



MAX-PLANCK-INSTITUT FÜR SOZIALRECHT UND SOZIALPOLITIK
MAX PLANCK INSTITUTE FOR SOCIAL LAW AND SOCIAL POLICY

mea *Munich Center for the Economics of Aging*

On the Optimal Provision of Social Insurance

Dirk Krueger, Alexander Ludwig

02-2013

MEA DISCUSSION PAPERS



Alte Nummerierung: 267-13

On the Optimal Provision of Social Insurance*

Progressive Taxation versus Education Subsidies in General

Equilibrium

Dirk Krueger

University of Pennsylvania, CEPR and NBER

Alexander Ludwig

SAFE at Goethe University Frankfurt, CMR and MEA

April 1, 2015

Abstract

In this paper we compute 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. Optimally chosen tertiary education subsidies mitigate these distortions. We highlight the importance of two different channels through which academic talent is transmitted across generations (persistence of innate ability vs. the impact of parental education) for the optimal design of these policies and model different forms of labor as imperfect substitutes, thereby generating general equilibrium feedback effects from policies to relative wages of skilled and unskilled workers. We show that subsidizing higher education has important redistributive benefits, by shrinking the college wage premium in general equilibrium. We also argue that a full characterization of the transition path is crucial for policy evaluation. We find that optimal education policies are always characterized by generous tuition subsidies, but the optimal degree of income tax progressivity depends crucially on whether transitional costs of policies are explicitly taken into account and how strongly the college premium responds to policy changes in general equilibrium.

Keywords: Progressive Taxation, Education Subsidy, Transitional Dynamics

J.E.L. classification codes: E62, H21, H24

*We thank Felicia Ionescu for helpful discussions and seminar participants at various institutions and conferences for many useful comments. Krueger gratefully acknowledges financial support from the NSF under grant SES-0820494. Ludwig gratefully acknowledges financial support by the Research Center SAFE, funded by the State of Hessen initiative for research LOEWE.

1 Introduction

In the presence of uninsurable idiosyncratic earnings risk, progressive taxation provides valuable social insurance among ex ante identical households. In addition it might enhance equity among ex ante heterogeneous households, which is beneficial if the social welfare function used to aggregate lifetime utilities values such equity. However, if high-earnings households face higher average tax rates than low-earnings households, this might discourage the incentives of these households to become earnings-rich through making conscious human capital accumulation decisions. 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 supply channel.

In this paper we compute the optimal tax and education policy transition, within a parametric class, in an economy where progressive taxes provide social insurance against idiosyncratic wage risk, but distort the education decision of households. Optimally chosen tertiary education subsidies mitigate these distortions. We highlight the importance of two different channels through which academic talent is transmitted across generations (persistence of innate ability vs. the impact of parental education) for the optimal design of these policies. We also model different forms of labor as imperfect substitutes, thereby generating general equilibrium feedback effects from policies to relative wages of skilled and unskilled workers.

We show that subsidizing higher education has important redistributive benefits, by shrinking the college wage premium in general equilibrium. We also argue that a full characterization of the transition path is crucial for policy evaluation. We find that optimal education policies are always characterized by generous tuition subsidies. The optimal degree of income tax progressivity however crucially depends on whether transitional costs of policies are explicitly taken into account and how strongly the college premium responds to policy changes in general equilibrium.

If unskilled and skilled labor are perfect substitutes (as in our previous work, Krueger and Ludwig, 2013, and in a substantial body of previous literature) and if we restrict ourselves to a steady state analysis, the optimal policy is characterized by a massive education subsidy of 170% of the college tuition cost¹, and a tax system more progressive than the current status quo calibrated to U.S. data.

Both policies complement each other in the steady state, in that *both* the public education subsidy θ (as a share of total tuition costs) and the relative tax deduction d (one measure of tax progressivity, measured as fraction of average income in the economy) increase from their status quo levels of $(\theta = 38.8\%, d = 27.1\%)$ to $(\theta = 170\%, d = 31\%)$. We also show that the level of the tax deduction (an alternative measure of tax progres-

¹ This implies that, effectively, the government not only provides free tertiary education, but also covers the living expenses of those going to college.

sivity, the product of d and average income) is higher in the economy with optimal θ, d than in one where we hold θ constant and maximize over d alone.² These findings point towards a strong complementarity of the two instruments. In addition to improved social insurance, the substantial welfare gains (in the order of 2.6% of lifetime consumption) stem from the fact that per capita output (and thus consumption) increases in the long run despite average hours worked falling. This is feasible since an education boom (the share of college educated individuals increases by 50% in the long run) improves the skill distribution in the economy.

However, building up the skill distribution takes time (as new college students are only a small part of the overall workforce), and along the transition path the economy undergoes a severe recession induced by the decline in labor supply and capital accumulation on account of the higher marginal tax rates required to pay for the more generous education subsidy and higher tax deduction. Therefore taking the transitional costs into account the steady state optimal policy actually entails significant *welfare losses* relative to the status quo, in the order of 2.8% of lifetime consumption. The optimal policy reform, taking the transition path explicitly into account, still calls for substantially higher education subsidies than the current status quo (of 125% of college tuition costs), but now calls for *less progressive* taxes (both relative to the steady state optimum of $d = 30\%$ as well as the status quo of $d = 27.1\%$) in order to avoid the short-run recession induced by higher tax marginal tax rates. The optimal tax deduction falls to $d = 10\%$ of average income and optimal marginal taxes decrease from a status quo of 27.5% to 23%. The finding that an explicit consideration of transitional dynamics in the analysis of education finance reform in models with endogenous human capital accumulation is potentially very important for optimal tax design is the first main quantitative conclusion of this paper.

If skilled and unskilled labor are perfect substitutes in production, the policy-induced college boom does not affect the college wage premium. Building on the substantial literature that has estimated the degree of substitutability of these two types of labor to be less than perfect (see e.g. Katz and Murphy (1992) or Borjas (2003)) we then relax this assumption and characterize optimal policy under our preferred substitution elasticity of 1.4. Our main optimal policy conclusions (large education subsidy, modest tax progressivity when the transition path is taken into account, substantial welfare gains in the order of 3% of lifetime consumption) remain intact: the optimal education subsidy rate is 150%, the optimal tax deduction now only 6% of average income and marginal taxes are at 22%.

However, there are important qualifications. First, with imperfect substitutes the strong rationale for progressive taxes disappears even in the steady state (the optimal steady state policy has a subsidy rate of 200% and a tax deduction of 10% of mean income). Since the generous education subsidy induces more individuals to go to college, the college wage premium falls in equilibrium which constitutes a *policy substitute* for re-

² The optimal level of d itself remains at $d=30\%$.

distributive tax progressivity.^{3,4} That general equilibrium wage effects can turn education subsidies and progressive taxes from policy complements to policy substitutes is the second main quantitative conclusion of this paper.

The paper is organized as follows. After relating our contribution to the literature in the next section, in section 3 we construct a simple, analytically tractable model to argue why progressive income taxes and education subsidies might simultaneously be part of an optimal government fiscal policy in the presence of an endogenous education decision. In order to make that argument most clearly, in that section of the paper we abstract from general equilibrium effects of these policies as well as the dynamics induced by asset accumulation and the intergenerational transmission of talent and wealth. These elements are then introduced in the quantitative model in section 4 where we set up the model and define equilibrium for a given fiscal policy of the government. Section 5 describes the optimal tax problem of the government, including its objective and the instrument available to the government. After calibrating the economy to U.S. data (including current tax and education policies) in section 6 of the paper, section 7 displays the results and interpretation of the optimal taxation analysis, first for the benchmark case of perfect substitutability of labor in production, and then for our preferred specification in which skilled and unskilled labor are imperfect substitutes and thus the policies affect the college wage premium. Section 8 concludes. The appendices contain the proofs of the propositions from section 3 as well as details of the calibration and the computation of the quantitative version of the model from section 4 of the main paper.

2 Relation to the Literature

Our paper aims at characterizing the optimal progressivity of the income tax code in a life cycle economy in which the public provision of redistribution and income insurance through taxation and education policies is desirable, but where progressive taxes not only distort consumption-savings and labor-leisure choices, but also household human capital accumulation choices. It is most closely related to the studies by Conesa and Krueger (2006), Conesa et al. (2009) and Karabarbounis (2012). Relative to their steady state analyses we provide a full quantitative *transition analysis* of the optimal tax code in a model with *endogenous education choices*. Bakis et al. (2014), Fehr and Kindermann (2015) and Kindermann and Krueger (2015) also compute optimal tax transitions, but abstract from endogenous human capital accumulation.

³ The policy-induced reduction in the college wage premium in turn weakens education incentives; thus a larger increase in subsidies is required to achieve a given expansion in educational attainments and the long-run effect of the policy on educational attainment is smaller.

⁴ Given that the optimal tax policy is already not very progressive in the steady state, the policy differences between the steady state and the transition are qualitatively similar to the perfect substitutes case, but quantitatively not very important.

In our earlier paper, Krueger and Ludwig (2013) we characterized the optimal policy transition in a model with college choice in which the innate ability distribution of children for college was exclusively determined by the education level of the parents, and in which tax and education policies have no general equilibrium impact on the relative wages of college versus non-college labor. Consequently education subsidies are a potent tool to encourage college attendance but have no redistributive benefits through reducing the college wage premium. Both features stacked the deck for finding large education subsidies and policy complementarity between these policies and redistributive taxation. Relative to this work here we use micro data to discipline the relative importance of parental education and parental innate ability in the intergenerational transmission of skills,⁵ and we model general equilibrium wage effects explicitly. We show that whereas the large education subsidies remain optimal, the optimal degree of tax progressivity declines substantially. Policy complements turn into policy substitutes.

Our paper follows a long tradition in the literature that uses quantitative overlapping generations models in the spirit of Hubbard and Judd (1986) and Auerbach and Kotlikoff (1987), but enriched by uninsurable idiosyncratic earnings risk as in Bewley (1986), Huggett (1993, 1997) and Aiyagari (1994), to study the optimal structure of the tax code in the Ramsey tradition, see Chamley (1986) and Judd (1985). The optimal tax code in life cycle economies with a representative household in each generation was characterized in important papers by Alvarez et al. (1992), Erosa and Gervais (2002), Garriga (2003), Gervais (2012) and Bovenberg and Jacobs (2010), and in economies with private information in the Mirrleesian (Mirrlees 1971) tradition, by Judd and Su (2006), Fukujima (2010), Bohacek and Kapicka (2008), Kapicka (2012), Findeisen and Sachs (2013, 2014)⁶ and Weinzierl (2011).⁷

In models in which progressive labor income taxes potentially distort education decisions a public policy that subsidizes these choices might be effective in mitigating the distortions from the tax code, as pointed out effectively by Bovenberg and Jacobs (2005). As in their theoretical analysis we therefore study such subsidies explicitly as part of the optimal policy mix in our quantitative investigation. Our focus of the impact of the tax code and education subsidies on human capital accumulation decisions strongly connects our work to the studies by Heckman et al. (1998, 1999), Benabou (2002), Caucutt et al. (2006), Bohacek and Kapicka (2012), Kindermann (2012), Ab-

⁵ Our model-implied impact of education subsidy payments on college attendance is consistent with Dynarski's (2003) empirical findings.

⁶ The focus of the last four papers on optimal income taxation in the presence of human capital accumulation make them especially relevant for our work, although they abstract from explicit life cycle considerations.

⁷ There is also large literature on the positive effects of various taxes on allocations and prices in life cycle economies. See e.g. Hubbard and Judd (1986) and Castañeda et al. (1999) for representative examples. The redistributive and insurance role of progressive taxation in models with heterogeneous households is also analyzed in Domeij and Heathcote (2004) and Heathcote et al. (2012).

bott et al. (2013), Holter (2014), Winter (2014) and Guvenen et al. (2014), although the characterization of the optimal tax code is not the main objective of these papers. In stressing the importance of the sources of the intergenerational transmission of talent (for college) and the general equilibrium wage effects of education policies for *optimal policies* we build especially strongly on the study by Abbott et al. (2013).

In our attempt to contribute to the literature on (optimal) taxation in life cycle economies with idiosyncratic risk and human capital accumulation we explicitly model household education decisions (and government subsidies thereof) in the presence of borrowing constraints and the intergenerational transmission of human capital as well as wealth. Consequently our works builds upon the massive theoretical and empirical literature investigating these issues, studied and surveyed in, e.g. Keane and Wolpin (2001), Cunha et al. (2006), Holmlund et al. (2011), Lochner and Monge (2011).⁸

3 A Simple Model

We now present a simple model⁹ that allows us to make precise the intuition that with incomplete financial markets progressive labor income taxes might be part of optimal fiscal policy because it implements a more equitable consumption distribution than the laissez faire competitive equilibrium, but that it distorts both the labor supply *and* the education decision. The latter distortion can be partially offset by an education subsidy which then becomes part of an optimal policy mix as well. Relative to the quantitative model used in the next sections, the model analyzed here abstracts from general equilibrium feedbacks and the two key sources of dynamics in that model, endogenous capital accumulation and the intergenerational transmission of talent and wealth.

3.1 The Environment

The economy lasts for one period and is populated by a continuum of measure one of households that differ by ability e . The population distribution of e is uniform on the unit interval, $e \sim U[0, 1]$. Households value consumption and dislike labor according to the utility function

$$\log \left(c - \mu \frac{l^{1+\frac{1}{\psi}}}{1+\frac{1}{\psi}} \right). \quad (1)$$

⁸ A comprehensive survey of this literature is well beyond the scope of this introduction. We will reference the papers on which our modeling assumptions or calibration choices are based specifically in sections 4 and 6.

⁹ The model builds on the analysis in Bovenberg and Jacobs (2005, 2009), but is tailored towards being a special case of the quantitative model studied below.

These Greenwood, Hercowitz and Huffman (1988) preferences rule out wealth effects on labor supply (which greatly enhances the analytical tractability of the model), but at the same time make utility strictly concave in consumption, which induces a redistribution/insurance motive for a utilitarian social planner or government.

A household can either go to college or not. A household with ability e that has gone to college produces $(1 + pe)w$ units of consumption per unit of labor, whereas a household without a college degree has labor productivity w . Here $w > 0$ and $p > 0$ (the college premium for the most able type) are fixed positive parameters. Going to college requires κw resources (but no time) where $\kappa > 0$ is a parameter.

3.2 Social Planner Problem

Prior to analyzing the competitive equilibrium without and with fiscal policy we establish, as a benchmark, how a social planner with utilitarian social welfare function would allocate consumption and labor across the population. The social planner problem chooses consumption and labor supply $c(e), l(e)$ for each type $e \in [0, 1]$ as well as the set I of types that are being sent to college to solve

$$\begin{aligned} & \max_{c(e), l(e), I} \int \log \left(c(e) - \mu \frac{l(e)^{1+\frac{1}{\psi}}}{1 + \frac{1}{\psi}} \right) de \\ & s.t \\ & \int c(e) de + \kappa w \int_{e \in I} de = \int_{e \notin I} wl(e) de + \int_{e \in I} (1 + pe)wl(e) de \end{aligned}$$

The following proposition summarizes its solution, under the following

Assumption 1

$$\frac{\left(\frac{\mu^\psi (1+\psi)\kappa}{w^\psi} + 1 \right)^{\frac{1}{1+\psi}} - 1}{p} < 1.$$

Proposition 2 *Suppose assumption 1 is satisfied. Then the solution to the social planner problem is characterized by an ability threshold e^{SP} such that all households with $e \geq e^{SP}$ are sent to college (and indexed with subscript c from now on) and the other households are not (and are indexed by n). Labor allocations are given by*

$$l_n = \left(\frac{w}{\mu} \right)^\psi \text{ for all } e < e^{SP} \quad (2)$$

$$l_c(e) = \left(\frac{(1 + pe)w}{\mu} \right)^\psi \text{ for all } e \geq e^{SP} \quad (3)$$

Consumption allocations are characterized by

$$\begin{aligned} c(e) &= c_n \text{ for all } e < e^{SP} \\ c(e) &= c_c(e) = c_n + \mu \frac{l_c(e)^{1+\frac{1}{\psi}}}{1+\frac{1}{\psi}} - \mu \frac{l_n^{1+\frac{1}{\psi}}}{1+\frac{1}{\psi}} \text{ for all } e \geq e^{SP} \end{aligned}$$

The optimal education threshold satisfies the first order condition

$$\kappa w = \frac{w^{1+\psi}}{\mu^\psi (1+\psi)} \left((1+pe^{SP})^{1+\psi} - 1 \right) \quad (4)$$

and is given in closed form as

$$e^{SP} = \frac{\left(\frac{\mu^\psi (1+\psi) \kappa}{w^\psi} + 1 \right)^{\frac{1}{1+\psi}} - 1}{p} = e^{SP}(\kappa, p, w, \mu) \quad (5)$$

Thus the larger is p, w and the smaller is κ, μ , the smaller is the education threshold and thus the more households are sent to college.

Proof 3 The threshold property of set I follows from the fact that the cost of college is independent of e and the productivity benefits pe of being college-educated are strictly increasing in e . The other results are directly implied by the first order conditions (which in the case of the education threshold e^{SP} involves applying Leibnitz' rule to the resource constraint, after having substituted in the optimal labor allocations). Assumption 1 assures that $e^{SP} \in (0, 1)$. Note that this assumption is purely in terms of the structural parameters of the model and requires that the college productivity premium p is sufficiently large, relative to the college cost κ , for the ablest households indeed be sent to college.

Equation (4) has an intuitive interpretation. The social planner chooses the optimal education threshold such that the cost κw of education for the marginal type e^{SP} equals the net additional resources this marginal type generates with a college education, relative to producing without having obtained a college education. The term on the right hand side of (4) takes into account that college educated households work longer hours (this explains the exponent $1 + \psi$) and the fact that college-educated households are compensated for their longer hours with extra consumption which explains the factor $\frac{1}{\mu^\psi(1+\psi)}$. This latter property of optimal allocations follows from the non-separability of consumption and labor in preferences (1).

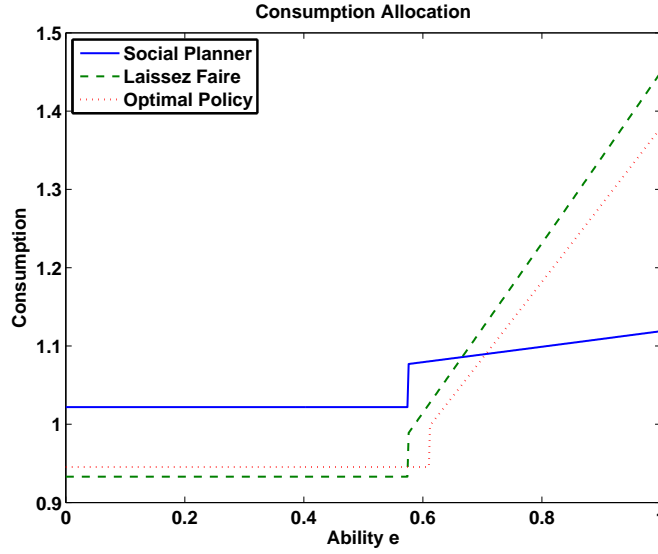
For future comparison with equilibrium consumption allocations we state:

Corollary 4 The optimal consumption allocation satisfies $c_n < c_c(e^{SP})$ and $c_c(e)$ is strictly increasing in e for all $e > e^{SP}$.

We depict the optimal consumption allocation in figure 1, together with two equilibrium allocations discussed in the next subsection.¹⁰

¹⁰ The specific parameter values that underly this figure are discussed below as well.

Figure 1: Optimal and Equilibrium Consumption Allocations



3.3 Competitive Equilibrium

Now we study the competitive equilibrium of this economy. To do so we first have to specify the market structure and the government policies. Households participate in two competitive markets, the goods market where they purchase consumption goods, and the labor market where they earn a wage per unit of labor supplied that equals their marginal product. In addition to choosing consumption and labor households decide whether to incur the cost κw of going to college. The benefit of doing so is a wage premium $epw > 0$. We denote by $\mathcal{I}(e) \in \{0, 1\}$ the college choice of household type e , with $\mathcal{I}(e) = 1$ if the household goes to college.

Financial markets, however, are assumed to be incomplete. Although there is no scope for intertemporal trade, in principle households would like to trade insurance contracts against the risk of being born as a low ability e type, prior to the realization of that risk. It is the insurance against this idiosyncratic wage risk that we rule out by assumption, and this fundamental market failure will induce a motive of insurance/redistribution for the benevolent, utilitarian government in this economy.¹¹

We assume that the benevolent, utilitarian government (which from now on we will frequently refer to as the Ramsey government) has access to three fiscal policy instrument, a flat labor income tax with tax rate τ , a lump sum transfer/tax $d \cdot w$ and a proportional education subsidy with rate θ . Note that by permitting $d < 0$ we allow the government to levy lump sum taxes; on the other hand it can also implement a progressive labor income tax schedule by setting $d, \tau > 0$. What we do not permit

¹¹ What is insurance ex ante (prior to the realization) of the e draws, is redistribution among different e types ex post.

are policies that make taxes or subsidies type-specific by conditioning τ, d, θ on type e . Given these restrictions on the tax code we would not expect the government to be able to implement the solution to the social planner problem as a competitive equilibrium with fiscal policies.

