable policies is strong.

So far we restricted the policy instruments to be constant along the transition path. We can study whether the results are driven by this assumption or not.

We study whether the government forgoes important welfare gains by not phasing in the policy reforms more slowly, permitting time-varying fiscal and education policies.

To do so we now consider transitional policies of the form:

$$\begin{aligned} \tau_{l,t} &= e^{-\omega(t-1)}\tau_{l,0}(1-e^{-\omega(t-1)}\tau_{l,T} \\ \theta_t &= e^{-\omega(t-1)}\theta_0(1-e^{-\omega(t-1)}\theta_T \\ d_t &= e^{-\omega(t-1)}d_0(1-e^{-\omega(t-1)}d_T \end{aligned}$$

where $(\tau_{l,0}, \theta_0, d_0)$ is the calibrated policy in the initial steady sate and $(\tau_{l,T}, \theta_T, d_T)$ is an arbitrary time constant new policy. Then $\omega \in [0, \infty)$ is a parameter that governs the speed of adjustment. If $\omega = \infty$ then adjustment is immediate; this is the assumption in the main text. The smaller is ω, the more gradual the new policy is being introduced.

(1) Is it optimal to more gradually introduce the policy we found to be optimal along the transition? To answer this question the authors set $(\tau_{l,T}, \theta_T, d_T)$ to the dynamic optimal triplet $(\tau_l^*, \theta^*, d^*)$ and maximize social welfare with respect to ω. They find $\omega = 4$ with no benefit from introducing the optimal policy more gradually than that. In this sense, results are robust towards permitting time variation in the policy instruments.

(2) Is the suboptimality of the policy that maximizes steady state welfare along the transition due to the fact that it is introduced too rapidly? To answer this question the authors set $(\tau_{l,T}, \theta_T, d_T)$ to the steady-state optimal triplet $(\tau_l^{ss}, \theta^{ss}, d^{ss})$ and again maximize welfare with respect to ω. The answer is yes. The steady state optimal policy leads to larger welfare gains along the transition if introduced more slowly and phased in over a time interval of about 16 years.

Two papers that are also very interesting in this literature:

- "How Does Tax Progressivity and Household Heterogeneity Affect Laffer Curves?," Holter, Krueger, and Stepanchuk (2017).
- "Optimal Tax Progressivity: An Analytical Framework," Heathcote, Storesletten and Violante (2016).