3.3.1 Definition and Characterization of Equilibrium

Now consider the problem of a generic household $e \in [0, 1]$. Given a fiscal policy the household's problem reads as

$$\max_{l(e), c(e) \geq 0, \mathcal{I}(e) \in \{0, 1\}} \log \left(c(e) - \mu \frac{l(e)^{1 + \frac{1}{\psi}}}{1 + \frac{1}{\psi}} \right) \quad (6)$$

s.t.

$$c(e) = (1 - \tau)(1 + \mathcal{I}(e)pe)wl(e) + dw - \kappa w(1 - \theta)\mathcal{I}(e) \quad (7)$$

Definition 5 For a given fiscal policy (τ, d, θ) a competitive equilibrium are consumption, labor and education allocations $c(e), l(e), \mathcal{I}(e)$ such that

1. For all $e \in [0, 1]$, the choices $c(e), l(e), \mathcal{I}(e)$ solve the household maximization problem (6).
2. The government budget constraint is satisfied:

$$dw + \kappa w \theta \int_{\{e: \mathcal{I}(e)=1\}} de = \tau \left(w \int_{\{e: \mathcal{I}(e)=0\}} l(e) de + w \int_{\{e: \mathcal{I}(e)=1\}} (1 + pe)l(e) de \right)$$

3. The goods market clears

$$\int c(e) de + \kappa w \int_{\{e: \mathcal{I}(e)=1\}} de = w \int_{\{e: \mathcal{I}(e)=0\}} l(e) de + w \int_{\{e: \mathcal{I}(e)=1\}} (1 + pe)l(e) de$$

We can completely characterize the competitive equilibrium, for a given fiscal policy. We summarize the results in the following

Proposition 6 Given a policy (τ, d, θ) , the optimal labor supply of households not going to college is given by

$$l_n(\tau) = \left(\frac{(1 - \tau)w}{\mu} \right)^\psi \quad (8)$$

whereas the optimal labor supply of households with a college education is given by

$$l_c(e; \tau) = \left(\frac{(1 - \tau)(1 + pe)w}{\mu} \right)^\psi. \quad (9)$$

The corresponding consumption allocations read as

$$c_n(\tau, d) = \frac{[(1-\tau)w]^{1+\psi}}{\mu^\psi} + d\tau \quad (10)$$

$$c_c(e; \tau, d, \theta) = \frac{[(1-\tau)w]^{1+\psi} (1+pe)^{1+\psi}}{\mu^\psi} + d\tau - \kappa w(1-\theta) \quad (11)$$

There is a unique education threshold e^{CE} such that all types with $e \geq e^{CE}$ go to college and the others don't. This threshold satisfies

$$c_n(\tau, d) - \frac{\mu l_n(\tau)^{1+\frac{1}{\psi}}}{1+\frac{1}{\psi}} = c_c(e^{CE}; \tau, d, \theta) - \frac{\mu l_c(\tau, e^{CE})^{1+\frac{1}{\psi}}}{1+\frac{1}{\psi}}$$

and is explicitly given by¹²

$$e^{CE} = \frac{\left(\frac{(1-\theta)\mu^\psi(1+\psi)\kappa}{(1-\tau)^{1+\psi}w^\psi} + 1 \right)^{\frac{1}{1+\psi}} - 1}{p} = e^{CE}(\tau, \theta; \kappa, p, w, \mu) \quad (12)$$

The threshold e^{CE} is strictly decreasing (the share of households going to college is strictly increasing) in θ , strictly increasing in τ and independent of d .

Proof 7 The equilibrium labor allocations follow directly from the first order conditions of the household problem. The equilibrium consumption allocations are then implied by plugging equilibrium labor supply into the household budget constraint (7). Thus lifetime utility conditional on not going to college is given by

$$\log \left(c_n(\tau, d) - \frac{\mu l_n(\tau)^{1+\frac{1}{\psi}}}{1+\frac{1}{\psi}} \right)$$

which is constant in ability e , and lifetime utility conditional on going to college reads as

$$\log \left(c_c(e; \tau, d, \theta) - \frac{\mu l_c(\tau, e)^{1+\frac{1}{\psi}}}{1+\frac{1}{\psi}} \right)$$

which is strictly increasing in e . The threshold result for the education decision thus follows, and the threshold itself is determined by the indifference between attending and not attending college at the threshold.

¹² This result assumes that $e^{CE} \leq 1$. An assumption similar to assumption 1 is required to assure this, and we assume (and ex-post check) such an assumption to hold for the range of policies considered below.

3.3.2 Optimality of Equilibrium without Government Intervention?

In this subsection we show that the unregulated competitive equilibrium displays an optimal (in the sense of solving the social planner problem) labor and education allocation and thus optimal production. However, the consumption distribution is suboptimally highly dispersed in the competitive equilibrium without government policies. The first result (a simple corollary to proposition 6) states that the competitive equilibrium without government intervention has an optimal labor and education allocation. It follows directly from comparing equations (2),(3) and (5) in the social planner problem to equations (8)-(12) in the competitive equilibrium, evaluated at $\tau = d = \theta = 0$.

Corollary 8

$$\begin{aligned} l_n^{CE}(\tau = 0) &= l_n^{SP} \\ l_c^{CE}(e; \tau = 0) &= l_c^{SP}(e) \\ e^{CE}(\tau = 0, \theta = 0) &= e^{SP} \end{aligned}$$

This last result also implies that aggregate output, defined in the competitive equilibrium as

$$\begin{aligned} L^{CE}(\tau, \theta) &= e^{CE}(\tau, \theta) w l_n^{CE}(\tau) + w \int_{e^{CE}}^1 (1 + pe) l_c^{CE}(e; \tau) de \\ &= e^{CE}(\tau, \theta) w \left(\frac{(1 - \tau)w}{\mu} \right)^\psi + w \int_{e^{CE}(\tau, \theta)}^1 (1 + pe) \left(\frac{(1 - \tau)w(1 + pe)}{\mu} \right)^\psi de \end{aligned}$$

is at the optimal level as well: $L^{CE}(\tau = 0, \theta = 0) = L^{SP}$. Note that aggregate output in the competitive equilibrium is strictly decreasing in the tax rate τ , strictly increasing in the college subsidy θ (since θ raises the share of households going to college and college-educated households are more productive and work longer hours), and independent of the lump-sum tax/subsidy d .

However, the next proposition shows that the consumption distribution in the competitive equilibrium is suboptimally dispersed since non-college households consume too little in the competitive equilibrium, so do non-productive college graduates. Furthermore, the dependence of consumption on individual ability e is suboptimally high in the competitive equilibrium without government intervention:

Proposition 9 *In the competitive equilibrium for policy $\tau = d = \theta = 0$ (and thus $e^{CE} = e^{SP}$) we have*

$$\begin{aligned} c_n^{CE} &< c_n^{SP} \\ c_c^{CE}(e^{SP}) &< c_c^{SP}(e^{SP}) \\ \frac{\partial c_c^{CE}(e)}{\partial e} &> \frac{\partial c_c^{SP}(e)}{\partial e} \end{aligned}$$

Proof 10 See Appendix A

The equilibrium consumption allocation in the absence of government policy is depicted in figure 1 and shows the excess consumption inequality proved in proposition 9. The social planner, relative to the competitive equilibrium without policies, provides additional consumption insurance, both between education groups, and within the high education group. As discussed in the beginning of this section, the fundamental market failure that leads to the suboptimality of the competitive equilibrium is the absence of insurance markets against e -risk.¹³

Also note that the socially optimal allocation cannot be implemented as a competitive equilibrium with the policies (τ, d, θ) unless policies can be made e -type specific. This can be seen from recognizing that, under the restricted policies, insuring $\frac{\partial c_c^{CE}(e)}{\partial e} = \frac{\partial c_c^{SP}(e)}{\partial e}$ requires a positive labor income tax $\tau > 0$ that satisfies

$$\frac{1}{1 + \frac{1}{\psi}} = (1 - \tau)^{1+\psi} \quad (13)$$

but such a tax distorts the labor supply decisions of households, a distortion that cannot be corrected with the existing set of instruments (see equations (8) and (9)).

3.4 Towards Optimal Policy

3.4.1 Macroeconomic Effects of Progressive Taxation and Education Subsidies

The previous section has shown that in the competitive equilibrium without policy consumption of households without college is suboptimally low and consumption of college educated depends suboptimally strongly on their ability e . Investigating the household budget constraint (7) we observe that these two concerns can both be mitigated by implementing a lump-sum transfer $d > 0$ financed by a proportional labor income tax, $\tau > 0$. It thus might be part of the optimal policy mix. We now study the consequences of such a policy. In order to do so we note that the government budget constraint reads as (from now on suppressing the CE label whenever unnecessary):

$$d + \theta\kappa(1 - e^{CE}(\tau, \theta)) = \tau L^{CE}(\tau, \theta)/w \quad (14)$$

Recall that neither the education threshold nor aggregate output is a function of the lump-sum tax/subsidy d . This observation immediately results in the following:

Proposition 11 *An increase in lump-sum transfers d , financed by a raise in the income tax rate τ (that is, an increase in the progressivity of the tax code) leads to*

¹³ Of course there exist *nonutilitarian* welfare weights $\mu(e) \neq 1$ under which the socially optimal allocation arises as a competitive equilibrium without government intervention.

1. A decline in the fraction e^{CE} of households attending college.
2. A reduction in individual and aggregate labor supply and thus output L^{CE} .

Proof 12 Follows directly from the fact that $l_n^{CE}(\tau), l_c^{CE}(e; \tau), e^{CE}(\tau, \theta)$ and $L^{CE}(\tau, \theta)$ are all strictly decreasing in τ and are independent of d .

Thus a τ -financed increase in lump-sum transfers $d > 0$ improves the consumption distribution by redistributing towards n -households (and low e college educated households), but it reduces aggregate output through reducing labor supply of all households and lowering the share $1 - e^{CE}$ of households that become more productive through a college education. The latter concern can be offset through education subsidies, as the next proposition shows.

Proposition 13 An increase in college subsidies θ financed by a reduction in the transfers d leads to

1. An increase in the fraction of households attending college (e^{CE} decreases)
2. An increase in aggregate labor supply and thus output L^{CE} .

Proof 14 Follows directly from the fact that e^{CE} is strictly decreasing in θ and $L^{CE}(\tau, \theta)$ is strictly decreasing in e^{CE} (output increases with more households going to college) and is independent of d .

Note that positive education subsidies, when financed by labor income taxes, not only increase aggregate output, but also redistribute from high e -types to low e -college types, but redistribute away from the very low e -types that do not go to college, hence do not enjoy the subsidy but still bear part of the income tax burden. To summarize, in light of an inefficiently dispersed consumption distribution in the unregulated equilibrium (relative to the one chosen by the utilitarian social planner) the implementation of a progressive tax system improves on the consumption distribution, but lowers average consumption by creating disincentives to work and go to college. The latter distortion can be offset by an appropriate education subsidy. It therefore is to be expected that the optimal fiscal policy in this model may feature progressive income taxes ($\tau, d > 0$) and a positive education subsidy, $\theta > 0$. The next subsection will demonstrate that this is indeed the case, at least for a non-empty subset of the parameter space.

3.4.2 The Optimal Policy Mix

Given the full characterization of a competitive equilibrium for a given fiscal policy (τ, d, θ) in proposition , we can now state the optimal fiscal policy problem of the Ramsey government as

$$\max_{\tau, d, \theta} \left\{ e^{CE}(\tau, \theta) \log \left(\frac{[(1-\tau)w]^{1+\psi}}{(1+\psi)\mu^\psi} + dw \right) + \int_{e^{CE}(\tau, \theta)}^1 \log \left(\frac{[(1-\tau)(1+pe)w]^{1+\psi}}{(1+\psi)\mu^\psi} + dw + \kappa(1-\theta)w \right) de \right\}$$

subject to

$$d + \theta\kappa(1 - e^{CE}(\tau, \theta)) = \tau L(\tau, \theta)/w$$

with

$$e^{CE}(\tau, \theta) = \frac{\left(\frac{(1-\theta)\mu^\psi(1+\psi)\kappa}{(1-\tau)^{1+\psi}w^\psi} + 1 \right)^{\frac{1}{1+\psi}} - 1}{p}$$

$$L(\tau, \theta)/w = \left(\frac{(1-\tau)w}{\mu} \right)^\psi \left(e^{CE}(\tau, \theta) + \int_{e^{CE}(\tau, \theta)}^1 (1+pe)^{1+\psi} de \right).$$

Unfortunately a full analytical characterization of the solution to the Ramsey problem is infeasible even in the simple model. We therefore here present a simple quantitative example, employing the parameter values summarized in table 1.

Table 1: Parameter Values

Parameter	σ	μ	ψ	w	p	κ
Value	1	2	0.1	1	1	0.55

Table 2 displays characteristics of the socially optimal and equilibrium allocations, the latter for three policy configurations. The second row displays equilibrium outcomes without government intervention, the third row for a restricted optimal policy where the education subsidy θ is constrained to equal zero. Finally, the last row summarizes the equilibrium under the optimal fiscal policy. Figure 2 below plots social welfare over the relevant range of fiscal policies, with τ and θ on the axes, and d adjusted to balance the government budget.

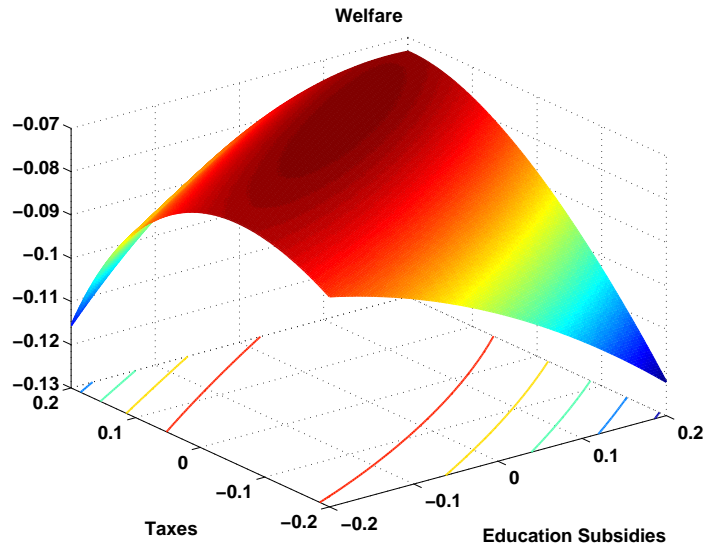
From table 2 we observe that the optimal policy indeed calls for progressive taxes and education subsidies. Comparing rows 1 and 2 shows, as proved above that the competitive equilibrium without government intervention has optimal labor supply, education and production allocations, but a consumption distribution in which low e types consume too little. Introducing a progressive income tax in row 3 leads to an improvement in that distribution, but at the expense of reduced output and a smaller fraction of households attending college. A positive (but quantitatively small) education subsidy raises this share, but requires extra government revenue and thus an even higher

Table 2: Results for Example

	τ	d	θ	e^*	L	l_n	$l_c(e^*)$	c_n	$c_c(e^*)$
SP	N/A	N/A	N/A	0.58	1.29	0.93	0.98	1.02	1.08
CE	0	0	0	0.58	1.29	0.93	0.98	0.93	0.98
CE	6.4%	8%	0	0.62	1.25	0.93	0.97	0.95	1.00
CE	9%	10.4%	4%	0.61	1.25	0.92	0.97	0.95	1.00

labor income tax rate τ . The resulting consumption distribution implied by the optimal fiscal policy is plotted in figure 1, alongside the socially optimal and the laissez faire allocation.

Figure 2: Welfare as a Function of τ and θ



3.5 Summary

To conclude this section, we have developed a simple model with risk-averse households that make endogenous labor supply and education decisions and used it to argue that progressive income taxes and positive education subsidies are part of an optimal second best policy mix, in the presence of incomplete insurance markets against income risk. However, we have abstracted from two key features that will be present in our quantitative analysis.

First, wages and thus the college wage premium were exogenously given and thus invariant to fiscal policy. If wages respond, in general equilibrium, to the change in the share of college-educated workers (as in the version of the model with imperfect

substitutability of skilled and unskilled labor), then education policies might be less potent in stimulating college attendance, but might have beneficial redistributive side effects. Second, the simple model ignored dynamics, and thus the insight that changing the skill distribution of the workforce takes time, with increased schooling costs materializing immediately but benefits accruing only later.

Our quantitative analysis with general equilibrium wage effects and transitional dynamics will uncover how these model elements shape the optimal degree of income tax progressivity and magnitude of optimal education subsidies.

4 The Quantitative Model

4.1 Demographics

Population grows at the exogenous rate χ . We assume that parents give birth to children at the age of j_f and denote the fertility rate of households by f , assumed to be the same across education groups.¹⁴ Notice that f is also the number of children per household. Further, let φ_j be the age-specific survival rate. We assume that $\varphi_j = 1$ for all $j = 0, \dots, j_r - 1$ and $0 < \varphi_j \leq 1$ for all $j = j_r, \dots, J - 1$, where j_r is the fixed retirement age ($j_r - 1$ is the last working age before retirement) and J denotes the maximum age (hence $\varphi_J = 0$). The population dynamics are then given by

$$N_{t+1,0} = f \cdot N_{t,j_f} \quad (15)$$

$$N_{t+1,j+1} = \varphi_j \cdot N_{t,j}, \quad \text{for } j = 0, \dots, J. \quad (16)$$

Observe that the population growth rate is then given by

$$\chi = f^{j_f+1} - 1. \quad (17)$$

4.2 Technology

We refer to workers that have completed college as skilled, the others as unskilled. Thus the skill level s of a worker falls into the set $s \in \{n, c\}$ where $s = c$ denotes college educated individuals. We assume that skilled and unskilled labor are imperfectly substitutable in production (see Katz and Murphy (1992) and Borjas (2003)) but that within skill groups labor is perfectly substitutable across different ages. Let $L_{t,s}$ denote aggregate labor of skill s , measured in efficiency units and let K_t denote the capital stock.

¹⁴ Note that due to the endogeneity of the education decision in the model, if we were to allow differences in the age at which households with different education groups have children, it is hard to assure that the model has a stationary joint distribution over age and skills.

Total labor efficiency units at time t , aggregated across both education groups, is then given by

$$L_t = \left(L_{t,n}^\rho + L_{t,c}^\rho \right)^{\frac{1}{\rho}} \quad (18)$$

where $\frac{1}{1-\rho}$ is the elasticity of substitution between skilled and unskilled labor.¹⁵ Note that as long as $\rho < 1$, skilled and unskilled labor are imperfect substitutes in production, and the college wage premium is not constant, but will endogenously respond to changes in government policy.

Aggregate labor is combined with capital to produce output Y_t according to a standard Cobb-Douglas production function

$$Y_t = F(K_t, L_t) = K_t^\alpha L_t^{1-\alpha} = K_t^\alpha \left[\left(L_{t,n}^\rho + L_{t,c}^\rho \right)^{\frac{1}{\rho}} \right]^{1-\alpha} \quad (19)$$

where α measures the elasticity of output with respect to the input of capital services.

As always, perfect competition among firms and constant returns to scale in the production function implies zero profits for all firms at all t , and an indeterminate size distribution of firms. Thus 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.

This representative firm rents capital and hires the two skill types of labor on competitive spot markets at prices $r_t + \delta$ and $w_{t,s}$, where r_t is the interest rate, δ the depreciation rate of capital and $w_{t,s}$ is the wage rate per unit of labor of skill s . Furthermore, denote by $k_t = \frac{K_t}{L_t}$ the “capital intensity”—defined as the ratio of capital to the CES aggregate of labor. Profit maximization of firms implies the standard conditions

$$r_t = \alpha k_t^{\alpha-1} - \delta \quad (20)$$

$$w_{t,n} = (1 - \alpha) k_t^\alpha \left(\frac{L_t}{L_{t,n}} \right)^{1-\rho} = \omega_t \left(\frac{L_t}{L_{t,n}} \right)^{1-\rho} \quad (21)$$

$$w_{t,c} = (1 - \alpha) k_t^\alpha \left(\frac{L_t}{L_{t,c}} \right)^{1-\rho} = \omega_t \left(\frac{L_t}{L_{t,c}} \right)^{1-\rho}, \quad (22)$$

where $\omega_t = (1 - \alpha) k_t^\alpha$ is the marginal product of total aggregate labor L_t . The college wage premium is then given by

$$\frac{w_{t,c}}{w_{t,n}} = \left(\frac{L_{t,n}}{L_{t,c}} \right)^{1-\rho} \quad (23)$$

and depends on the relative supplies of non-college to college labor (unless $\rho = 1$) and the elasticity of substitution between the two types of skills, and thus is endogenous in our model.

¹⁵ Katz and Murphy (1992) report an elasticity of substitution across education groups of $\sigma = 1.4$. This is also what Borjas (2003) finds, using a different methodology and dataset.

4.3 Household Preferences and Endowments

4.3.1 Preferences

Households are born at age $j = 0$ and form independent households at age j_a , standing in for age 18 in real time. Households give birth at the age j_f and children live with adult households until they form their own households. Hence for ages $j = j_f, \dots, j_f + j_a - 1$ children are present in the parental household. Parents derive utility from per capita consumption of all household members and leisure that are representable by a standard time-separable expected lifetime utility function

$$E_{j_a} \sum_{j=j_a}^J \beta^{j-j_a} u \left(\frac{c_j}{1 + \mathbf{1}_{\mathcal{J}_s} \zeta_f}, \ell_j \right) \quad (24)$$

where c_j is total consumption, ℓ_j is leisure and $\mathbf{1}_{\mathcal{J}_s}$ is an indicator function taking the value one during the period when children are living in the respective household, that is, for $j \in \mathcal{J}_s = [j_f, j_f + j_a - 1]$, and zero otherwise. $0 \leq \zeta \leq 1$ is an adult equivalence parameter. Expectations in the above are taken with respect to the stochastic processes governing mortality and labor productivity risk as well as with respect to survival risk.

We model an additional form of altruism of households towards their children. At parental age j_f , when children leave the house, the children's' expected lifetime utility enters the parental lifetime utility function with a weight $v\beta^{j_f}$, where the term β^{j_f} simply reflects the fact that children's' lifetime utility enters parental lifetime utility at age j_f , and the parameter v measures the strength of parental altruism.¹⁶

4.3.2 Initial Endowments and Human Capital Accumulation Technology

At age $j = j_a$, before any decision is made, households draw their innate ability to go to college, $e \in \{e_1, e_2, \dots, e_N\}$ according to a distribution $\pi(e|\cdot)$ that may depend on the characteristics of their parents, including parental education s_p and parental labor productivity to be described below.¹⁷ Innate ability also affects future wages directly and independent of education, in a stochastic way, also described below. A young household with ability e incurs a per-period resource cost of going to college $w_{t,c}\kappa$ that

¹⁶ Evidently the exact timing when children lifetime utility enters that of their parents is inconsequential. We can simply rescale v to offset changes in the time discount factor β^{j_f} and leave the effective degree of altruism $v\beta^{j_f}$ unchanged. Similarly, the parameter v captures the utility parents receive from all of their f (identical) children. One could write $v = \tilde{v}f$, where \tilde{v} is per-child altruism factor, but this of course leaves both the dynamic programming problem as well as the calibration of the model unchanged (since \tilde{v} would turn out to equal v/f in our calibration).

¹⁷ Ability e in our model does not only capture innate ability in the real world since it also stands in for all characteristics of the individual at the age of the college decision, that is, everything learned in primary and secondary education. In our model one of the benefits of going to college is to be able to raise children that will (probabilistically) be more able to go to college.

is proportional to the aggregate wage of the high-skilled, $w_{t,c}$.¹⁸ In case the government chooses to implement education subsidies, a fraction θ_t of the resource cost is borne by the government. In addition, a constant fraction θ_{pr} of the education costs is borne by private subsidies, paid from accidental requests described below. We think of θ_{pr} as a policy invariant parameter to be calibrated, and introduce it to capture the fact that, empirically, a significant share of university funding comes from alumni donations and support by private foundations.

Going to college also requires a fraction $\xi(e) \in [0, 1]$ of time at age j_a , in the period in which the household attends school.¹⁹ The dependence of the time cost function ξ on innate ability to go to college reflects the assumption that more able people require less time to learn and thus can enjoy more leisure time or work longer hours while attending college (the alternative uses of an individual's time).²⁰ A household that completed college has skill $s = c$, a household that did not has skill $s = n$.

Households 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.²¹ Parents make these transfers, assumed to be noncontingent on the child's education decision²², at their age j_f , after having observed their child's ability draw e . This transfer is restricted to be nonnegative. In addition to this one-time intentional intergenerational transfer b , all households receive transfers from accidental bequests. We assume the assets of households that die at age j are redistributed uniformly across all households of age

¹⁸ Instead of a monetary cost Abbott et al. (2013) use a "psychic stress" formulation of costs based on Heckman, Lochner, Todd (2006). Our specification is closer to Caucutt et al. (2006) where the costs stand in for hiring a teacher to acquire education.

¹⁹ In the quantitative implementation of the model a period will last four years, and thus households attend college for one model period.

²⁰ With this time cost we also capture utility losses of poorer households who have to work part-time to finance their college education.

²¹ This is similar to Abbott et al. (2013). We model this as a one time payment only. The transfer payment captures the idea that parents finance part of the higher education of their children. Our simplifying assumptions of modeling these transfers are a compromise between incorporating directed inter-generational transfers of monetary wealth in the model and computational feasibility.

If we were to model flexible inter-vivo transfers at all ages $j = j_f, \dots, j_f + j_c$, we would have to deal with two continuous state variables. Both their own as well as their parents' assets would be relevant for children's decisions at all ages $j = j_a, \dots, j_f$. An additional continuous state variable is also required if we were to assume that parents commit to pay constant transfers b at all ages $j_f, \dots, j_f + j_c$ which would perhaps have a more realistic flavor than assuming a one-time transfer. During those years b is a state variable for the children's' problem. Note that if parental borrowing constraints are not binding one-time transfers are equivalent to a commitment to transfers for many periods (as long as the contingency of parental death is appropriately insured). Thus the issue whether our assumption is quantitatively important depends on the specification of the borrowing constraint, and, given this specification, whether the constraint often binds for households at age j_f .

²² Note that parents of course understand whether, given b , children will go to college or not, and thus can affect this choice by giving a particular b .

$j - j_f$, that is, among the age cohort of their children. Let these age dependent transfers be denoted by $Tr_{t,j}$

4.3.3 Labor Productivity

In each period of their lives households are endowed with one unit of productive time. A household of age j with skill $s \in \{n, c\}$ earns a wage

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

per unit of time worked. Wages depend on a deterministic age profile $\epsilon_{j,s}$ that differs across education groups, on the skill-specific average wage $w_{t,s}$, a fixed effect $\gamma \in \Gamma_s = \{\gamma_{l,s}, \gamma_{h,s}\}$ that spreads out wages within each education group and remains constant over the life cycle, and an idiosyncratic stochastic shock η . The probability of drawing the high fixed effect prior to labor market entry is a function of the ability of the household, and denoted by $\pi_s(\gamma|e)$. The stochastic shock η is mean-reverting and follows an education-specific Markov chain with states $\mathcal{E}_s = \{\eta_{s1}, \dots, \eta_{sM}\}$ and transitions $\pi_s(\eta'|\eta) > 0$. Let Π_s denote the invariant distribution associated with π_s . Prior to making the education decision a household's idiosyncratic shock η is drawn from Π_η . We defer a detailed description of the exact forms for $\pi_s(\gamma|e)$ and $\pi_s(\eta'|\eta)$ to the calibration section.²³ Thus at the beginning of every period in working life the individual state variables of the household include (j, γ, s, η, a) , the household's age j , fixed effect γ , education s , stochastic labor productivity shock η and assets a .

4.4 Market Structure

We assume that financial markets are incomplete in that there is no insurance available against idiosyncratic mortality and labor productivity shocks. Households can self-insure against this risk by accumulating a risk-free one-period bond that pays a real interest rate of r_t . In equilibrium the total net supply of this bond equals the capital stock K_t in the economy, plus the stock of outstanding government debt B_t .

Furthermore, we severely restrict the use of credit to self-insure against idiosyncratic labor productivity and thus income shocks by imposing a strict credit limit. The only

²³ The purpose of introducing the fixed effect γ instead of making wages directly depend on ability e is mainly computational (although we think it is plausible to make ability to succeed in college and ability in the labor market imperfectly correlated).

In order to permit the share of households that go to college to vary smoothly with economic policy it is important that the set $\{e_1, e_2, \dots, e_N\}$ is sufficiently large. On the other hand, given the large state space for households of working age keeping track of the state variable e is costly; by stochastically mapping e into the fixed effect γ after the education decision and restricting γ to take only two possible values (for each education group) reduces this burden significantly.

borrowing we permit is to finance a college education through student loans. Households that borrow to pay for college tuition and consumption while in college face age-dependent borrowing limits of $\underline{A}_{j,t}$ (whose size depends on the degree to which the government subsidizes education) and also face the constraint that their balance of outstanding student loans cannot increase after they have completed school. This assumption rules out that student loans are used for general consumption smoothing over the life cycle.

The constraints $\underline{A}_{j,t}$ are set such that student loans need to be fully repaid by age j_r at which early mortality sets in. This insures that households can never die in debt and we do not need to consider the possibility and consequences of personal bankruptcy. Beyond student loans we rule out borrowing altogether. This, among other things, implies that households without a college degree can never borrow

4.5 Government Policies

The government needs to finance an exogenous stream G_t of non-education expenditures and an endogenous stream E_t of education expenditures. It can do so by issuing government debt B_t , by levying linear consumption taxes $\tau_{c,t}$ and income taxes $T_t(y_t)$ which are not restricted to be linear. The initial stock of government debt B_0 is given. We restrict attention to a tax system that discriminates between the sources of income (capital versus labor income), taxes capital income $r_t a_t$ at the constant rate $\tau_{k,t}$, but permits labor income taxes to be progressive or regressive. We will take consumption and capital income tax rates $\tau_{c,t}$, $\tau_{k,t}$ as exogenously given, but will optimize over labor income tax schedules within a simple parametric class.

Specifically, the total amount of labor income taxes paid takes the following simple linear form

$$T_t(y_t) = \max \left\{ 0, \tau_{l,t} \left(y_t - d_t \frac{Y_t}{N_t} \right) \right\} \quad (25)$$

$$= \max \{ 0, \tau_{l,t} (y_t - Z_t) \} \quad (26)$$

where y_t is household taxable labor income, $\frac{Y_t}{N_t}$ is 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. Note that the tax system is potentially progressive (if $d_t > 0$) or regressive (if $d_t < 0$). Therefore for every period there are two policy parameters on the tax side, $(\tau_{l,t}, d_t)$.

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

$$G_t = gy \cdot Y_t$$

where the share of output $gy = \frac{G_t}{Y_t}$ commanded by the government is a parameter to

be calibrated from the data.²⁴

In addition, the government administers a pure pay-as-you-go social security system that collects payroll taxes $\tau_{ss,t}$ and pays benefits $p_{t,j}(\gamma, s)$, which depend on the wages a household has earned during her working years, and thus on her characteristics (γ, s) as well as on the time period in which the household retired (which, given today's date t can be inferred from the current age j of the household). In the calibration section we describe how we approximate the current U.S. system with its progressive benefit schedule through the function $p_{t,j}(\gamma, s)$. Since we are interested in the optimal progressivity of the income tax schedule *given* the current social security system it is important to get the progressivity of the latter right, in order to not bias our conclusion about the desired progressivity of income taxes. In addition, the introduction of social security is helpful to obtain more realistic life cycle saving profiles and an empirically more plausible wealth distribution.

Since the part of labor income that is paid by the employer as social security 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

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

4.6 Competitive Equilibrium

We deal with time sequentially, both in our specification of the model as well as in its computation. For a given time path of prices and policies it is easiest to formulate the household problem recursively, however. In order to do so for the different stages of life we first collect the key decisions and state variables in a time line.

4.6.1 Time Line

1. Newborn individuals are economically inactive but affect parental utility until they form a new household at age j_a .
2. At age j_a a new adult household forms. Initial state variables are age $j = j_a$, parental education s_p and parental productivity γ_p , own education $s = n$ (the household does not have a college degree before having gone to college). 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 and thus immediately constitute the initial endowment of assets (generically denoted by a) for other ages.

²⁴ Once we turn to the determination of optimal tax and subsidy policies we will treat G rather than gy as constant. A change in policy changes output Y_t and by holding G fixed we assume that the government does not respond to the change in tax revenues by adjusting government spending (if we held gy constant it would).

Then initial idiosyncratic labor productivity η is drawn according to Π_n . Thus the state of a household prior to the college decision is $z = (j_a, e, s = n, \eta, a = b/(1 + r(1 - \tau_k)))$.²⁵

3. Given state z , at age j_a the educational decision is made. If a household decides to go to college, she immediately does so at age j_a , and her education state switches to $s = c$ at that age. Then households draw their labor productivity fixed effect γ from the education- and ability-contingent distribution $\pi(\gamma|s, e)$.
4. At age j_a , but after the education decision has been made, the household problem differs between non-college and college households since the latter need to spend time and resources on college. A household that goes to college but works part time does so for non-college wages:

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

where η is drawn as described above. Observe that γ is fixed whereas η is drawn from the non-college distribution. At the end of the college period j_a the idiosyncratic shock η of college-bound households is re-drawn from the college distribution Π_c and now evolves according to $\pi_c(\eta'|\eta)$ for those with $s = c$. Furthermore college-educated households draw their fixed effect from the distribution $\pi(\gamma|c, e)$ prior to entering the labor market.

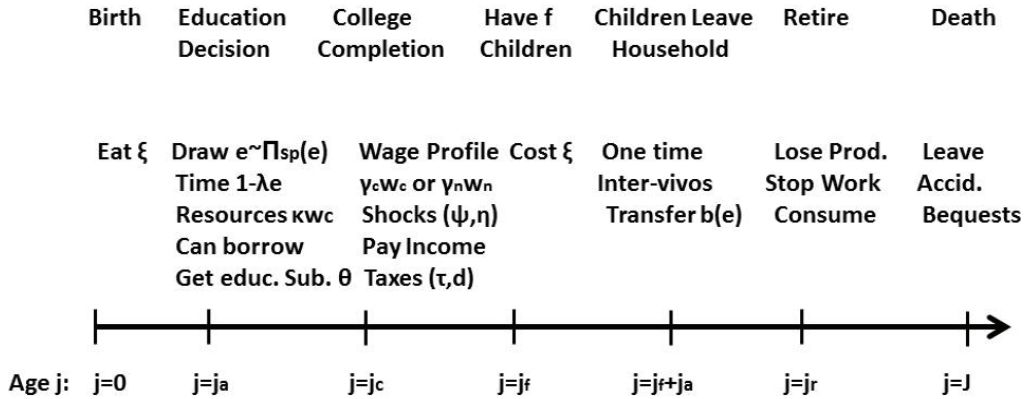
5. Ages $j_a + 1, \dots, j_f - 1$: Between age of $j_f - 1$ and j_f the decision problem changes because children now enter the utility function and households maximize over per capita consumption $c_j/(1 + \zeta_f)$.
6. Ages $j_a + j_f, \dots, j_a + j_f - 1$: Between age of $j_a + j_f - 1$ and $j_a + j_f$ the decision problem changes again since at age $j_a + j_f$ 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 : Households make transfers b to their children conditional on observing the skills e of their children.
8. Age $j_a + j_f + 1, \dots, j_r - 1$: Only utility from own consumption and leisure enters the lifetime utility at these ages. Labor productivity falls to zero at retirement which is at age j_r .
9. Ages $j = j_r, \dots, J$: Households are now in retirement and only earn income from capital and from social security benefits $p_{t,j}(e, s)$.

The key features of this time line are summarized in figure 3.

²⁵ For all ages $j > j_a$ assets a brought into the period generate gross revenue $(1 + r(1 - \tau_k))a$. Given our timing assumption inter-vivo transfers b generate gross revenue of b . Thus the initial asset state of households of age j_a is $a = b/(1 + r(1 - \tau_k))$.

Figure 3: Time Line in the Model

Life Cycle



4.6.2 Recursive Problems of Households

We now spell out the dynamic household problems at the different stages in the life cycle recursively.

Child at $j = 0, \dots, j_a - 1$ Children live with their parents and command resources, but do not make own economic decisions.

Education decision at j_a Before households make the education decision households draw ability e , their initial labor productivity η and receive inter-vivos transfers b . We specify an indicator function for the education decision as $\mathbf{1}_s = \mathbf{1}_s(e, \eta, b)$, where a value of 1 indicates the household goes to college. Recall that households, as initial condition, are not educated in the first period, $s = n$ and that age is $j = j_a$. The education decision solves

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

where $V_t(j_a, e, s, \eta, b / (1 + r(1 - \tau_k)))$ is the lifetime utility at age $j = j_a$, conditional on having chosen (but not necessarily completed) education $s \in \{n, c\}$. It is formally

given by

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

where $V_t(j_a, e, \gamma, s, \eta, b / (1 + r(1 - \tau_k)))$ is defined below and is the value function at age j_a after the fixed effect has been drawn from $\pi_s(\gamma | e)$.

Problem at $j = j_a$ After having made the education decision at age j_a and having drawn the fixed effect γ households choose how much to work, how much to consume and how much to save. The dynamic programming problem of college-bound and non-college bound households differs. Households first draw the fixed effect γ from distribution $\pi(\gamma | s, e)$ and then solve

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

subject to the budget constraint²⁶

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

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

Note that ability e is a redundant state variable for non-college bound households at age j_a , but not for households going to college, since the time loss for doing so still depends on e . It does become a redundant state variable at age $j_a + 1$ and thus does not appear on the right hand side of the Bellman equation above.²⁷

Problem at $j_a + 1, \dots, j_f - 1$ At these ages education is completed, thus no time and resource cost for education is being incurred. The problem reads as

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

subject to the budget constraint

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

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

²⁶ At age j_a assets a equal to the transfers b from parents. Since these enter the budget constraint of children in the period they are given, for j_a the first term on the right hand side of the budget constraint reads as

$$b + (1 + (1 - \tau_{k,t})r_t)Tr_{t,j}.$$

²⁷ Furthermore we slightly abused notation in that for college-bound households η' at age j_a is drawn from $\Pi_c(\eta')$ rather than $\pi_c(\eta' | \eta)$.

Problem at ages $j_f, \dots, j_f + j_a - 1$ At these ages children live with the household and thus resource costs of children are being incurred. The problem reads as

$$V_t(j, \gamma, s, \eta, a) = \max_{\substack{c, l \in [0, 1] \\ a' \geq -\mathbf{1}_s \Delta_{j,t}}} \left\{ u \left(\frac{c}{1 + \zeta_f}, 1 - l \right) + \beta \varphi_j \sum_{\eta'} \pi_s(\eta' | \eta) V_{t+1}(j + 1, \gamma, s, \eta', a') \right\}$$

subject to the budget constraint

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

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

Problem at $j_f + j_a$ This is the age of the household where children leave the home, parents give them an inter-vivos transfer b and the children's lifetime utility enters that of their parents. The dynamic problem becomes

$$\begin{aligned} V_t(j, \gamma, s, \eta, a) = & \max_{\substack{c(e'), l(e') \in [0, 1], b(e') \geq 0 \\ a'(e') \geq -\mathbf{1}_s \Delta_{j,t}}} \sum_{e'} \pi(e' | s, \gamma) \{ u(c(e'), 1 - l(e')) \\ & + \beta \varphi_j \sum_{\eta'} \pi_s(\eta' | \eta) V_{t+1}(j + 1, \gamma, s, \eta', a'(e')) \\ & + v \sum_{\eta'} \Pi_n(\eta') \max \left[V_t \left(j_a, e', n, \eta', \frac{b(e')}{1 + r(1 - \tau_k)} \right), V_t \left(j_a, e', c, \eta', \frac{b(e')}{1 + r(1 - \tau_k)} \right) \right] \} \end{aligned}$$

subject to

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

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

Note that since parents can observe the ability of their children e' before giving the transfer, the transfer b (and thus all other choices in that period) are contingent on e' . Also notice that all children in the household are identical. Since parents do not observe the initial labor productivity of their children, parental choices cannot be made contingent on it, and expectations over η' have to be taken in the Bellman equation of the parents over the lifetime utility of their children.²⁸

²⁸ Note that we make parents choose their transfers noncontingent on the schooling choice of their children. Mechanically it is no harder to let this choice be contingent on the schooling choice (it then simply would be two numbers). Note that permitting such contingency affects choices, since making transfers contingent permits parents to implicitly provide better insurance against η -risk. If the transfers also could be conditioned on η , then we conjecture that it does not matter whether they in addition are made contingent on the education decision of the children or not. Note that in any case, parents can fully think through what transfer induced what education decision.

Problem at $j_f + j_a + 1, \dots, j_r - 1$ Now children have left the household, and the decision problem exactly mimics that in ages $j \in \{j_c + 1, \dots, j_f - 1\}$. Observe 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 $j_f + j_a$.

Problem at j_r, \dots, J Finally, in retirement households have no labor income (and consequently no labor income risk). Thus the maximization problem is given by

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

subject to the budget constraint

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

4.7 Definition of Equilibrium

Let $\Phi_{t,j}(\gamma, s, \eta, a)$ denote the share of agents, at time t of age j with characteristics (γ, s, η, a) .²⁹ For each t and j we have $\int d\Phi_{t,j} = 1$

Definition 15 *Given an initial capital stock K_0 , initial government debt level B_0 and initial measures $\{\Phi_{0,j}\}_{j=0}^J$ of households, and given a stream of government spending $\{G_t\}$, a competitive equilibrium is sequences of household value and policy functions $\{V_t, a'_t, c_t, l_t, \mathbf{1}_{s,t}, b_t\}_{t=0}^\infty$, production plans $\{Y_t, K_t, L_{t,n}, L_{t,c}\}_{t=0}^\infty$, sequences of tax policies, education policies, social security policies and government debt levels $\{T_t, \tau_{l,t}, \tau_{c,t}, \theta_t, \tau_{ss,t}, p_{t,j}, (\cdot), B_t\}_{t=0}^\infty$, sequences of prices $\{w_{t,n}, w_{t,c}, r_t\}_{t=0}^\infty$, sequences of transfers $\{Tr_{t,j}\}_{t=0}^\infty$ and sequences of measures $\{\Phi_{t,j}\}_{t=1}^\infty$ such that*

1. *Given prices, transfers and policies, $\{V_t\}$ solve the Bellman equations described in subsection 4.6.2 and $\{V_t, a'_t, c_t, l_t, \mathbf{1}_{s,t}, b_t\}$ are the associated policy functions.*
2. *Interest rates and wages satisfy (20).*
3. *Transfers are given by*

$$Tr_{t+1, j-j_f+1} = \frac{N_{t,j}}{N_{t+1, j-j_f+1}} \int (1 - \varphi_j) a'_t(j, \gamma, s, \eta, a) d\Phi_{t,j} - \frac{1}{\sum_{\ell=j_f}^J N_{t+1, \ell-j_f+1}} \frac{PE_{t+1}}{1 + r_{t+1}(1 - \tau_{t+1}^k)} \quad (27)$$

²⁹ For age j_a and $s = c$ the state space also includes the ability e of the household, but not the fixed effect γ . To simplify notation in the definition below we keep this case distinction implicit whenever there is no room for confusion.

for all $j \geq j_f$, where private aggregate education subsidies are given by

$$PE_{t+1} = \theta_{pr} \kappa \omega_{t+1,c} N_{t+1,j_a} \int_{\{(e,s,\eta,a):s=c\}} d\Phi_{t+1,j_a} \quad (28)$$

4. Government policies satisfy the government budget constraints

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

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

where, for each household, taxable income y_t was defined in the recursive problems in subsection 4.6.2 and aggregate consumption and government education expenditures are given by

$$E_t = \theta_t \kappa \omega_{t,c} N_{t,j_a} \int_{\{(e,\gamma,s,\eta,a):s=c\}} d\Phi_{t,j_a} \quad (29)$$

$$C_t = \sum_j N_{t,j} \int c_t(j, \gamma, s, \eta, a) d\Phi_{t,j} \quad (30)$$

5. Markets clear in all periods t

$$L_{t,s} = \sum_j N_{t,j} \int \epsilon_{j,s} \gamma \eta l_t(j, \gamma, s, \eta, a) d\Phi_{t,j} \text{ for } s \in n, c \quad (31)$$

$$K_{t+1} + B_{t+1} = \sum_j N_{t,j} \int a'_t(j, \gamma, s, \eta, a) d\Phi_{t,j} \quad (32)$$

$$K_{t+1} = Y_t + (1 - \delta) K_t - C_t - CE_t - G_t - E_t. \quad (33)$$

where Y_t is given by (19) and it is understood that the integration in (31) is only over individuals with skill s . Also

$$CE_t = (1 - \theta_t) \kappa \omega_{t,c} N_{t,j_a} \int_{\{(e,s,\eta,a):s=c\}} d\Phi_{t,j_a} \quad (34)$$

is aggregate private spending on education.

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

The law of motion for the measures explicitly states as follows. Define the Markov transition function at time t for age j as

$$Q_{t,j}((\gamma, s, \eta, a), (\Gamma \times \mathcal{S} \times \mathcal{E} \times \mathcal{A})) =$$

$$\begin{cases} \sum_{\eta' \in \mathcal{E}} \pi_s(\eta' | \eta) & \text{if } \gamma \in \Gamma, s \in \mathcal{S}, \text{ and } a'_t(j, \gamma, s, \eta, a) \in \mathcal{A} \\ 0 & \text{else} \end{cases}$$

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 education type, and \mathcal{A} does not include the optimal asset choice.³⁰ If it does, then the transition probability is purely governed by the stochastic shock process for η .

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

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

The initial measure over types at age $j = j_a$ (after the college decision has been made) is more complicated.³¹ Households start with assets equal to bequests from their parents determined by the bequest function b_t , draw initial mean reverting productivity according to $\Pi_n(\eta')$, determine education 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 $\pi(\gamma' | s, e')$:

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

Definition 16 *A stationary equilibrium is a competitive equilibrium in which all individual functions and all aggregate variables are constant over time.*

³⁰ There is one exception: at age $j = j_c$ college-educated households 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

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

³¹ Part of the complication is that at age j_a the individual state space includes ability e which then becomes a redundant state variable. 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_a} from age j_a to age $j_a + 1$ (and only at this age) has as first argument (e, γ, s, η, a) .

5 Thought Experiment

5.1 Social Welfare Function

The social welfare function is Utilitarian for people initially alive

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

where $V_1(\cdot; \mathcal{T})$ 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 households in the stationary equilibrium under the status quo policy.³²

5.2 Optimal Tax System

In our optimal policy analysis we hold constant the capital income and consumption tax rate as well as the pension contribution rate 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 and a cross-section of households Φ_0 determined by a stationary (to be calibrated) policy $\tau_{l,0}, \theta_0, d_0, b_0 = B_0/Y_0$, the optimal tax reform is defined as the sequence $\mathcal{T}^* = \{\tau_{l,t}, \theta_t, d_t, B_t\}_{t=1}^{\infty}$ that maximizes the social welfare function, i.e. that solves

$$\mathcal{T}^* \in \arg \max_{\mathcal{T} \in \Gamma} SWF(\mathcal{T}).$$

Here Γ is the set of policies for which an associated competitive equilibrium exists. Unfortunately the set Γ 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

$$\begin{aligned} \tau_{l,t} &= \tau_{l,1} \\ \theta_t &= \theta_1 \\ d_t &= d_1 \end{aligned}$$

for all $t \geq 1$. Note that the associated debt to GDP ratio will of course not be constant over time. Since all admissible policies defined by $(\tau_{l,1}, \theta_1, d_1)$ have to lie in Γ , from the definition of equilibrium 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 $(\tau_{l,1}, \theta_1, d_1)$ over which the optimization of the social welfare function is carried out.

³² Note that future generations' lifetime utilities are implicitly valued through the value functions of their parents. Of course there is nothing wrong in principle to additionally include future generations' lifetime utility in the social welfare function with some weight, but this adds additional free parameters (the social welfare weights).

6 Calibration

6.1 Demographics

We take survival probabilities from the Social Security Administration life tables. The total fertility rate f in the economy is assumed to be $f = 1.14$, reflecting the fact that a mother on average has about $2f = 2.28$ children. This number also determines the population growth rate, cf. equation (17). Each period in the model has a length of four years. Children are born with age 0 and form households at biological age 18. We discard the first two years of childhood and accordingly set $j_a = \frac{18-2}{4} = 4$. Households require 4 actual years to complete a college education and therefore exit college at model 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 is the last working period of life), hence $j_r = 16$. The maximum life span is 101 years, i.e., the last period households are alive is biological age bin 98-101 and accordingly $J = 24$.

6.2 Labor Productivity Process

Recall that a household of age j with education $s \in \{n, c\}$, fixed effect γ and idiosyncratic shock η earns a wage of

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

where w_s is the skill-specific wage per labor efficiency unit in period t .

We estimate the deterministic, age- and education-specific component of labor productivity $\{\epsilon_{j,s}\}$ from PSID data (cf. Ludwig, Schelkle and Vogel 2012) and normalize the mean productivity at the age of college completion, $j_c = j_a + 1$, for $s = n$ to $\epsilon_{j_c,n} = 1$. The estimated profile $\epsilon_{j,c}$ is scaled up by a fixed constant, ϵ , such that the average college wage premium in the model of 80%, in line with U.S. data for the later part of the 2000's (see, e.g., Heathcote et al. 2010).

We choose the Markov chain driving the stochastic mean reverting component of wages η as a two state Markov chain with education-specific states for log-wages $\{-\sigma_s, \sigma_s\}$ and transition matrix

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

In order to parameterize this Markov chain we first estimate the following process on the education-specific PSID samples selected by Karahan and Ozkan (2012):

$$\begin{aligned} \log w_t &= \alpha + z_t \\ z_t &= \rho z_{t-1} + \eta_t, \end{aligned}$$

where α is an individual-specific fixed effect that is assumed to be normally distributed

(with cross-sectional variance σ_α^2). The estimation results are summarized in table 3.³³

Table 3: Estimates for Earnings Process

Group	ϱ	σ_η^2	σ_α^2
College	0.969	0.0100	0.0474
Non-College	0.928	0.0192	0.0644

For each education group we choose the two numbers (π_s, σ_s) such that the two-state Markov chain for wages we use has exactly the same persistence and conditional variance as the AR(1) process estimated above.³⁴ This yields parameter choices given in table 4.

Table 4: Markov Chain for Wages

Group	π_s	σ_s	\mathcal{E}_s
College	0.9408	0.191	{0.8113, 1.1887}
No College	0.8713	0.250	{0.7555, 1.2445}

After de-logging, the wage states were normalized so that the mean of the stochastic component of wages equals 1. We observe that college educated agents face somewhat smaller wage shocks, but that these shocks are slightly more persistent than for non-college educated households.

This leaves us with 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 wage component so that our model under the status quo policy matches selected wage or earnings observations from the data. We assume that

$$\begin{aligned}\pi(\gamma = \gamma_h | e, s = c) &= e \\ \pi(\gamma = \gamma_h | e, s = n) &= \nu \cdot e\end{aligned}$$

and, of course, $\pi(\gamma = \gamma_l | e, s) = 1 - \pi(\gamma = \gamma_h | e, s)$. Here ν is a parameter. The distribution of e itself is discussed in subsection 6.7. Note that since the γ 's are education

³³ For the details of the sample selection we refer the reader to Karahan and Ozkan (2012) and we thank the authors for providing us with the estimates for the process specified in the main text. In their paper they estimate a richer stochastic process (which, if implemented in our framework, would lead to at least one additional state variable).

³⁴ The (unconditional) persistence of the AR(1) process is given by ϱ and the conditional variance by σ_η^2 whereas the corresponding statistics for the Markov chain read as $2\pi_s - 1$ and σ_s^2 , respectively.

For a model where a period lasts 4 years and the AR(1) process is estimated on yearly data, the corresponding statistics are ϱ^4 and $(1 + \varrho^2 + \varrho^4 + \varrho^6)\sigma_\eta^2$.

specific and the probability of drawing the (education-specific) high $\gamma_{h,s}$ is a function of ability e , the wage benefits of going to college will be ability-specific as well. In appendix B we show that this premium is lower for the marginal college attendee (and thus for the households drawn newly into college by an increase in the college subsidy θ) than the average college wage premium.

The parameters $\{\gamma_{l,n}, \gamma_{h,n}, \gamma_{l,c}, \gamma_{h,c}, \nu\}$ are chosen jointly such that the stationary equilibrium of the status quo economy attains the following targets:

- Normalizations: the average γ is equal to one for each $s \in \{n, c\}$ [2 targets].
- The estimated variances of the fixed effect for both education groups σ_α^2 displayed in the last column of table 3. Note that the variances in the model are determined, for each s , by the spread between $\gamma_{l,s}$ and $\gamma_{h,s}$ (as well as the probabilities of drawing them, see next bullet point) (a_s, b) . [2 targets]
- The college earnings premium of marginal households (those close to indifferent between attending and not attending under the benchmark policy), as empirically measured by Findeisen and Sachs (2014). In our model this statistic is primarily governed by the parameter ν . [1 target].

6.3 Technology

The parameters to be calibrated are (α, δ, ρ) . We choose the parameter $\rho = 0.285$, corresponding to an elasticity of substitution elasticity between unskilled and skilled labor of $\frac{1}{1-\rho} = 1.4$, as estimated by Katz and Murphy (1992); see also Borjas (2003). We also consider a version of the model in which both types of labor are perfect substitutes, $\rho = 1$, and thus $\frac{1}{1-\rho} = \infty$. In this case a change in the relative supply of college-educated labor, $\frac{L_c}{L_n}$, will have no impact on its relative price, $\frac{w_c}{w_n}$. In appendix C.2 we show that when moving from the perfect to the imperfect substitution case we do not have to recalibrate any of the other parameters (apart from a TFP scaling factor in the production function) for the model to attain the same steady state statistics.

The capital share is set to $\alpha = 1/3$. Furthermore we target an investment to output ratio of 20% and a capital-output ratio of 2.65. Accounting for population growth this implies a yearly depreciation rate of 7.55% and thus a yearly interest rate of about 5.4%. The capital-output ratio (equivalently, the real interest rate) will be attained by appropriate calibration of the preference parameters (especially the time discount factor β), as discussed below.

6.4 Government Policy

In the initial steady state the policy parameters to be chosen are $(\tau_k, \tau_l, \tau_c, \tau_p, d, b, gy)$. We pick $b = 0.6$ and $gy = 0.17$ to match a government debt to GDP ratio of 60% and government consumption (net of tertiary education expenditure) to GDP ratio of 17%. Consumption taxes can be estimated from NIPA data as in Mendoza, Razin and Tesar (1994) who find $\tau_c \approx 0.05$. For the capital income tax rate, we adopt Chari and Kehoe's (2006) estimate of $\tau_k = 28.3\%$ for the early 2000's.

The payroll tax $\tau_{ss} = 12.4\%$ is chosen to match the current social security payroll tax (excluding Medicare). We model social security benefits $p_{t,j}(e, s)$ as concave function of average wages earned during a household's working life, in order to obtain a reasonably accurate approximation to the current progressive US benefit formula, but without the need to add a continuous state variable to the model. The details of the calibration of social security benefits are contained in appendix C.1.

We calibrate the labor income tax deduction to match the sum of standard deductions and exemptions from the US income tax code. Both median income as well as the size of the standard exemption and deduction varies by household size and type, but their ratio is roughly constant at 35%. Thus we calibrate the deduction in the benchmark economy to 35% of the (endogenous) median income in the model. That is, we choose the policy parameter d such that $\frac{d \cdot Y/N}{med(y^{gross})} = 35\%$, where Y/N is output per capita in the model.³⁵ Finally the marginal tax rate on labor income τ_l is chosen to balance the government budget.

6.5 Preferences

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

³⁵ In 2009, according to the U.S. Census Bureau Statistical Abstract of the United States 2012 (table 692), median household money income of a household of 4 members was \$73,071, relative to a sum of standard deduction (\$11,400) and four times the exemption ($4 \cdot 3,650$) of \$26,000. The corresponding numbers for a two person household are \$53,676 and \$18,700 and for a single person of \$26,080 and 10,350. The corresponding ratios are $d = 35.6\%$, $d = 34.8\%$ and $d = 35.9\%$.

We approximate money income in the model as

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

with social security contributions by the employer not part of the measure of income to which we relate the size of the deduction.

period utility function as

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

We 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 3 and households on average work 1/3 of their time.^{36,37}

6.6 Education Costs and Subsidies

We choose the resource cost for college education κ and the share of expenses borne by the government and private sources, θ and θ_{pr} , in the benchmark model to match the total average yearly cost of going to college, as a fraction of GDP per capita, $\frac{\kappa w_c}{y}$, and the cost net of subsidies, $\frac{(1-\theta-\theta_{pr})\kappa w_c}{Y/N}$.

To calculate the corresponding numbers from the data we turn to Ionescu and Simpson (2014) who 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 that 67% of all students that finish college completed a 4 year college and 33% a two year college. Thus the average net cost of tuition and fees for one year of college is

$$0.67 \cdot 58,654/4 + 0.33 \cdot 20,535/2 = \$13,213.$$

Average GDP per capita during this time span was, in constant 2005 dollars, \$42,684. Thus

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

Furthermore education at a glance (OECD 2012, Table B3.2b) reports that the share of tertiary education expenditures borne by public and private subsidies is $\theta = 38.8\%$ and $\theta_{pr} = 16.6\%$, so that

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

Thus the cost parameter κ is calibrated so that the equilibrium of the benchmark model has to be calibrated within the model so that in the model $\frac{\kappa w_c}{Y/N} = 0.694$.

³⁶ These preferences imply a Frisch elasticity of labor supply of $\left(\frac{1-\mu(1-\sigma)}{\sigma} \right) \left(\frac{1-l}{l} \right)$, and with an average labor supply of $l = 1/3$ one could be worried that the Frisch labor supply elasticity, which, given the parameter estimates will be around 1 for most households, is implausibly high. But note that this elasticity of labor supply of entire households, not that of white prime age males on which many lower empirical estimates are based. Also, the average Frisch elasticity for prime age workers in the age bin 24-54 in our model is at 0.6 which we view as a conservative estimate.

³⁷ The coefficient of relative risk aversion with this formulation equals $\sigma\mu + 1 - \mu \approx 2$.

6.7 Ability Transitions and College Time Costs

Newly formed households 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 the latter is in turn determined by parental ability e_p . We interpret $e \in [0, 1]$ as 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 $e_p \in [0, 1]$

$$\pi(e|s_p, \gamma_p) = \frac{\psi\left(\frac{e - \mu(s_p, \gamma_p)}{\sigma_e}\right)}{\Psi\left(\frac{1 - \mu(s_p, \gamma_p)}{\sigma_e}\right) - \Psi\left(\frac{-\mu(s_p, \gamma_p)}{\sigma_e}\right)}, \quad (35)$$

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

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

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

Note that this parametrization of the intergenerational ability transmission gives both a role to parental education and to parental ability (through their draws of γ) for shaping children's ability e . Setting $\chi = 0$ shuts down the effects of parental on offspring innate ability e . In contrast, setting $\nu = 1$ in the draw of parental γ and setting $\zeta = 0$ eliminates the impact of parental education s on the ability e of their offspring.

We then restrict e to take on a discrete set of $n_e = 31$ values that are evenly spaced in the unit interval.³⁸ 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 thus utility) cost of going to college. We calibrate λ to match the overall share of households completing college in the data.

³⁸ We do not directly target intergenerational ability persistence. Our calibration implies model intergenerational ability persistence of 0.22. The ability persistence appears to be somewhat lower than in the data.

To obtain college completion rates of students by parental education we turn to the National Education Longitudinal Study (NELS:88).³⁹ We compute the percent of individuals from this nationally representative sample who were first surveyed as eighth-graders in the spring of 1988, that by 2000 had obtained at least a Bachelors degree, conditional on the highest education level of their parents. We identify $s_p = c$ in our model with the highest education of a parent being at least a Bachelors degree (obtained by 1992). We find that for students with parents in the $s_p = c$ category 63.3% have completed a Bachelors degree. The corresponding number for parents with $s_p = n$ is 28.8%.

6.8 Borrowing Constraints

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

$$\underline{A}_{j_c,t} = (1 + r_t)\underline{A}_{j_a,t-1} + \phi(1 - \theta_t - \theta_{pr})\kappa w_{t,c}.$$

For $j = j_c + 1, \dots, j_r$ we specify

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

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

The parameter ϕ to be calibrated determines how tight the borrowing constraint for college is. Note that in contrast rp is not a calibration parameter but an endogenously determined repayment amount that insures that households don't 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 for independent undergraduate students (the more relevant number given that our students are independent households).⁴⁰ Relative to GDP per capita in 2008 of \$48,000, this given maximum debt constitutes 14% and 23.4% of GDP per capita. Compare that to the 31% of total costs computed above, this indicates that independent undergraduate students can borrow at most approximately 75% of the cost of college, and thus we set $\phi = 0.75$.

Table 5 summarizes the parameters used in our optimal tax computations.

³⁹ <http://nces.ed.gov/surveys/nels88/>

⁴⁰ Note that about 66% of students finishing four year colleges have debt, and *conditional* on having debt the average amount is \$23,186 and the median amount is \$20,000.

Table 5: Calibration

Parameter	Interpretation	Value
Exogenously Calibrated Parameters		
<i>Population</i>		
j_a	Age at HH form. (age 18)	4
j_c	Age, coll. compl. (age 21)	5
j_f	Fertility Age (age 30)	7
j_r	Retirement Age (age 66)	16
J	Max. Lifetime (age bin 98-101)	24
χ	Population Growth Rate	1.008
$\{\varphi_i\}$	Survival Probabilities	Life Tables SSA
<i>Labor Productivity</i>		
$\{\varepsilon_{j,s}\}$	Age Profile	Estimates (PSID)
\mathcal{E}_s and $\pi_s(\eta' \eta)$	Stochastic Part of Wages	Estimates (PSID)
<i>Preferences</i>		
σ	Coef. of Rel. Risk Aversion = 2	4
ζ	Equivalence Scale	0.3
<i>Technology</i>		
α	Capital Share	33.3%
δ	Depreciation	7.55%
ρ	Subst. Elasticity $(1/(1-\rho)) \in \{1.4, \infty\}$	$\{0.285, 1\}$
<i>Ability and Education</i>		
ϕ	Tightness of Borrowing Constraint	75.0%
<i>Government Policy</i>		
θ	Public Education Subsidy	38.8%
θ_p	Private Education Subsidy	16.6%
τ_c	Consumption Tax Rate	5.0%
τ_k	Capital Income Tax Rate	28.3%
b	Debt-GDP Ratio	60.0%
gy	Gov. Cons to GDP Ratio	17.0%
τ^p	Social Security Payroll Tax	12.4%
Parameters Calibrated in Equilibrium (Targets in Brackets)		
<i>Preferences</i>		
β	Time Discount Rate (K/Y)	0.988
v	Altruism Parameter (Avg. Transfers)	0.482
μ	Leisure Share (Fraction of h worked)	0.374
<i>Ability and Education</i>		
$\varepsilon \equiv \frac{\varepsilon_{j,c,c}}{\varepsilon_{j,c,n}}$	Relative Productivity at age j_c (College Wage Premium, 1.8)	1.03
v	Probability to draw γ_h , $\pi(\gamma = \gamma^h e, s = n) = b \cdot e$ (College Wage Premium for Marginal HH)	0.5
λ	Time Costs of College (College Completion)	1.351
σ_e	Std. Dev of e (Distribution Coverage 90%)	0.266
κ	Resource Cost of Coll. (Spend. on Tert. Educ.)	0.203
θ_{1s}	(s -specific variance of fixed effect)	[0.338, 0.243]
χ	$\mu(e)$ -shifter for γ^h (Intergenerational Earnings Persistence of 0.4)	0.134
ζ	$\mu(e)$ -shifter for $s_p = c$ (Intergenerational Education Persistence)	0.057
<i>Government Policy</i>		
τ_l	Labor Income Tax Rate (Budget Bal.)	27.5%
d	Tax Deduction Rate (Median Rate of 35%)	27.1%

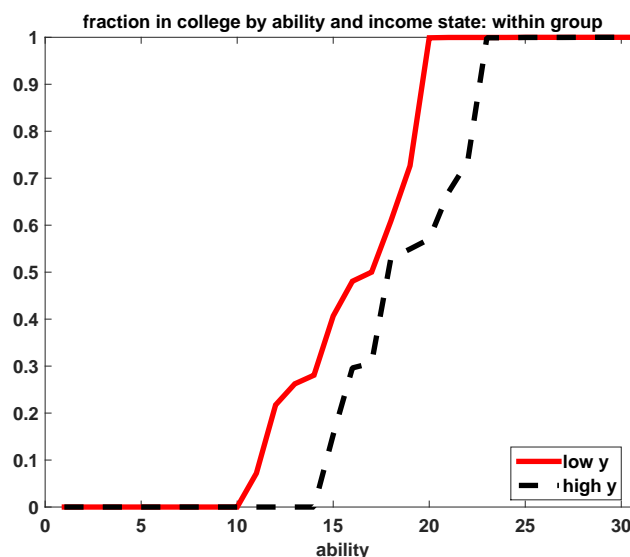
7 Results

In this section we present our results. We first discuss what determines the education decision in the initial steady state; since that steady state is identical across the perfect and imperfect substitutability case no distinction is needed. We then turn to the determination of the optimal tax-subsidy policy showing that considering transitional dynamics and general equilibrium wage effects are crucial for the normative policy analysis. Finally we discuss the sources of the welfare gains from the optimal policy, relative to the initial status quo.

7.1 How the Model Works: The Education Decision

Prior to presenting the optimal tax results it is instructive to discuss how households make their key economic decisions *for a given policy*. Ours is a fairly standard life cycle model with idiosyncratic wage risk, and thus the life cycle profiles of consumption, asset and labor supply are consistent with those reported in the literature (see e.g. Conesa et al. (2009), figure 1). Instead, here we explore how the optimal education decision is made, as a function of the initial characteristics of the household. This focus is further warranted by the observation that the optimal policy will have a strong impact on this decision and will result in a significant change in the share of households obtaining an education in the aggregate, which is in turn important for understanding the optimality of the policy in the first place.

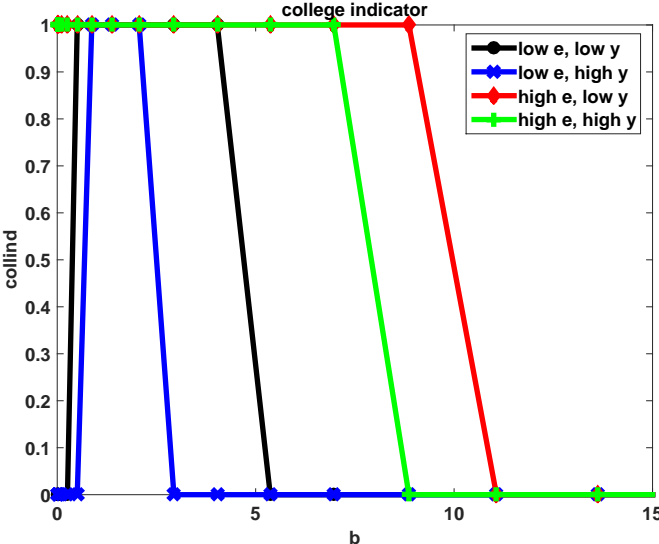
Figure 4: Fraction of Households Deciding to Go to College



Recall that households, at the time of the college decision (that is, at age j_a) differ according to (e, η, b) , that is, their ability to go to college e , their wages outside college (as

determined by the idiosyncratic shock η), and their initial asset levels resulting from parental transfers b . In figure 4 we display the share of households deciding to go to college, under the status quo policy, as a function of e , both for households with low and with high η realizations and the associated low and high incomes y . All households with high abilities ($e \geq e_{23}$) go to college, and non of the households with very low ability ($e \leq e_{10}$) do. For households in the middle of the ability distribution, their decision depends on the attractiveness of the outside option of working in the labor market: a larger share of households with lower opportunity costs (low η and thus y) attends college. Finally, a share strictly between zero and one, conditional on η , indicates that wealth heterogeneity among the youngest cohort (which in turn stems from wealth and thus transfer heterogeneity of their parents) is an important determinant of the college decision for those in the middle of the ability distribution ($e \in [e_{11}, e_{22}]$).

Figure 5: College Decision by Initial Assets



Notes: The low (high) ability group is group 17 (25) out of 31 groups.

This point is further reinforced by figure 5 which displays the college decision indicator function in dependence of initial assets b , and conditional on the non-college wage realization. A value of 0 on the y -axis stands for not attending college, a value of 1 represents the decision to go to college. Assets on the x -axis are normalized such that a value of $b = 1$ stands for assets equal to one time average asset holdings of the parental generation at the age intergenerational transfers are given. We display the policy function for those with relatively low ability ($e = e_{17}$) and those with relatively high ability ($e = e_{25}$).

We make several observations. First, and not surprisingly, lower-ability households go to college for a smaller range of initial assets than do high ability households (blue and black line vs. green and red line). Second, as discussed above, a higher non-college wage (high y) reduces the incidence of attending college. Finally and perhaps most

interestingly, the effects of initial wealth on the college decision are non-monotone. For households at the low end of the wealth distribution (and with sufficiently low e) the borrowing constraint is important. Although the government subsidizes college (in the status quo it covers a 38.8% share of the costs) and although households can borrow 75% of the remaining resource costs, at zero or close to zero wealth a household might still not be able to afford college. That is, either it is impossible for these households to maintain positive consumption even by working full time while attending college, or the resulting low level of consumption and/or leisure make such a choice suboptimal. As parental transfers increase the borrowing constraint is relaxed and even the less able households decide to go to college. Finally, sufficiently wealthy households that expect to derive a significant share of their lifetime income from capital income find it suboptimal to invest in college and bear the time and resource cost in exchange for larger labor earnings after college. Note, however, that although this last result follows from the logic of our model, it is not important quantitatively since the stationary asset transfer distribution puts essentially no mass on initial assets $b \geq 5$.

7.2 Analysis of Optimal Policy Transitions

To summarize our main results right at the beginning: starting from the status quo, the optimal policy transition is obtained by a significantly larger education subsidy and a significantly less progressive tax system. The welfare gains relative to the status quo are substantial, in the order of 1.6% to 3.5% of lifetime consumption. This statement is robust across different degrees of substitutability of labor but it does depend on the explicit consideration of the transition path.

Tables 6 and 7 display the optimal policy combinations, both for $\rho = 1$ and $\rho = 0.285$, and both for maximizing steady state and transitional welfare.⁴¹ In addition, the tables provide summary measures of aggregate economic activity⁴² and inequality statistics for the initial steady state with status quo policy (column 2), and the final state of the optimal policy induced transition (column 3). Column 4 shows the change in these variables between the initial and the final steady state.⁴³ Finally, columns 5 and 6 do the same, but for the *optimal steady state policy* that ignores transitional dynamics and

⁴¹ In order to meaningfully compare the optimal steady state policies and the optimal transition policies we adopt the same type of welfare criterion when maximizing steady state welfare as when maximizing transitional welfare: the integral of lifetime utilities, weighted by the cross-sectional distribution of state variables in the initial steady state. In contrast to maximizing expected utility of a newborn household (as often done in the literature, see e.g. Conesa et al., 2009), this welfare measure places positive weight on older households even in the steady state.

⁴² All variables are denoted in per capita terms.

⁴³ For variables that are already in % units we report the percentage *point* changes. For the Gini coefficients we simply report the point changes.

hence transitional welfare.⁴⁴

Table 6: Long-Run Effects: Perfect Substitutes ($\rho = 1$)

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	-0.24 p
$Gini(a)$	0.5558	0.5203	-3.55 p	0.5336	-2.22 p
CEV			1.6565%		2.6126%

7.2.1 The Optimal Policies

The first five rows of tables 6 and 7 display the fiscal constitution in the economy, both in the initial steady state as well as in the optimum. Recall that τ_l is the marginal labor income tax rate, θ the public subsidy rate, $Z = d \frac{Y}{N}$ is the size of the labor income tax deduction and d measures the size of the deduction relative to income per capita.

Focusing first on the optimal transition policy (rows 3 and 4 of both tables) we observe

⁴⁴ When characterizing optimal steady state policies we hold constant the government debt to GDP ratio at its initial steady state level.

that regardless of whether policies affect the relative price of college labor (table 6 vs. 7) the government finds it optimal to heavily subsidize college education, in the order of 120% to 150% of the college tuition cost. That subsidy rate becomes even higher when transitional dynamics (and thus the transitional cost of building up a more educated workforce) is ignored completely and steady state welfare is maximized (170% in the perfect and 200% in the imperfect substitutability case). At the same time the tax system becomes less progressive, with the optimal tax deduction being cut more than in half.

Table 7: Long-Run Effects: Imperfect Substitutes ($\rho = 0.285$)

Var.	Status Quo	Trans. Dynamics		Steady State	
		Opt. Pol.	Change	Opt. Pol.	Change
τ_l	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%
$\tau_l \cdot Z$	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%

For both specifications of technology the new policy mix increases the share of those going to college, and thus, in the long run, the share of the population with a college degree very substantially. However, the exact size of the college boom depends massively on the extent to which the policy-induced college boom triggers changes in the relative wage $\frac{w_c}{w_n}$ of skilled labor. If this effect is absent (perfect substitutes, $\rho = 1$,

table 6), then the college subsidy of 120% (and the reduction of tax progressivity) implies that in the long run 3/4 of each cohort goes to college and the same share of the working age population is college-educated.

In contrast, in our preferred specification with imperfect substitutes ($\rho = 0.285$, table 7) an even more sizable increase in education subsidies also encourages college attendance, but at 10.5% points the increase is not nearly as large (relative to the 32% point college boom in the perfect substitutes case). The reduction in $\frac{w_c}{w_n}$ in general equilibrium strongly mitigates the increased incentives to go to college stemming from the policy reform.

7.2.2 Long-Run Impact of the Policy Reform on the Macro Economy

Both tables 6 (perfect substitutes) and table 7 (imperfect substitutes) show that the optimal policy reform has a strong positive effect on output and consumption per capita (15% and 18.8% in table 6, still 8% and 10.3% in table 7), despite the fact that hours worked only move moderately in response to the lower marginal labor income taxes. Nevertheless labor efficiency units per person $\frac{L}{N}$ increase drastically in the new, relative to the old steady state, on account of a now more skilled workforce. Quantitatively this effect is much less potent with imperfect substitutability of skilled and unskilled labor ($\frac{L}{N}$ increases only by 7.6%, compared to 16% with $\rho = 1$).

Even though in both economies the capital stock and government debt per capita increase along the policy-induced transition towards the new steady state, the difference in the *magnitude* of the college boom is so great that the capital-labor ratio falls and the real interest rate rises in the economy with perfect substitutes whereas the opposite is true if skilled and unskilled labor are imperfect substitutes.

On the distributional side, 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, if the college wage premium shrinks in response to the policy reform, as is the case in our economy with imperfect substitutes, consumption inequality declines more strongly than in the perfect substitutability case. Overall, in both economies the policy reforms increase educational attainments, raise output and consumption per capita significantly on account of a more productive workforce and at the same time reduce consumption inequality.⁴⁵

Note however, that these effects rely on the long-run expansion of college attainment within the workforce, which takes time to materialize and requires higher investment (in terms of time and resources) into tertiary education. A full evaluation of the costs and benefits therefore requires an explicit characterization of the transitional dynamics

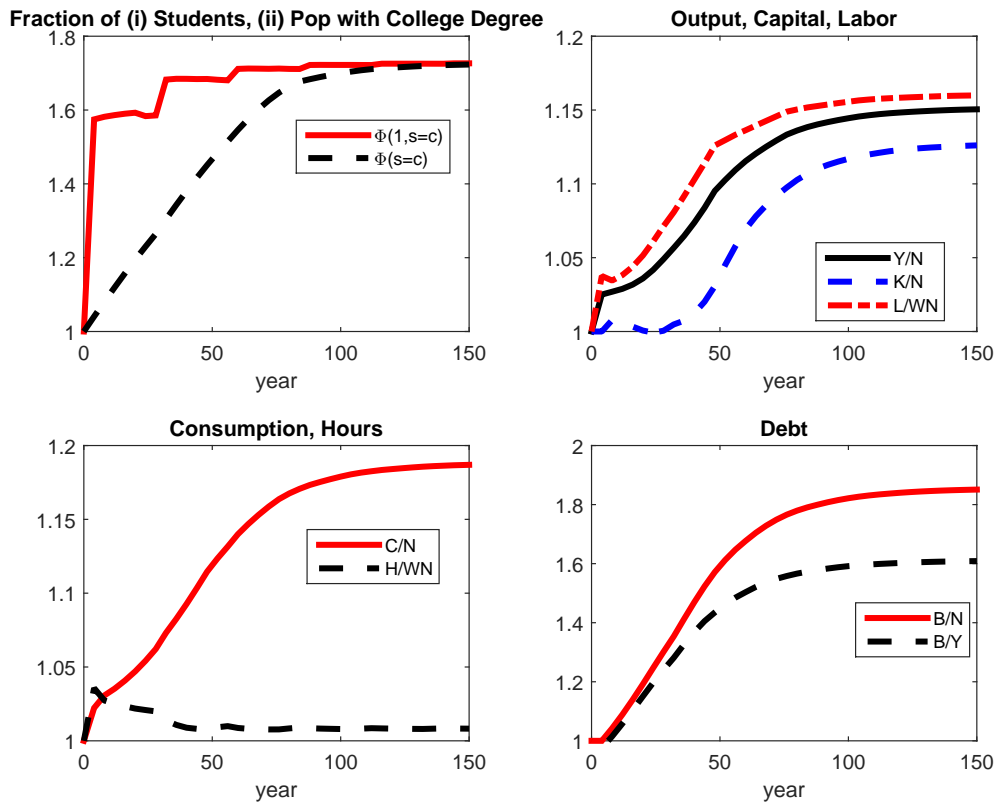
⁴⁵ Hours worked increase slightly and hours (and thus leisure) inequality rises somewhat, but both effects are quantitatively rather modest.

that any (and therefore the optimal) policy reform induces. We turn to this analysis in the next section.

7.2.3 Transitional Dynamics

The discussion of the positive consequences and normative benefits of the policy reforms have so far ignored the fact that it takes time (and resources) to build up a more skilled workforce, suggesting that an explicit consideration of the transition path is important. At any point in time, the youngest cohort constitutes just a small share of the overall workforce, so even if the education decision of this cohort is changed drastically on impact in favor of more college education, it takes years, if not decades, until the skill composition of the entire workforce changes significantly (as older, less skilled cohorts retire and younger, more skilled cohorts take over).

Figure 6: Evolution of Macroeconomic Aggregates: Perfect Substitutes



In figures 6 and 8 we plot the evolution of the key macroeconomic variables along the policy-induced transition path for both economies. In this section we focus on the case with perfect substitutes, whereas in the next section we stress the importance of general equilibrium relative wage effects implied by the imperfect substitutability of skilled and unskilled labor.

The upper left panel of figure 6 displays both the share of the youngest cohort going

to college as well as the overall fraction of the population. Whereas the share of the youngest cohort going to college increases immediately by close to 60% on policy impact, it takes approximately two generations (roughly 60 years) until the overall skill distribution has reached a level close to its new steady state value. It is this sluggish dynamics of the skill and thus labor productivity distribution that a restriction to a long-run steady state policy analysis misses completely.⁴⁶

As explained in the previous section, the long-run expansion of per capita output and consumption is driven by the improved skill distribution in the population. The upper right and the lower left panel of figure 6 show the corollary of this result: 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 capita (upper right panel) increase slowly along the transition as well. The lower right panel documents the same for consumption, and also shows that average hours worked increase on impact by about 3%, mainly 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.⁴⁷

In fact, without the reduction in marginal taxes from 27.55% to 22.9% (and the implied necessary reduction in the tax deduction) the economy would have gone through a fairly severe recession prior to output growth resuming on account of a more productive workforce. Abstracting from the transitional costs, the last column of table 6 shows that the optimal policy when maximizing *steady state* welfare is characterized by *higher* tax progressivity and associated *higher* marginal labor income taxes. Figure 7 shows how the economy goes through a recession if the decrease of taxes and the reduction of the tax deduction would not have taken place when the education subsidy is increased ($d = 0.27, \theta = 1.2$). As households are drawn from the labor market into college, the economy experiences a transitional decline in output and per capita consumption. The figure also shows the outcome along the transition if the government would implement the optimal steady state combination of instruments ($d = 0.31, \theta = 1.7$). In this case the recession is even more severe because the subsidy is substantially higher and so are marginal labor income taxes as well as the tax deduction.⁴⁸

⁴⁶ We also observe a further discrete jump in the share of the newborn cohort attending college 30 years after the initial policy reform. The distribution of ability of children is partially determined by the persistence of innate ability, but also impacted by parental education. The boost of college attendance on impact generates more highly educated parents thirty years later, and thus a more favorable college ability distribution which in turns drives a higher share of children to attend college themselves. Further, but smaller, echo effects are observed for later generations.

⁴⁷ College students can and do work while studying, but the hours they supply are quantitatively minor.

⁴⁸ Interestingly, the long-run increase of per capita consumption is smaller than in the steady state comparison. Taking the transition into account when implementing the optimal steady state policy leads to a long-run increase of per capita consumption by only about 2% whereas the last column of table 6 shows an increase of about 5%.

Figure 7: Evolution of Per-Capita Consumption: Perfect Substitutes

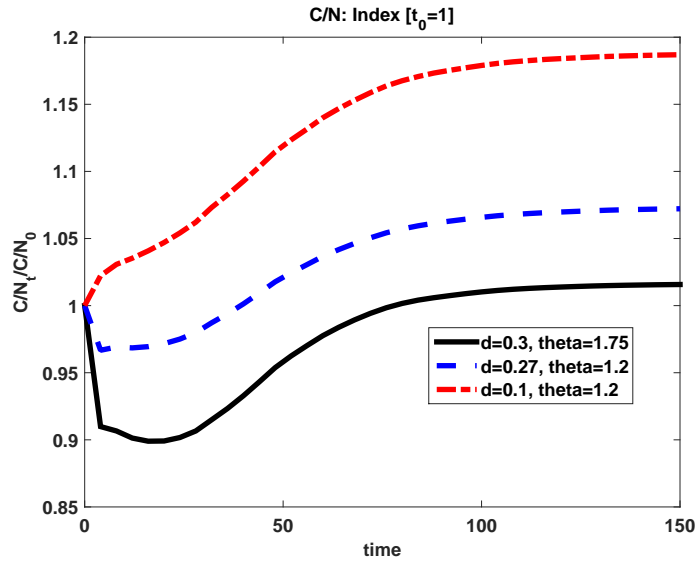


Table 8 summarizes the associated welfare consequences from the policy reform. With the increase of the tax subsidy alone, there would still be welfare gains but those are much smaller than in the optimum. An implementation of the optimal steady state policy would lead to very substantial welfare losses along the transition.

Table 8: Welfare Consequences for Different Policy Combinations: Perfect Substitutes

	$d = 0.1, \theta = 1.2$	$d = 0.27, \theta = 1.2$	$d = 0.31, \theta = 1.7$
CEV	1.6565%	0.0981%	-2.8027%

These findings represent our first main result perhaps most clearly: an explicit consideration of the transition path when conducting a normative welfare analysis of education and tax policies can change the qualitative prescriptions for optimal fiscal policy rather dramatically.

Finally, the lower right panel of figure 6 displays the evolution of government debt per capita and the debt-to-output ratio. Given the tax and education policy and the initial debt level the sequence of government debt is determined from the sequence of 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 revenues due to low—relative to the final steady state—economic activity) by borrowing along the transition. As a consequence government debt per capita increases from 60% to 97% (an increase of about 60%) along the transition. The debt-GDP ratio rises by less than public debt per capita since GDP is also increasing along the transition, although not at the same

speed as debt itself; hence the increase in the debt to GDP ratio.

7.2.4 Importance of General Equilibrium Wage Effects

In figure 8 we collect the dynamics of macroeconomic aggregates for the version of the model with general equilibrium effects in relative wages for skilled versus unskilled labor. Although most of figure 8 looks similar to the plots in figure 6, there are two important distinctions. First, whereas the college boom on impact is comparable across the two economies, now relative wages of the college educated workforce fall as the relative supply of college labor increases.

Figure 8: Evolution of Macroeconomic Aggregates: Imperfect Substitutes

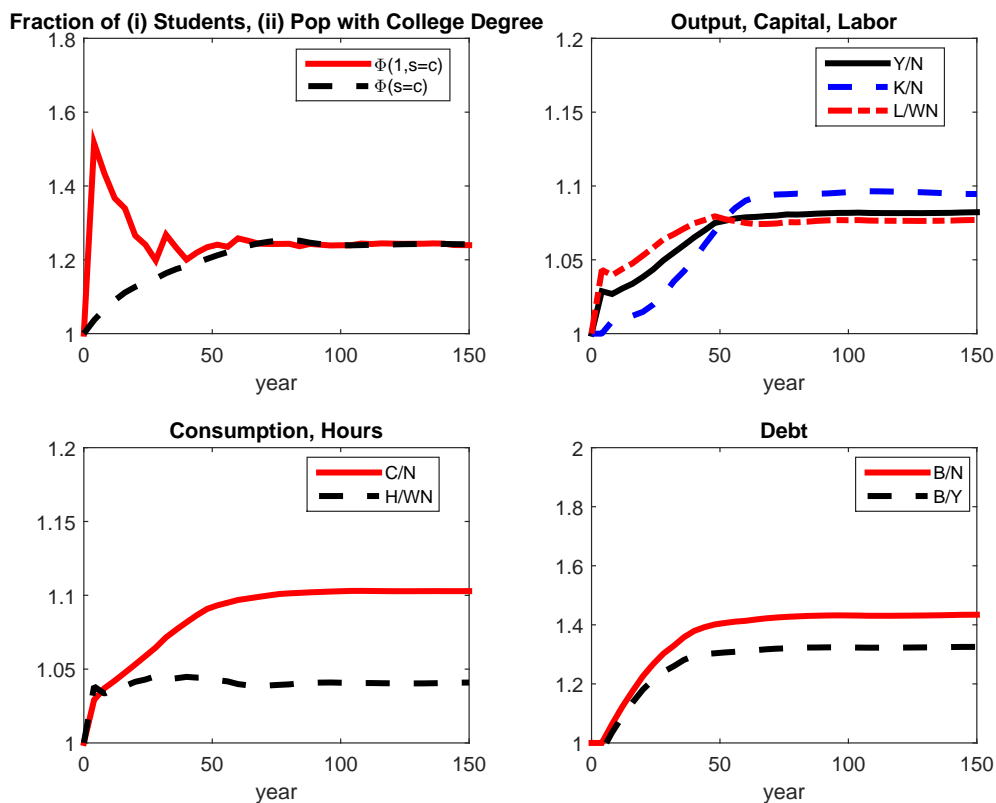
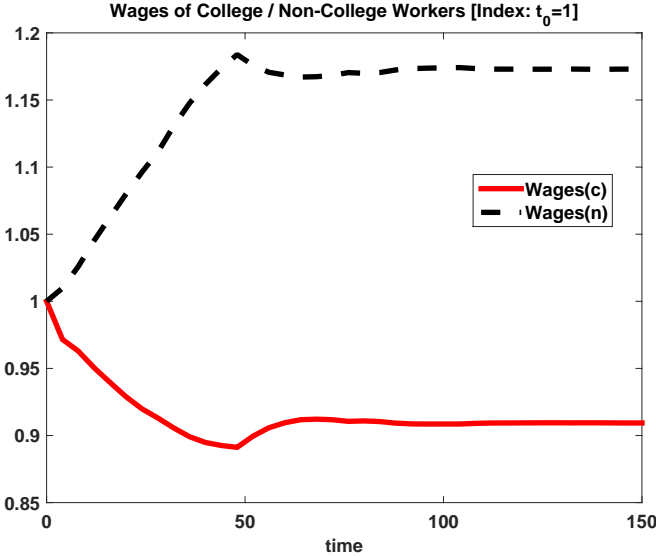


Figure 9 shows these effects by displaying wages of college and non-college households relative to their initial steady state levels. From the figure it is clear that the college wage premium decreases.⁴⁹ Consequently the college boom lessens over time (although the share of young cohorts going to college remains about 23% about its initial steady state level). As a consequence, the skill distribution improves much less

⁴⁹ Also note that, at least in the long-run, the overall wage level per efficiency unit of labor increases, see table 7. In consequence, absolute wages of the unskilled increase and this increase is stronger than the fall of wages of the skilled. Put differently, the fall of wages of the skilled is dampened by the increase of the overall wage level.

pronouncedly in this economy, and the expansion of output, consumption and debt per capita are much smaller, as already discussed for the long run in the context of tables 6 and 7.

Figure 9: Wages of College and Non-College Households: Imperfect Substitutes



To gain further insight into the importance of the general equilibrium response under imperfect substitutes, table 9 compares the welfare gain in the economy with imperfect substitutes (also see last column in row 3 in table 7) to the gain obtained in the economy with perfect substitutes with the optimal $d = 0.06, \theta = 1.5$. The gain is only at 1.61%. This shows that the difference in general equilibrium response, in particular the reduction of the college wage premium, is very important for the welfare assessment of the optimal policy.

Table 9: Welfare Consequences: Importance of GE Response

Substitutability	Imperfect ($\rho = 0.285$)	Perfect ($\rho = 1$)
$CEV(d = 0.06, \theta = 1.5)$	3.4652%	1.6079%

7.2.5 Progressive Income Taxation and Education Subsidies: Complements or Substitutes?

In the simple model we demonstrated how progressive taxes and education subsidies can be complementary second best policies in a world where private insurance is imperfect, thus public insurance is potentially beneficial but distorts human capital accumulation decisions, with education subsidies mitigating this distortions. This

model (and indeed the full quantitative model with perfect substitutability of labor) abstracted from the general equilibrium relative wage effects induced by these policies which curbs the effectiveness of education subsidies for encouraging college attendance. With the two types of labor being imperfect substitutes, the education policy now has indirect beneficial redistributive effects, by reducing the average wage gap between skilled and unskilled workers, and thus might potentially be a *substitute* for redistributive tax policies in general equilibrium.

In this section we explore this theme further, by computing optimal policies along one policy dimension, holding the other dimension constant. We restrict ourselves to steady states to make our points most clearly. Table 10 presents results for a policy decomposition analysis in the steady state for perfect substitutes. The first column restates our findings for the steady state optimum already familiar from table 6. The other two columns document constrained-optimal policies when one instrument (either the education subsidy rate θ or the size of the deduction (as parameterized by d) is held constant.

Comparing our measures for the progressivity of the income tax code, Z as well as $\tau_l \cdot Z$, for the full optimum (column 2) with the constrained optimum (column 4) where we hold θ constant, we observe the complementarity of policy instruments similar to the one we documented in the simple model in table 2. When θ is fixed at its status quo value of $\theta = 38.8\%$, which is lower than the full steady state optimum of $\theta = 170\%$, the constrained optimal effective tax deduction (as measured either by Z or $\tau_l Z$) is lower as well.

Table 10: Final Steady State: Decomposition for Perfect Substitutes ($\rho = 1$)

Var.	Baseline	Constant d	Constant θ
τ_l	36.98%	35.06%	29.34%
d	31%	27.1%	31%
θ	170%	175%	38.8%
Z	0.0866	0.0775	0.08
$\tau_l \cdot Z$	0.032	0.0272	0.0235
CEV	2.6126%	2.5085%	0.1673%

In addition, table 10 also shows that most of the (steady state) welfare gains stem from an adjustment of the education subsidy: the best steady state policy reform brings welfare gains of 2.6% of lifetime consumption, and only adjusting the education subsidy generates 2.5%, whereas adjusting the progressivity of the labor income tax code alone⁵⁰ yields only gains of 0.17%. Of course, absent general equilibrium wage effects

⁵⁰ Also note that the constrained-optimal tax progressivity is rather close to the status quo whereas the optimal education subsidy is substantially higher than the initial policy.

the education policy is very effective in stimulating college attendance and thus labor productivity in the long run; along the transition this policy is not nearly as beneficial, which also explains why the transitional welfare gains are lower than the long-run benefits, compare the last row of columns 4 and 6 of table 6.

Table 11 repeats the same decomposition exercise for imperfect substitutes. Interestingly, we observe that the univariate optima favor a higher degree of tax progressivity and education subsidies, just as for the case with perfect substitutes, cf. table 10. However, once both policy instruments are in operation, the two instruments become substitutes: in the unconstrained optimum, the degree of education subsidies is higher than in the univariate optimum and tax progressivity decreases. Now the additional effect of a decrease of the college wage premium allows the government to achieve a more equal consumption distribution (at least across college- vs. non-college household) *either* through college education subsidies or progressive income taxes. Also, in this case transitional welfare gains are larger than long-run welfare gains since the GE wage reactions only play out in the long-run and reduce college attendance.

Table 11: Final Steady State: Decomposition for Imperfect Substitutes ($\rho = 0.285$)

Var.	Baseline	Constant d	Constant θ
τ_l	26.03%	32.51%	29.34%
d	10%	27.1%	31%
θ	200%	175%	38.8%
Z	0.0281	0.0716	0.0803
$\tau_l \cdot Z$	0.0073	0.0233	0.0236
CEV	2.7569%	2.4135%	0.1309%

We conclude this section by summarizing our second main finding: whether the two policies complement each other or become substitutes in providing redistribution crucially depends on the importance of general equilibrium effects, which in turn are determined by the form of the aggregate production function. Since we find the empirical evidence for imperfect substitutability of skilled and unskilled labor compelling we conclude that the case for education subsidies and progressive labor income taxes being substitutable policies is strong.

8 Conclusions

In this paper we characterized the optimal mix of progressive income taxes and education subsidies and argued that a large education subsidy and a moderately progressive labor income tax and constitute part of the optimal fiscal constitution once household

college attendance decisions are endogenous and transitional dynamics are modeled explicitly. The latter aspect (transitional dynamics) can have a crucial impact on the optimal policy design, as the case of perfect substitutes indicates. Given the important differences between steady state optimal and transition-optimal policy in it conceivable, in fact likely, that policies that are *time-varying* over the transition provide further welfare gains.⁵¹

In our thought experiment we also took the tax on capital income as exogenously given. Future work will need to determine whether our policy conclusions, especially the high subsidy rates for human capital, remain robust once the government chooses not only the progressivity of the labor income tax, but also the optimal mix between capital and labor income taxes.

Finally we determined the optimal tax policy as one which maximizes Utilitarian social welfare among households currently alive.⁵² We also documented that both policy instruments are potentially effective in generating more equally distributed consumption and lifetime welfare. We leave for future work a detailed analysis which elements of our optimal fiscal constitution remains intact if social preferences for insurance and redistribution deviate from the Utilitarian benchmark we have analyzed here.

References

- [1] Abbott, B., G. Gallipoli, C. Meghir and G. Violante (2013), "Education Policy and Intergenerational Transfers in Equilibrium," Working Paper.
- [2] Aiyagari, R. (1994), "Uninsured Idiosyncratic Risk and Aggregate Saving," *Quarterly Journal of Economics*, 109(3): 659-684.
- [3] Alvarez, Y., J. Burbidge, T. Farrell, and L. Palmer (1992), "Optimal Taxation in a Life-Cycle Model," *Canadian Journal of Economics*, 25, 111-122.
- [4] Auerbach, A. and L. Kotlikoff (1987), *Dynamic Fiscal Policy*, Cambridge University Press.
- [5] Bakis, O., B. Kaymak and M. Poschke (2014), "Transitional Dynamics and the Optimal Progressivity of Income Redistribution," *Review of Economic Dynamics*, forthcoming.

⁵¹ Because of the high dimensionality of the policy space these fully time-varying might be hard to characterize numerically for the near future, however.

⁵² Although Utilitarian social welfare is commonly used in the literature, it is of course but one choice for the social welfare function. For alternative criteria and their merits, see e.g. Weinzierl (2012) or Saez and Stantcheva (2013).

- [6] Benabou, R. (2002), "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?" *Econometrica*, 70, 481-517.
- [7] Bewley, T. (1986), "Stationary Monetary Equilibrium with a Continuum of Independently Fluctuating Consumers," in Werner Hildenbrand and Andreu Mas-Colell (eds.) *Contributions to Mathematical Economics in Honor of Gerard Debreu*, North-Holland, 79-102.
- [8] Bohacek, R. and M. Kapicka (2008), "Optimal Human Capital Policies," *Journal of Monetary Economics*, 55, 1-16.
- [9] Bohacek, R. and M. Kapicka (2012), "A Quantitative Analysis of Educational Reforms in a Dynastic Framework," *Manuscript*, University of California Santa Barbara.
- [10] Borjas, G. (2003), "The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market," *Quarterly Journal of Economics*, 118, 1335-1374
- [11] Bovenberg, L., and B. Jacobs (2005), "Redistribution and Education Subsidies are Siamese Twins," *Journal of Public Economics*, 89, 2005–2035.
- [12] Brown, M., M. Mazzocco, J. Scholz and A. Seshadri (2006), "Tied Transfers," Working Paper.
- [13] Castañeda, A., J. Díaz-Giménez and V. Ríos-Rull (1999), "Earnings and Wealth Inequality and Income Taxation: Quantifying the Trade-offs of Switching to a Proportional Income Tax in the U.S." Working Paper.
- [14] Caucutt, E., K. Kumar and S. Imrohorglu (2006), "Does the Progressivity of Income Taxes Matter for Human Capital and Growth?," *Journal of Public Economic Theory*, 8, 95-118.
- [15] Chamley, C. (1986), "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," *Econometrica*, 54, 607-622.
- [16] Chari, V. V., and P. Kehoe (1999), "Optimal Fiscal and Monetary Policy," In *Handbook of Macroeconomics*, ed. John B. Taylor and Michael Woodford, 1671-1745. Amsterdam: North Holland.
- [17] Chari, V. V., and P. Kehoe (2006), "Modern Macroeconomics in Practice: How Theory Is Shaping Policy," *Journal of Economic Perspectives*, 20, 3-28.
- [18] Conesa, J. and D. Krueger (2006), "On the Optimal Progressivity of the Income Tax Code," *Journal of Monetary Economics*, 53, 1425-1450.
- [19] Conesa, J., S. Kitao and D. Krueger (2009), "Taxing Capital: Not a Bad Idea after All!," *American Economic Review*, Vol 99, 25-48.

- [20] Cunha, F., J. Heckman, L. Lochner and D. Masterov (2006), Interpreting the Evidence on Life Cycle Skill Formation, *Handbook of the Economics of Education*, Vol. 1, 697-812.
- [21] Dynarski, S. (2003), "Does Aid Matter? Measuring the Effect of Student Aid on College Attendance and Completion," *American Economic Review*, 93, 279-288.
- [22] Domeij, D. and J. Heathcote (2004), On the Distributional Effects of Reducing Capital Taxes, *International Economic Review*, 45, 523-554.
- [23] Erosa, A. and M. Gervais (2002), "Optimal Taxation in Life Cycle Economies," *Journal of Economic Theory*, 105, 338-369.
- [24] Fehr, H. and F. Kindermann (2015), "Taxing Capital along the Transition - Not a Bad Idea After All?," *Journal of Economic Dynamics and Control*, 51, 64-77.
- [25] S. Findeisen and D. Sachs (2013), "Education and Optimal Dynamic Taxation: The Role of Income-Contingent Student Loans," Working Paper.
- [26] S. Findeisen and D. Sachs (2014), "Designing Efficient College and Tax Policies," Working Paper.
- [27] Fukushima, K. (2010), "Quantifying the Welfare Gains from Flexible Dynamic Income Tax Systems," *Working paper*, University of Wisconsin.
- [28] Garriga, C. (2003), "Optimal Fiscal Policy in Overlapping Generations Models," Working Paper.
- [29] Gale, W. and J. Scholz (1994), "Intergenerational Transfers and the Accumulation of Wealth," *Journal of Economic Perspectives, American Economic Association*, 8, 145-60.
- [30] Gervais, M. (2012), "On the Optimality of Age-Dependent Taxes and the Progressive U.S. Tax System," *Journal of Economic Dynamics and Control*, 36, 682-691.
- [31] Greenwood, J, Z. Hercowitz, and G. Hoffman (1988), "Investment, Capacity Utilization, and the Real Business Cycle," *American Economic Review*, 78, 402-417.
- [32] Guvenen, F., B. Kuruscu and S. Ozkan (2014), "Taxation of Human Capital and Wage Inequality: A Cross-Country Analysis," *Review of Economic Studies*, 81, 818-850.
- [33] Heckman, J., L. Lochner and C. Taber (1998), "Tax Policy and Human-Capital Formation," *American Economic Review*, 88, 293-297.
- [34] Heckman, J., L. Lochner and C. Taber (1999), "Human Capital Formation and General Equilibrium Treatment Effects: A Study of Tax and Tuition Policy," *Fiscal Studies*, 20, 25-40.

- [35] Heckman, J., L. Lochner and P. Todd (2006), "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond," *Handbook of the Economics of Education*, 307-458.
- [36] Heathcote, J., F. Perri and G. Violante (2010), "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States, 1967-2006," *Review of Economic Dynamics*, 13, 15-51
- [37] Heathcote, J., S. Storesletten and G. Violante (2012), "Redistributive Taxation in a Partial-Insurance Economy," mimeo.
- [38] Holmlund, H., M. Lindahl and E. Plug (2011), "The Causal Effect of Parents' Schooling on Children's Schooling: A Comparison of Estimation Methods," *Journal of Economic Literature*, 49, 615-651.
- [39] Holter, H. (2014), "Accounting for Cross-Country Differences in Intergenerational Earnings Persistence: The Impact of Taxation and Public Education Expenditure," Working Paper.
- [40] Hubbard, G. and K. Judd (1986), "Liquidity Constraints, Fiscal Policy, and Consumption," *Brookings Papers on Economic Activity*, 1986, 1-50.
- [41] Huggett, M. (1993), "The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies," *Journal of Economic Dynamics and Control*, 17, 953-969.
- [42] Huggett, M. (1997), "The One-Sector Growth Model with Idiosyncratic Shocks: Steady States and Dynamics," *Journal of Monetary Economics*, 39, 385-403.
- [43] Ionescu, F. and N. Simpson (2014), "Credit Scores and College Investment: Implications for Student Loan Policies," *Working paper*, Colgate University.
- [44] Jacobs, B. and L. Bovenberg (2010), "Human Capital and Optimal Positive Taxation of Capital Income," *International Tax and Public Finance*, 17, 451-478.
- [45] Judd, K. (1985), "Redistributive Taxation in a Simple Perfect Foresight Model," *Journal of Public Economics*, 28, 59-83.
- [46] Judd, K. and C. Su (2006), "Optimal Income Taxation with Multidimensional Taxpayer Types," Working Paper.
- [47] Kapicka, M. (2011), "The Dynamics of Optimal Taxation when Human Capital is Endogenous," Working Paper.
- [48] Karabarbounis, M. (2012), "Heterogeneity in Labor Supply Elasticity and Optimal Taxation," Working Paper.
- [49] Katz, L. and K. Murphy (1992), "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *Quarterly Journal of Economics*, 107, 35-78.

- [50] Keane, M. and K. Wolpin (2001), "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment," *International Economic Review*, 42, 1051-1103.
- [51] Kindermann, F. (2012), "Welfare Effects of Privatizing Public Education When Human Capital Investments Are Risky," *Journal of Human Capital*, 6, 87-123.
- [52] Krueger, D. and A. Ludwig (2013), "Optimal Progressive Labor Income Taxation and Education Subsidies When Education Decisions and Intergenerational Transfers are Endogenous," *American Economic Review*, 103, 496-501.
- [53] Lochner, L. and A. Monge (2011), "Credit Constraints in Education," forthcoming, *Annual Review of Economics*.
- [54] Ludwig, A., T. Schelkle and E. Vogel (2012), "Demographic Change, Human Capital and Welfare," *Review of Economic Dynamics*, 15(1), 94-107.
- [55] Mendoza, E., A. Razin and L. Tesar (1994), "Effective Tax Rates in Macroeconomics: Cross-Country Estimates on Factor Income and Consumption," *Journal of Monetary Economics*, 34, 297-323.
- [56] Mirrless, J. (1971), "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38, 175-208.
- [57] Saez, E. and S. Stantcheva (2013), "Optimal Tax Theory with Endogenous Social Marginal Welfare Weights," Working paper, UC Berkeley.
- [58] Weinzierl, M. (2011) "The Surprising Power of Age-Dependent Taxes," *Review of Economic Studies*, 78, 1-29.
- [59] Weinzierl, M. (2012) "The Promise of Positive Optimal Taxation: A Generalized Theory Calibrated to Survey Evidence on Normative Preferences Explains Puzzling Features of Policy," *NBER Working Paper* 18599.
- [60] Winter, C. (2014) "Accounting for the Changing Role of Family Income in Determining College Entry," *The Scandinavian Journal of Economics*, 116, 909-963

A Theoretical Appendix: Proof of Proposition 9

Proof 17 Recall that $e^{CE}(\tau = \theta = 0) = e^{SP}$ and thus aggregate output net of education costs coincide in the unregulated equilibrium and the solution to the social planner problem. Consequently aggregate consumption is identical as well:

$$\begin{aligned} C^{CE}(\tau = 0, \theta = 0) &= L(\tau = 0, \theta = 0) - \kappa w(1 - e^{CE}(\tau = \theta = 0)) = \\ &L^{SP} - \kappa w(1 - e^{SP}) = C^{SP}. \end{aligned}$$

In the social optimum the utilitarian social planner equalizes lifetime utility across all household types (this follows directly from the first order conditions of the planning problem) and therefore

$$c_c^{SP}(e) = c_n^{SP} + \mu \frac{(I_c^{SP}(e))^{1+\frac{1}{\psi}}}{1 + \frac{1}{\psi}} - \mu \frac{(I_n^{SP})^{1+\frac{1}{\psi}}}{1 + \frac{1}{\psi}} \quad (36)$$

for all e , and thus especially for $e = e^{SP}$. In the competitive equilibrium, at the education threshold (and only there) we have

$$c_c^{CE}(e^{CE}) = c_n^{CE} + \mu \frac{(I_c^{CE}(e^{CE}))^{1+\frac{1}{\psi}}}{1 + \frac{1}{\psi}} - \mu \frac{(I_n^{CE})^{1+\frac{1}{\psi}}}{1 + \frac{1}{\psi}}$$

But for $\tau = d = \theta = 0$ we have $e^{CE} = e^{SP}$ and thus

$$c_c^{CE}(e^{CE}) - c_n^{CE} = c_c^{SP}(e^{SP}) - c_n^{SP} \quad (37)$$

that is, the consumption premium of the marginal type going to college is the same in the unregulated equilibrium and in the social planner problem. Now we show that for all types $e \geq e^{SP}$ we have $\frac{\partial c_c^{CE}(e)}{\partial e} > \frac{\partial c_c^{SP}(e)}{\partial e}$. For this we note that from (36) and (11), evaluated at $\tau = 0$,

$$\begin{aligned} \frac{\partial c_c^{SP}(e)}{\partial e} &= \frac{(1 + \psi) p w^{1+\psi} [(1 + pe)]^\psi}{\left(1 + \frac{1}{\psi}\right) \mu^\psi} \\ \frac{\partial c_c^{CE}(e)}{\partial e} &= \frac{(1 + \psi) p w^{1+\psi} (1 + pe)^\psi}{\mu^\psi} \end{aligned}$$

Thus as long as $\psi < \infty$ we have $\frac{\partial c_c^{CE}(e)}{\partial e} > \frac{\partial c_c^{SP}(e)}{\partial e}$ for all $e \geq e^{CE}(\tau = \theta = 0) = e^{SP}$. But since $C^{CE}(\tau = 0, \theta = 0) = C^{SP}$ it then follows from (37) that

$$\begin{aligned} c_n^{CE} &< c_n^{SP} \\ c_c^{CE}(e^{SP}) &< c_c^{SP}(e^{SP}). \end{aligned}$$

Otherwise we would have $c^{CE}(e) > c^{SP}(e)$ for all e , which violates the resource constraint.

B College Premium for Average and Marginal Attendees

Suppose households with $e \geq e^*$ go to college and households with $e < e^*$ don't. Also suppose the γ -process is

$$\pi(\gamma = \gamma_{h,s} | e, s) = \nu_s e$$

with $\nu_c = 1$ and $\nu_n = \nu < 1$. Then the average college wage premium is proportional to⁵³

$$\overline{wp} = \frac{E(\gamma | s = c)}{E(\gamma | s = n)} = \frac{\gamma_{l,c} + (\gamma_{h,c} - \gamma_{l,c}) \bar{e}_c}{\gamma_{l,n} + (\gamma_{h,n} - \gamma_{l,n}) \nu \bar{e}_n}$$

where \bar{e}_n, \bar{e}_c are the average abilities of non-college and college households, respectively. Note that because of the threshold property we have $\bar{e}_n < e^* < \bar{e}_c$. The premium for the marginal type is given by

$$\overline{wp}(e^*) = \frac{\gamma_{l,c} + (\gamma_{h,c} - \gamma_{l,c}) e^*}{\gamma_{l,n} + (\gamma_{h,n} - \gamma_{l,n}) \nu e^*}$$

Assume that $\gamma_{h,c} > \gamma_{l,c}$ as well as $\gamma_{h,n} > \gamma_{l,n}$ (which will turn out to be the case in our calibration). Then

$$\overline{wp}(e^*) < \overline{wp}$$

since

$$\frac{\gamma_{l,c} + (\gamma_{h,c} - \gamma_{l,c}) e^*}{\gamma_{l,n} + (\gamma_{h,n} - \gamma_{l,n}) \nu e^*} < \frac{\gamma_{l,c} + (\gamma_{h,c} - \gamma_{l,c}) \bar{e}_c}{\gamma_{l,n} + (\gamma_{h,n} - \gamma_{l,n}) \nu e^*} < \frac{\gamma_{l,c} + (\gamma_{h,c} - \gamma_{l,c}) \bar{e}_c}{\gamma_{l,n} + (\gamma_{h,n} - \gamma_{l,n}) \nu \bar{e}_n}.$$

Thus as long as the education decision has the alleged threshold property such that low e households don't do go to college whereas high e households do, the wage premium for the marginal type e^* of going to college is smaller than the average college premium. Ceteris paribus (that is, keeping $\bar{e}_n, \bar{e}_c, e^*$ and all other parameters constant,

$$\frac{\overline{wp}(e^*)}{\overline{wp}} = \frac{\frac{\gamma_{l,c} + (\gamma_{h,c} - \gamma_{l,c}) e^*}{\gamma_{l,n} + (\gamma_{h,n} - \gamma_{l,n}) \nu e^*}}{\frac{\gamma_{l,c} + (\gamma_{h,c} - \gamma_{l,c}) \bar{e}_c}{\gamma_{l,n} + (\gamma_{h,n} - \gamma_{l,n}) \nu \bar{e}_n}} = \frac{\gamma_{l,n} + (\gamma_{h,n} - \gamma_{l,n}) \nu \bar{e}_n}{\gamma_{l,n} + (\gamma_{h,n} - \gamma_{l,n}) \nu e^*} * \frac{\gamma_{l,c} + (\gamma_{h,c} - \gamma_{l,c}) e^*}{\gamma_{l,c} + (\gamma_{h,c} - \gamma_{l,c}) \bar{e}_c}$$

is decreasing in ν , since

$$\begin{aligned} \frac{\partial \overline{wp}(e^*)}{\partial \nu} &\doteq (\gamma_{h,n} - \gamma_{l,n}) \bar{e}_n [\gamma_{l,n} + (\gamma_{h,n} - \gamma_{l,n}) \nu e^*] - \\ &\quad (\gamma_{h,n} - \gamma_{l,n}) e^* [\gamma_{l,n} + (\gamma_{h,n} - \gamma_{l,n}) \nu \bar{e}_n] \\ &= (\gamma_{h,n} - \gamma_{l,n}) \bar{e}_n \gamma_{l,n} + (\gamma_{h,n} - \gamma_{l,n}) \bar{e}_n (\gamma_{h,n} - \gamma_{l,n}) \nu e^* - \\ &\quad (\gamma_{h,n} - \gamma_{l,n}) e^* \gamma_{l,n} - (\gamma_{h,n} - \gamma_{l,n}) e^* (\gamma_{h,n} - \gamma_{l,n}) \nu \bar{e}_n \\ &= (\gamma_{h,n} - \gamma_{l,n}) (\bar{e}_n - e^*) \gamma_{l,n} < 0. \end{aligned}$$

⁵³ The proportionality factor depends on the deterministic age profiles $\{\varepsilon_{j,s}\}$.

Finally note that

$$\begin{aligned}
\frac{\overline{wp}(e^*)}{\partial e^*} &\doteq (\gamma_{h,c} - \gamma_{l,c}) [\gamma_{l,n} + (\gamma_{h,n} - \gamma_{l,n}) \nu e^*] - (\gamma_{h,n} - \gamma_{l,n}) \nu [\gamma_{l,c} + (\gamma_{h,c} - \gamma_{l,c}) e^*] \\
&= (\gamma_{h,c} - \gamma_{l,c}) \gamma_{l,n} + (\gamma_{h,c} - \gamma_{l,c}) (\gamma_{h,n} - \gamma_{l,n}) \nu e^* - \\
&\quad (\gamma_{h,n} - \gamma_{l,n}) \nu \gamma_{l,c} - (\gamma_{h,n} - \gamma_{l,n}) (\gamma_{h,c} - \gamma_{l,c}) \nu e^* \\
&= (\gamma_{h,c} - \gamma_{l,c}) \gamma_{l,n} - (\gamma_{h,n} - \gamma_{l,n}) \gamma_{l,c} \nu
\end{aligned}$$

so this can take either sign, even if $(\gamma_{h,c} - \gamma_{l,c}) > (\gamma_{h,n} - \gamma_{l,n})$, but is positive (again *ceteris paribus*) as long as ν is sufficiently small.

C Calibration Appendix

C.1 Details of the Calibration of Social Security Benefits

The U.S. system is characterized by an indexation to “average indexed monthly earnings” (AIME). This sum the 35 years of working life with the highest individual earnings relative to average earnings. Social security benefits are then calculated as a concave function of AIME.

We approximate this system as follows. First, we define AIME of a type (\hat{e}, \hat{s}) household that retires in year t_r as

$$\bar{y}_{t_r}(\hat{e}, \hat{s}) = \frac{\sum_{j=j_c}^{j_r-1} w_{t-(j_r-1-j), \hat{s}} \epsilon_{j, \hat{s}} \gamma_{\hat{s}}(\hat{e})}{\sum_{e, s} \sum_{j=j_c}^{j_r-1} w_{t-(j_r-1-j), s} \epsilon_{j, s} \gamma_s(e)} \quad (38)$$

as the sum of yearly wages, averaged across all η , for the cohort entering into retirement in year t_r , normalized such that $\sum \bar{y}_{t_r}(e, s, k) = 1$. For simplicity, we start the sum in (38) after college completion and thereby do not account for the lower wages of college attendees while in college.

The primary insurance amount (PIA) of the cohort entering into retirement in period t_r , $pia_{t_r}(e, s)$, is then computed as

$$pia_{t_r}(e, s) = \begin{cases} s_1 \bar{y}_{t_r}(e, s, k) & \text{for } \bar{y}_{t_r}(e, s, k) < b_1 \\ s_1 b_1 + s_2 (\bar{y}_{t_r}(e, s, k) - b_1) & \text{for } b_1 \leq \bar{y}_{t_r}(e, s, k) < b_2 \\ s_1 b_1 + s_2 (b_2 - b_1) + s_3 (\bar{y}_{t_r}(e, s, k) - b_2) & \text{for } b_2 \leq \bar{y}_{t_r}(e, s, k) < b_3 \\ s_1 b_1 + s_2 (b_2 - p b_1) + s_3 (b_3 - b_2) & \text{for } \bar{y}_{t_r}(e, s, k) \geq b_3 \end{cases}$$

for slopes $s_1 = 0.9$, $s_2 = 0.32$, $s_3 = 0.15$ and bend points $b_1 = 0.24$, $b_2 = 1.35$ and $b_3 = 1.99$.

Pensions for all pensioners of age $j \geq j_r$ in period t are then given by

$$p_{t,j}(e, s) = \omega_t w_t (1 - \tau_{ss,t}) \cdot pia_{t_r}(e, s)$$

where ω , the net pension benefit level, governs average pensions.

Budget balance requires that

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

and thus

$$\tau_{ss,t} \sum_s w_{t,s} L_{t,s} = \omega_t w_t (1 - \tau_{ss,t}) \sum_{j=j_r}^J N_{t,j} \int pia_{t_r}(e, s) d\Phi_{t,j}$$

C.2 Calibration of Model with Imperfect Substitutes

We calibrate the economy to the model variant with perfect substitutes in production. For the model variant with imperfect substitutes, we make some adjustments of the model that enable us to avoid recalibration.

First, we adjust equation (19) by a technology scaling parameter Y_0 to the effect that output writes as

$$Y_t = F(K_t, Y_0 L_t) = K_t^\alpha (Y_0 L_t)^{1-\alpha} = K_t^\alpha \left[Y_0 \left(L_{t,n}^\rho + L_{t,c}^\rho \right)^{\frac{1}{\rho}} \right]^{1-\alpha}.$$

This implies that the first-order conditions in (20) rewrite as

$$w_{t,n} = Y_0 \omega_t \left(\frac{L_t}{L_{t,n}} \right)^{1-\rho} \quad (39a)$$

$$w_{t,c} = Y_0 \omega_t \left(\frac{L_t}{L_{t,c}} \right)^{1-\rho} \quad (39b)$$

where $\omega_t = (1 - \alpha)k_t^\alpha$ is the marginal product of the CES aggregate of labor, L_t . Observe that the average wage level in the economy is given by

$$\bar{w}_t = \frac{L_{t,n}}{L_{t,n} + L_{t,c}} \cdot w_{t,n} + \frac{L_{t,c}}{L_{t,n} + L_{t,c}} \cdot w_{t,c} \quad (40)$$

In the economy with perfect substitutes, where $Y_0^{ps} = 1$, we have $w_{t,n}^{ps} = w_{t,c}^{ps} = \bar{w}_t^{ps} = \omega_t^{ps}$. In the economy with imperfect substitutes, we determine Y_0^{ips} such that both economies feature the same average wage level in the calibration period 0. This requirement implies that $\omega_0^{ips} = \bar{w}_0^{ips}$. Setting (40) equal to ω_t and using (39) in the resulting equation gives

$$Y_0^{ips} = \left(\frac{L_{0,n}}{L_{0,n} + L_{0,c}} \cdot \left(\frac{L_0}{L_{0,n}} \right)^{1-\rho} + \frac{L_{0,c}}{L_{0,n} + L_{0,c}} \left(\frac{L_0}{L_{0,c}} \right)^{1-\rho} \right)^{-1}$$

Next, recall from Section 4.3.3 that the deterministic component of a household's life-cycle wage profile is determined by

$$w_{t,s} \epsilon_{j,s} \quad \text{for} \quad s \in \{n, c\}$$

and that the average college wage premium in the perfect substitutes economy is calibrated by appropriate choice of the productivity shifting parameter

$$\epsilon \equiv \frac{\epsilon_{j,c}^{ps}}{\epsilon_{j,c}^{ips}}.$$

In the economy with imperfect substitutes we next normalize the age profiles in the initial steady state such that

$$\epsilon_{j_c,n}^{ips} = \frac{\bar{w}_0^{ips}}{w_{0,n}^{ips}} \quad \text{and} \quad \epsilon_{j_c,c}^{ips} = \frac{\bar{w}_0^{ips}}{w_{0,c}^{ips}} \epsilon$$

to the effect that

$$\frac{\epsilon_{j_c,c}^{ips} w_{0,c}^{ips}}{\epsilon_{j_c,n}^{ips} w_{0,n}^{ips}} = \epsilon.$$

Finally, we define college costs in the economy with imperfect substitutes to be given by

$$\kappa \frac{\epsilon_{j_c,c}^{ips}}{\epsilon} w_{t,c}^{ips}$$

so that, in the initial steady state, college costs are equal to

$$\kappa \bar{w}_0^{ips}$$

as in the model with perfect substitutes.

With these adjustments, households are given the same wages, interest rates, etc. and calibration parameters across the two model variants so that the same fixed point (and the same moments for calibration) will be reached in equilibrium.

D Computational Appendix

D.1 Aggregate Problem

An important element to understand the solution of the aggregate problem is the government budget constraint. Along the transition it is given by

$$\begin{aligned} B_{t+1} &= (1 + r_t)B_t + G_t + E_t - \left(\sum_j N_{t,j} \int T_t(y_t) d\Phi_{t,j} + \tau_{k,t} r_t (K_t + B_t) + \tau_{c,t} C_t \right) \\ &= (1 + r_t)B_t + G_t + E_t - T_t, \quad B_0 \text{ given,} \end{aligned} \quad (41)$$

where T_t denotes total tax revenue. The steady state version is given by

$$b_t = \frac{t_t - (e_t + g_t)}{r_t - n} \text{ for } t \in \{0, T\} \quad (42)$$

where $b_t = \frac{B_t}{Y_t}$, $t_t = \frac{T_t}{Y_t}$, $e_t = \frac{E_t}{Y_t}$ and $g_t = \frac{G_t}{Y_t}$ are the government debt, tax revenue, educational expenditures and government consumption to GDP ratios, respectively.

In the transition these two equations—together with the initial condition that B_0 is determined from the initial steady state of the model—impose restrictions on the system of equations characterizing the transitional dynamics. This implies that the government loses two degrees of freedom in setting its policy along the transition. This insight drives the computational implementation of our model. Specifically, out of the policy parameters, we treat d and θ as the two free policy instruments of the government whereas τ as well as $\{B_t\}_{t=1}^T$ are pinned down by equations (41) and (42).

We then implement separate nested procedures to solve the initial and the final steady states as well as the transitional dynamics of our model. In each case, we take government expenditures to GDP, b_t for $t = 0, \dots, T$ as “given”, either as a calibration target in the initial steady state, i.e., for $t = 0$, or as a separate variable to loop over when we compute the final steady state and the transitional dynamics, i.e., for periods $t = 1, \dots, T$. Observe from (41) and (42) that the final steady state depends on the transition. Accordingly, our following description of the algorithm first treats the solution of the initial steady state followed by the final steady state together with the transition.

D.1.1 Initial Steady State

In the (initial) steady state solutions in period 0, we have three nested fixed point problems. First, as in any aggregative model, we have to solve for market clearing prices. Second, we have to compute a fixed point in measures because initial measures at age j_a , Φ_{0,j_a} , depend on parental measures at age $j_f + j_a$, Φ_{0,j_f+j_a} . Third, we have to compute a fixed point in value functions because the inter-vivo transfer decision of

parents depends on the utility of children which itself has to be consistent with the value function of parents.⁵⁴

We approach this with the following nesting of loops:

1. Outer loop: Loop over initial measures Φ_{0,j_a}
2. Inner loop: Solution for market clearing prices
3. Interior loop: Solution of household problem, including fixed point in value functions.

As to the inner loop of step 2 we iterate on the 6×1 steady state vector of aggregate variables, $\vec{P}_{ss} = [p_1, \dots, p_6]'$. p_1 is the capital intensity, $k_0 = \frac{K_0}{L_0}$, p_2 is the labor share, $l_0 = \frac{L_0}{N_0}$, p_3 is the share of households with a college degree, $\frac{L_0^c}{L_0}$ which is a relevant variable only in the model variant with imperfect substitutes, p_4 are accidental bequests (as a fraction of GDP) $tr_0 = \frac{Tr_0}{Y_0}$ ⁵⁵, p_5 is the debt to GDP ratio, b_0 , and p_6 is the net pension benefit level, ω_0 . Observe that all these elements of \vec{P}_{ss} are defined such that they are constant in the steady state.

The iteration is then as follows:

1. Guess the initial distribution of children, Φ_{0,j_a}^1 , an initial vector of outer loop variables, \vec{P}_{ss}^1 , and an initial value function of children V_{0,j_a}^1 .
2. In measure iteration m , for distribution of children Φ_{0,j_a}^m , loop as follows:
 - (a) In iteration q for guess \vec{P}_{ss}^q , loop as follows:
 - i. Compute all variables that households need to solve their problem. These are the interest rate, r_0 , the relative supply of college and non-college workers which are needed to compute college and non-college wages, $w_{0,s}$, $s \in \{n, c\}$, transfers from accidental bequests, the tax deduction, $\tau_0^l Z_0 = d\tau_0^l \frac{Y_0}{N_0}$ and pension payments.
 - ii. Given these variables solve the household problem. This requires finding the fixed point in value functions. We implement a Howard-type improvement algorithm and accordingly proceed as follows:
 - A. In iteration k for the "outer loop" value function $V_0(j_a, \cdot)^k$ solve for decision functions.

⁵⁴ Observe that the fixed point problems in measures and value functions are specific to the steady state solutions. In the transition, we iterate backward in time to compute value functions and decision rules. When updating measures in each time period t , we iterate forward in time and backward in age. Consequently, the distribution of parents and their inter-vivo transfer decisions in any period t are known when solving the problem of children in t .

⁵⁵ We also update the distribution of bequests in the population but leave out the details here.

- B. Hold decision functions constant and denote by $V_0^{i1}(j_a, \cdot) = V_0^k(j_a, \cdot)$ the “inner loop” value function. In “inner loop” iteration l on the value function, loop backward on the value functions from $j = j_f + j_a$ to j_a to update the inner loop value function $V_0^{i+1}(j_a, \cdot)$. Continue until convergence of “inner loop” value functions with criterion $\|V_0^l(j_a, \cdot) - V_0^{l+1}(j_a, \cdot)\| < \epsilon$.
- C. Denote by $V_0^{k+1}(j_a, 0) = V_0^{i+1}(j_a, \cdot)$ the update of the “outer loop” value function. If $\|V_0^k(j_a, \cdot) - V_0^{k+1}(j_a, \cdot)\| < \epsilon$ STOP, ELSE continue with step 2(a)iiA.
- iii. Aggregate across all households and compute updates of all aggregate variables. Specifically, use (42) to compute an update of the steady state debt to GDP ratio. Collect the updated variables in \vec{P}_{ss} . Notice that $\vec{P}_{ss} = H(\vec{P}_{ss})$, where H is a vector-valued non-linear function.
- (b) Define the root-finding problem $G(\vec{P}_{ss}) = \vec{P}_{ss} - H(\vec{P}_{ss})$, where G is a vector-valued non-linear function. If $\|\vec{P}_{ss}^q - H(\vec{P}_{ss}^q)\| < \epsilon$ STOP, ELSE form an update \vec{P}_{ss}^{q+1} and continue with step 2a. We use Broyden’s method to solve the problem.
3. Update measures $\tilde{\Phi}_{0,j_a}^{m+1}$. If $\|\Phi_{0,j_a}^m - \tilde{\Phi}_{0,j_a}^{m+1}\| < \epsilon$ STOP, ELSE form an update $\Phi_{0,j_a}^{m+1} = \omega \tilde{\Phi}_{0,j_a}^{m+1} + (1 - \omega) \Phi_{0,j_a}^m$ and continue with step 2.⁵⁶

D.1.2 Final Steady State and Transition

For the transition to the finale steady state we start by guessing a sequence of a $7 \cdot (T - 1) \times 1$ vector of equilibrium prices, $\vec{P} = [\vec{p}'_1, \dots, \vec{p}'_7]'$, where $p_i, i = 1, \dots, 7$ are vectors of length $(T - 1) \times 1$. The time series are (i) capital labor ratio, (ii) labor share, (iii) the fraction of workers with a college degree, (iv) accidental bequests, (v) the debt to GDP ratio, (vi) the pension benefit level, and (vii) labor taxes.

Our steps are as follows:

1. Start with an initial guess for the price vector \vec{P}^0 .
2. In iteration q for guess \vec{P}^q solve the final steady state of the model. Relative to the previous description for the initial steady state, do so by artificially holding constant the debt to GDP at the initial guess, i.e., do not include b_T in the set of outer loop variables.

⁵⁶ In our application, we further improve stability of the problem by using an additional outer loop via the price vector, \vec{P}_{ss} , and initial measures. We accordingly apply a nested fixed point iteration over both objects. The reason is the circular relationship in that prices are functions of initial measures and initial measures are functions of prices.

3. Scale all entries in \vec{P}^q (except debt to GDP and taxes) such that the time paths are consistent with the final steady state. Denote the scaled vector by \vec{P}_s^q
4. Solve the household problem. We do so by iterating backwards in time for $t = T - 1, \dots, 1$ to get the decision rules and forward for $t = 1, \dots, T - 1$ (and backward in age in each t) for aggregation.
5. Update variables as in the steady state solutions. Denote by $\vec{\tilde{P}}_s^q = H(\vec{P}_s^q)$ the $7 \cdot (T - 1) \times 1$ vector of updated variables. The non-trivial updating of the debt to GDP ratio, B_t/Y_t for all $t = 1, \dots, T$ is based on an inner loop iteration. We search for labor income taxes $\tau_t^l = \tau_1^l$ for $t = 1, \dots, T$ such that the time path of the debt to GDP ratio is consistent with initial conditions and a final steady state constant b_T . Observe that the initial debt level in period 1 is implied by the initial steady state solution because of the recursive nature of debt, cf. equation (41).
 - (a) Start by defining $\bar{Y}_t = T_t/\tau_t^l$, for $t = 1, \dots, T$ as some measure of average income. We use this in the debt iteration to compute an update of the deficit to GDP ratio. Denote current GDP by Y_t and the current deficit by D_t for all $t = 1, \dots, T$.
 - (b) In each debt iteration k
 - i. Guess τ_T^{l0} .
 - ii. Set $\tau_t^{lk} = \tau_T^{lk}$, for $t = 1, \dots, T$.
 - iii. Compute updated aggregate tax income as $\tau_t^{lk} \cdot \bar{Y}_t$. Compute update of deficit to output ratio, $\frac{D_t + (\tau_t^{l0} - \tau_t^{lk}) \cdot \bar{Y}_t}{Y_t}$ for all $t = 1, \dots, T$.
 - iv. Compute steady state debt to GDP ratio from the steady state debt condition, cf. equation (42).
 - v. Iterate backward (=back-shooting) on the government budget constraint, equation (41). Label new time path of debt to GDP ratio as $\widetilde{B/Y}$.
 - vi. If $\|B_1/Y_1 - \widetilde{B_1/Y_1}\| < \epsilon$ STOP, ELSE continue with step 5(b)ii. We use Brent's method to form updates of τ_T^l .
6. Define the root-finding problem as $G(\vec{P}_s^q) = \vec{P}_s^q - H(\vec{P}_s^q)$. Observe that we take the distance in terms of the scaled vector \vec{P}_s^q . Since T is large, this problem is substantially larger than in the steady state iterations. We use the Gauss-Seidel-Quasi-Newton algorithm suggested in Ludwig (2006) to update \vec{P}^q and proceed with step 2 with each update.

D.2 The Household Problem

D.2.1 Recursive Problems of Households: Reformulation

We apply the endogenous grid method of Carroll (2005) and therefore re-formulate the recursive problem of households by introducing “cash-on-hand” as the relevant state variable. For practical reasons, our definition of cash-on-hand differs slightly from Deaton (1990). As we solve the model backwards, we also start here with the retirement period and continue backwards from there.

We work with an exogenous grid for savings, a' , denoted by \mathcal{G}^a . Precisely, we work with age and time dependent grids but, for sake of simplicity, we leave out a detailed description on how these are constructed.⁵⁷

Problem at $J, J - 1, \dots, j_r$. Denote cash-on-hand by:

$$x = (1 + r_t(1 - \tau_t^k))(a + Tr_{t,j}) + p_{t,j}. \quad (43)$$

Recall that $Tr_{t,j}$ are transfers from accidental bequests and $p_{t,j}$ is pension income.

Given that there is no labor income risk we do not have to take expectations of continuation values and the problem reads as:

$$W_t(j, \gamma, s, x) = \max_{c, x' \geq 0} \{u(c, 1) + \beta \varphi_j W_{t+1}(j + 1, \gamma, s, x')\} \quad (44)$$

subject to the constraints

$$\begin{aligned} x' &= (1 + r_{t+1}(1 - \tau_{t+1}^k))(a' + Tr_{t+1}) + p_{t+1} \\ a' &= x - c(1 + \tau^c) \geq 0. \end{aligned}$$

First-order and envelope conditions:

$$\frac{u_c}{1 + \tau^c} - \beta \varphi_j W_{t+1, x'}(\cdot)(1 + r_{t+1}(1 - \tau_{t+1}^k)) - \mu^a = 0 \quad (45)$$

$$W_{t,x}(\cdot) = \beta \varphi_j W_{t+1, x'}(\cdot)(1 + r_{t+1}(1 - \tau_{t+1}^k)) + \mu^a \quad (46)$$

μ^a is the multiplier associated with the liquidity constraint. Observe that the above equations imply the standard Euler equation.

Solution of this problem gives c . To compute x, a and tax payments we then proceed as follows. Given c, a' , compute x, a and income tax payments as

$$\begin{aligned} x &= a' + c(1 + \tau^c) \\ a &= \frac{x - p_t}{1 + r_t(1 - \tau_t^k)} - Tr_t \\ T_t &= r_t \tau_t^k a \end{aligned}$$

⁵⁷ Our construction of age dependent grids insures that we never end outside the bounds of grids in the computation of policy and value functions so that no extrapolation methods are needed.

Problem at $j_r - 1, \dots, j_f + j_a + 1$. Now let

$$x = (1 + r_t(1 - \tau_t^k))(a + Tr_{t,j}) + w_{t,j,s}(1 - \tau_{ss})$$

be the maximal cash on hand where

$$w_{t,j,s} = \epsilon_{j,s} \gamma_s(e) \eta.$$

This is cash-on-hand before labor income taxes are being paid and with maximum labor income possible.

Notice that the constraint $1 - l \geq 0$ will always be non-binding by the lower Inada condition. Then

$$W_t(j, \gamma, s, \eta, x) = \max_{c, x' \geq 0, l \in [0,1]} \left\{ u(c, l) + \beta \varphi_j \sum_{\eta'} \pi_s(\eta' | \eta) W_{t+1}(j+1, \gamma, s, \eta', x') \right\}$$

subject to the constraints⁵⁸

$$\begin{aligned} x' &= (1 + r_{t+1}(1 - \tau_{t+1}^k))(a' + Tr_{t+1}) + w_{t+1,j+1,s}(1 - \tau_{ss}) \\ \text{where } a' &= x - w_{t,j,s} \tau_t^l (1 - 0.5\tau_{ss}) + \tau_t^l Z_t \\ &\quad - c(1 + \tau^c) - w_{t,j,s}(1 - \tau_{ss} - \tau_t^l(1 - 0.5\tau_{ss}))(1 - l) \end{aligned}$$

Problem at $j_f + j_a$ We now define by \bar{a}' savings including the transfer payments, hence savings net of transfer payments is $a'(e') = \bar{a}' - b(e')f$. As we work with the endogenous grid method, \bar{a}' is exogenous and not contingent on e' . All other variables (decisions) are. We further have, as before, cash on hand as:

$$x(e') = (1 + r_t(1 - \tau_{t+1}^k))(a(e') + Tr_{t,j}) + w_{t,j,s}(1 - \tau_{ss})$$

The problem then is:

$$\begin{aligned} W_t(j, \gamma, s, \eta, x) &= \max_{c(e'), x'(e') \geq 0, l(e') \in [0,1], b(e') \geq 0} \sum_{e'} \pi(e' | s, \gamma) \left\{ u(c(e'), 1 - l(e')) \right. \\ &\quad \left. + \beta \varphi_j \sum_{\eta'} \pi_s(\eta' | \eta) W_{t+1}(j+1, \gamma, s, \eta', x'(e')) \right. \\ &\quad \left. + v \sum_{\eta'} \Pi_n(\eta') \max [W_t(j_a, e', n, \eta', x(b(e'))), W_t(j_a, e', c, \eta', x(b(e')))] \right\}, \quad (47) \end{aligned}$$

where $x(b(e'))$ is cash-on-hand of children, to be defined below.

⁵⁸ Observe that at age $j_r - 1$ there is no need to evaluate the expectation over idiosyncratic income states because η becomes a redundant state variable and the continuation value is $W_{t+1}(j_r, \gamma, s, x')$.

Maximization is subject to

$$\begin{aligned}
x'(e') &= (1 + r_{t+1}(1 - \tau_{t+1}^k)) (\bar{a}' - b(e')f + Tr_{t+1}) + w_{t+1,j+1,s}(1 - \tau_{ss}) \\
a'(e') &= \bar{a}' - b(e')f = x(e') - w_{t,j,s}\tau_t^l(1 - 0.5\tau_{ss}) + \tau_t^l Z_t \\
&\quad - c(e')(1 + \tau^c) - w_{t,j,s}(1 - \tau_{ss} - \tau_t^l(1 - 0.5\tau_{ss})) \cdot \\
&\quad \cdot (1 - l(e')) - b(e')f \geq 0 \\
1 - l(e') &\leq 1
\end{aligned}$$

By the max operator in the above, the problem is non-convex along the b -dimension. Conditional on the choice of $b(e')$, the first-order conditions for $c(e')$, $l(e')$ and the envelope condition for $x'(e')$ are as previously.

We then search over a grid of b to bracket the maximum and subsequently determine the optimal b by a one-dimensional bracketing method (golden search). For each b we compute optimal consumption and labor using standard first-order conditions. Notice that the inequality constraint endogenously fixes a maximum b and that the solution for $\bar{a}' = 0$ is $b = 0$.

Problem at ages $j_f + j_a - 1, \dots, j_f$ Definition of cash-on-hand x is as for the age interval $j_f + j_a + 1, \dots, j_r - 1$.

$$\begin{aligned}
W_t(j, \gamma, s, \eta, x) &= \max_{c, x' \geq 0, l \in [0,1]} \left\{ u \left(\frac{c}{1 + \zeta_f}, 1 - l \right) + \right. \\
&\quad \left. \beta \varphi_j \sum_{\eta'} \pi_s(\eta' | \eta) W_{t+1}(j + 1, \gamma, s, \eta', x') \right\}
\end{aligned}$$

subject to the constraints

$$\begin{aligned}
x' &= (1 + r_{t+1}(1 - \tau_{t+1}^k)) (a' + Tr_{t+1}) + w_{t+1,j+1,s}(1 - \tau_{ss}) \\
a' &= x - w_{t,j,s}\tau_t^l(1 - 0.5\tau_{ss}) + \tau_t^l Z_t \\
&\quad - w_{t,j,s}(1 - \tau_{ss} - \tau_t^l(1 - 0.5\tau_{ss}))(1 - l) - c(1 + \tau^c) \geq 0 \\
1 - l &\leq 1
\end{aligned}$$

Problem at $j = j_a$ Consistent with the notation that cash-on-hand is maximum resources available to the household, we now also deduct college expenditures for those households who have $s = c$. We accordingly have:

$$x = (1 + r_t(1 - \tau_{t+1}^k))(a + Tr_{t,j}) + w_{t,j,s}(1 - \mathbf{1}_s \zeta(e))(1 - \tau_{ss}) - \mathbf{1}_s(1 - \theta_t - \theta_{pr})\kappa w_{t,c}$$

The problem now reads as

$$W_t(j, e, \gamma, s, \eta, x) = \max_{c, a' \geq 0, l \in [0, 1 - \mathbf{1}_s \zeta(e)]} \left\{ u(c, 1 - \mathbf{1}_s \zeta(e) - l) + \beta \varphi_j \sum_{\eta'} \pi_n(\eta' | \eta) W_{t+1}(j + 1, \gamma, s, \eta', x') \right\}$$

subject to the constraints⁵⁹

$$\begin{aligned} x' &= (1 + r_{t+1}(1 - \tau_{t+1}^k)) (a' + Tr_{t+1}) + w_{t+1, j+1, s} (1 - \mathbf{1}_s \zeta(e)) (1 - \tau_{ss}) - \\ &\quad \mathbf{1}_s (1 - \theta_{t+1} - \theta_{pr}) \kappa w_{t+1, c} \\ a' &= x - w_{t, j, s} (1 - \mathbf{1}_s \zeta(e)) \tau_t^l (1 - 0.5 \tau_{ss}) + \tau_t^l Z_t \\ &\quad - c(1 + \tau^c) - w_{t, j, s} (1 - \tau_{ss} - \tau_t^l (1 - 0.5 \tau_{ss})) (1 - \mathbf{1}_s \zeta(e) - l) \geq 0 \\ 1 - l &\leq 1 \end{aligned}$$

Education decision at $j = j_a$ We define the set

$$\begin{aligned} \mathcal{B}(e, \eta) &= \{b \mid W_t(j = j_a, e, s = c, \eta, x) > W_t(j = j_a, e, s = n, \eta, x)\} \\ &= [\underline{b}(e, \eta), \bar{b}(e, \eta)] \end{aligned} \quad (48)$$

where

$$x = (1 + r_t(1 - \tau_t^k)) \left(\frac{b}{1 + r_t(1 - \tau_t^k)} + Tr_{t, j} \right) + w_{t, j, s} (1 - \mathbf{1}_s \zeta(e)) (1 - \tau_{ss})$$

We then have

$$\mathbf{1}_{\mathcal{B}}(e, \eta, b) = \begin{cases} 1 & \text{if } b \in \mathcal{B} \\ 0 & \text{else} \end{cases} \quad (49)$$

as the indicator function.

D.2.2 Computational Implementation

We now provide a more detailed description of the solution of the household problem.

Solving for policy functions. We loop backwards in age using standard methods for finite horizon models. To describe solution at one particular age j , we focus on the case with endogenous labor supply, hence $j < j_r$. We look at an arbitrary age and use indicator variables to denote whether children are living in the household (during age bin $j = j_f + j_a - 1, \dots, j_f$) or whether households attend college (age $j_a, s = c$).

⁵⁹ As in the main text, we slightly abuse notation in that for college-bound households η' at age j_a is drawn from $\Pi_c(\eta')$ rather than $\pi_c(\eta' | \eta)$.

We have two regions, with two subregions each, depending on constraints. The first region is for sufficiently asset rich households who work little and hence do not pay labor taxes. Here we have the (weak inequality) constraint $l \geq 0$.

The second region is less asset rich households who work more and hence pay labor taxes. Here we have the (weak inequality) constraint $l \geq \frac{Z_t}{w_{t,j,s}(1-0.5\tau_{ss})}$ (because otherwise total labor income taxes, taxes net of deduction, would be negative). There is some threshold cash-on-hand level \bar{x} , separating the two regions.

To simplify notation, define

$$\begin{aligned} cex &\equiv (1 - \theta - \theta_{pr})\kappa w_{t,c} \\ w_{t,j,s} &\equiv w_{t,s}\epsilon_{j,s}\gamma\eta \\ w_{t,j,s}^n &\equiv \begin{cases} w_{t,j,s}(1 - \tau_{ss} - \tau_t^l(1 - 0.5\tau_{ss})) & \text{for } x \leq \bar{x} \\ w_{t,j,s}(1 - \tau_{ss}) & \text{otherwise} \end{cases} \end{aligned}$$

and

$$\bar{l} \equiv \begin{cases} \frac{Z_t}{w_{t,j,s}(1-0.5\tau_{ss})} & \text{for } x < \bar{x} \\ 0 & \text{otherwise.} \end{cases}$$

Using this notation, the first-order conditions are:

$$\frac{u_c}{1 + \tau^c} - \beta\varphi_j \sum_{\eta'} \pi_s(\eta'|\eta) W_{t+1x'}(\cdot)(1 + r_{t+1}(1 - \tau_{t+1}^k)) - \mu^a = 0 \quad (50a)$$

$$u_{1-\mathbf{1}_s\tilde{\zeta}(e)-l} - w_{t,j,s}^n \left(\beta\varphi_j \sum_{\eta'} \pi_s(\eta'|\eta) W_{t+1x'}(\cdot)(1 + r_{t+1}(1 - \tau_{t+1}^k)) + \mu^a \right) - \mu^l = 0 \quad (50b)$$

$$W_{tx}(\cdot) = \beta\varphi_j \sum_{\eta'} \pi_s(\eta'|\eta) W_{t+1x'}(\cdot)(1 + r_{t+1}(1 - \tau_{t+1}^k)) + \mu^a, \quad (50c)$$

where μ^l denotes the multiplier on the constraint $1 - \mathbf{1}_s\tilde{\zeta}(e) - \bar{l} - (1 - \mathbf{1}_s\tilde{\zeta}(e) - l) \geq 0$. μ^a denotes the multiplier on the borrowing constraint.

We first look at the case where $\mu^a = 0$.⁶⁰ We solve this problem in four steps:

1. Assume $l = 0$, $\tau_t^l = 0$, hence $a' = x - w_{t,j,s}(1 - \tau_{ss})(1 - \mathbf{1}_s\tilde{\zeta}(e)) - c(1 + \tau^c)$. Solve for c using (50a). Compute μ_l from (50b). If $\mu_l > 0$ then

$$\begin{aligned} x &= a' + w_{t,j,s}(1 - \tau_{ss})(1 - \mathbf{1}_s\tilde{\zeta}(e)) + c(1 + \tau^c) \\ a &= \frac{x - w_{t,s}(1 - \tau_{ss})(1 - \mathbf{1}_s\tilde{\zeta}(e)) + \mathbf{1}_s cex}{1 + r_t(1 - \tau_t^k)} - Tr_t \end{aligned}$$

⁶⁰ Recall from Carroll's description of the endogenous grid method that borrowing constraints are easy to deal with, see below.

$$T_t = r_t \tau_t^k a$$

else proceed to next step.

2. Assume $l \in \left(0, \frac{Z_t}{w_{t,j,s}(1-0.5\tau_{ss})}\right]$, hence $\tau_t^l = 0$ and $a' = x - c(1 + \tau^c) - w_{t,j,s}(1 - \tau_{ss})(1 - \mathbf{1}_s \tilde{\zeta}(e) - l)$. Solution of first-order conditions gives $c, 1 - l$. If $l < \frac{Z_t}{w_{t,j,s}(1-0.5\tau_{ss})}$ then

$$x = a' + w_{t,j,s}(1 - \tau_{ss})(1 - \mathbf{1}_s \tilde{\zeta}(e)) - l + c(1 + \tau^c)$$

$$a = \frac{x - w_{t,s}(1 - \tau_{ss})(1 - \mathbf{1}_s \tilde{\zeta}(e)) + \mathbf{1}_s c e x}{1 + r_t(1 - \tau_t^k)} - Tr_t$$

$$T_t = r_t \tau_t^k a$$

else proceed to next step.

3. Assume $l = \frac{Z_t}{w_{t,j,s}(1-0.5\tau_{ss})}$, $\tau_t^l = 0$. Recall that $a' = x - c(1 + \tau^c) - w_{t,j,s}(1 - \tau_{ss})(1 - \mathbf{1}_s \tilde{\zeta}(e) - l)$. Compute c from (50a). Compute μ_l from (50b). In so doing set $w_{t,j,s}^n = w_{t,j,s}(1 - \tau_{ss} - \tau_t^l(1 - 0.5\tau_{ss}))$ because the net tax is what is relevant at the margin. If $\mu_l > 0$ then

$$x = a' + w_{t,j,s}(1 - \tau_{ss})(1 - \mathbf{1}_s \tilde{\zeta}(e) - l) + c(1 + \tau^c)$$

$$a = \frac{x - w_{t,s}(1 - \tau_{ss})(1 - \mathbf{1}_s \tilde{\zeta}(e)) + \mathbf{1}_s c e x}{1 + r_t(1 - \tau_t^k)} - Tr_t$$

$$T_t = r_t \tau_t^k a$$

else proceed to next step.

4. We have $l > \frac{Z_t}{w_{t,j,s}(1-0.5\tau_{ss})}$, hence $\tau_t^l > 0$. Recall that

$$\begin{aligned} a' &= x - w_{t,j,s}(1 - \mathbf{1}_s \tilde{\zeta}(e))\tau_t^l(1 - 0.5\tau_{ss}) + \tau_t^l Z_t \\ &\quad - c(1 + \tau^c) - w_{t,j,s}(1 - \tau_{ss} - \tau_t^l(1 - 0.5\tau_{ss}))(1 - \mathbf{1}_s \tilde{\zeta}(e) - l) \end{aligned}$$

Solution gives $c, 1 - l$ and

$$\begin{aligned} x &= a' + w_{t,j,s}(1 - \mathbf{1}_s \tilde{\zeta}(e))\tau_t^l(1 - 0.5\tau_{ss}) - \tau_t^l Z_t + \\ &\quad + w_{t,j,s}\left(1 - \tau_{ss} - \tau_t^l(1 - 0.5\tau_{ss})\right)(1 - \mathbf{1}_s \tilde{\zeta}(e) - l) + c(1 + \tau^c) \end{aligned}$$

$$a = \frac{x - w_{t,s}(1 - \tau_{ss})(1 - \mathbf{1}_s \tilde{\zeta}(e)) + \mathbf{1}_s c e x}{1 + r_t(1 - \tau_t^k)} - Tr_t$$

$$T_t = r_t \tau_t^k a + \tau_t^l(1 - 0.5\tau_{ss})w_{t,j,s}l - \tau_t^l Z_t.$$

This description makes clear that, in order to solve the first-order conditions, we basically have to deal with two cases. One is with interior solution for l , either with

positive or zero tax payments. The other one is where l is restricted, either to $\bar{l} = 0$ or to $\bar{l} = \frac{Z_t}{w_{t,j,s}(1-0.5\tau_{ss})}$.

We next show the closed form solutions by the application of the endogenous grid method:

1. Interior solution for l . The intra-temporal Euler equation writes as

$$\begin{aligned} u_{1-\mathbf{1}_s\bar{\zeta}(e)-l} &= \frac{w_{t,j,s}^n}{1+\tau^c} u_c \\ \Leftrightarrow (1-\phi)(1-\mathbf{1}_s\bar{\zeta}(e)-l)^{-1} &= \frac{w_{t,j,s}^n}{1+\tau^c} \phi c^{-1} \end{aligned}$$

Therefore

$$1-\mathbf{1}_s\bar{\zeta}(e)-l = \frac{1-\phi}{\phi} \frac{(1+\tau^c)}{w_{t,j,s}^n} c. \quad (51)$$

Using the above in (50a), denoting by $EDMV \equiv \beta\phi_j \sum_{\eta'} \pi_s(\eta'|\eta) W_{t+1,x'}(\cdot)(1+r_{t+1}^n)$ gives

$$\frac{1}{1+\tau^c} u_c \left(\frac{c}{1+\mathbf{1}_j\bar{\zeta}f'} \cdot \right) = EDMV$$

Therefore:

$$c = \left(\frac{1+\tau^c}{\phi} (1+\mathbf{1}_j\bar{\zeta}f)^{\phi(1-\theta)} \left(\frac{\phi}{1-\phi} \frac{w_{t,j,s}^n}{1+\tau^c} \right)^{(1-\phi)(1-\theta)} EDMV \right)^{-\frac{1}{\theta}}$$

and $1-\mathbf{1}_s\bar{\zeta}(e)-l$ follows from (51).

2. Corner solution, $l = \bar{l}$. Using $l = \bar{l}$ in (50a), again denoting by $EDMV \equiv \beta\phi_j \sum_{\eta'} \pi_s(\eta'|\eta) W_{t+1,x'}(\cdot)(1+r_{t+1}^n)$ gives

$$\frac{1}{1+\tau^c} u_c \left(\frac{c}{1+\mathbf{1}_j\bar{\zeta}f'} \cdot \right) = EDMV$$

Therefore:

$$c = \left((1+\mathbf{1}_j\bar{\zeta}f)^{\phi(1-\theta)} \frac{1+\tau^c}{\phi} \left(\frac{1}{1-\mathbf{1}_s\bar{\zeta}(e)-\bar{l}} \right)^{(1-\phi)(1-\theta)} EDMV \right)^{\frac{1}{\phi(1-\theta)-1}}$$

Next, look at the case with binding borrowing constraints (hence $\mu^a > 0$). In this case $a' = \bar{a}' \leq 0$ is given. Key object is the intra-temporal Euler equation: First-order

conditions (for interior solution with weakly positive labor supply and positive tax payments):

$$u_{1-\mathbf{1}_s\tilde{\zeta}(e)-l} - w_{t,j,s}^n \frac{u_c}{1+\tau^c} - \mu^l = 0$$

which we combine with the resource constraint to characterize the solution, adopting the same steps described above. Solutions again come as closed form expressions for our nested CD-CRRA preferences.

1. Assume $l = 0, \tau_t^l = 0$, hence total resources used for expenditures on consumption and leisure are $ress = x - a'$. Recall that $a' = x - w_{t,j,s}(1 - \tau_{ss})(1 - \mathbf{1}_s\tilde{\zeta}(e)) - c(1 + \tau^c)$. Accordingly, we get $c = \frac{ress - w_{t,j,s}(1 - \tau_{ss})(1 - \mathbf{1}_s\tilde{\zeta}(e))}{1 + \tau^c}$. Compute μ_l from (50b). If μ_l positive, stop, else proceed to next case.
2. Assume $l \in \left(0, \frac{Z_t}{w_{t,j,s}(1 - 0.5\tau_{ss})}\right]$, hence $\tau_t^l = 0$ and $a' = x - c(1 + \tau^c) - w_{t,j,s}(1 - \tau_{ss})(1 - \mathbf{1}_s\tilde{\zeta}(e) - l)$. Resources to be used for consumption and leisure are again $ress = x - a'$. From the intra-temporal Euler equation (51) we get

$$lcr = \frac{1 - \mathbf{1}_s\tilde{\zeta}(e) - l}{c} = \frac{1 - \phi}{\phi} \frac{(1 + \tau^c)}{w_{t,j,s}^n}.$$

From the budget constraint it then follows that

$$c = \frac{ress}{1 + w_{t,j,s}^n + \tau^c}.$$

and $1 - \mathbf{1}_s\tilde{\zeta}(e) - l = lcr \cdot c$. Compute labor tax payments. If labor income taxes are positive, proceed to next case.

3. Assume $l = \frac{Z_t}{w_{t,j,s}(1 - 0.5\tau_{ss})}, \tau_t^l = 0$. Then $ress = x - a'$. If $ress > w_{t,j,s}(1 - \tau_{ss})$ then $c = \frac{ress - w_{t,j,s}(1 - \tau_{ss})(1 - \mathbf{1}_s\tilde{\zeta}(e) - l)}{1 + \tau^c}$, else go to next case. Compute μ_l from (50b) using $w_{t,j,s}^n = w_{t,j,s}(1 - \tau_{ss} - \tau_t^l(1 - 0.5\tau_{ss}))$. If μ_l positive, stop, else proceed with next case.
4. We have $l > \frac{Z_t}{w_{t,j,s}(1 - 0.5\tau_{ss})}$, hence $\tau_t^l > 0$. Recall that

$$\begin{aligned} a' &= x - w_{t,j,s}(1 - \mathbf{1}_s\tilde{\zeta}(e))\tau_t^l(1 - 0.5\tau_{ss}) + \tau_t^l Z_t \\ &\quad - c(1 + \tau^c) - w_{t,j,s}(1 - \tau_{ss} - \tau_t^l(1 - 0.5\tau_{ss}))(1 - \mathbf{1}_s\tilde{\zeta}(e) - l) \end{aligned}$$

Resources to be used for consumption and leisure are therefore $ress = x - a' - w_{t,j,s}\tau_t^l(1 - 0.5\tau_{ss})(1 - \mathbf{1}_s\tilde{\zeta}(e)) + \tau_t^l Z_t$.

From the intra-temporal Euler equation (51) we get

$$lcr = \frac{1 - \mathbf{1}_s\tilde{\zeta}(e) - l}{c} = \frac{1 - \phi}{\phi} \frac{(1 + \tau^c)}{w_{t,j,s}^n}.$$

From the budget constraint it then follows that

$$c = \frac{ress}{1 + w_{t,j,s}^n + \tau^c}$$

and $1 - \mathbf{1}_s \zeta(e) - l = lcr \cdot c$.