Discussion Paper 138

Institute for Empirical Macroeconomics
Federal Reserve Bank of Minneapolis
90 Hennepin Avenue
Minneapolis, Minnesota 55480-0291

December 2000

Does the Progressivity of Taxes Matter for Economic Growth?

Elizabeth M. Caucutt*

University of Rochester

Selahattin Imrohoroglu*

University of Southern California

Krishna B. Kumar*

University of Southern California

*Caucutt: ecau@troi.cc.rochester.edu; Imrohoroglu: selo@marshall.usc.edu; Kumar: kbkumar@usc.edu. We are grateful to Jeremy Greenwood, Robert Lucas, and Harald Uhlig, and seminar participants at the 2000 SED meeting in San Jose, Costa Rica, and USC for helpful comments. Kumar acknowledges financial support from the USC Marshall School's research fund. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

ABSTRACT

A sizeable literature has argued that the growth effects of changes in flat rate taxes are small. In this paper, we investigate the relatively unexplored area of the growth effect of changes in the tax *structure,* in particular, in the progressivity of taxes. Considering such a tax reform seems empirically more relevant than considering changes in flat tax rates. We construct a general equilibrium model of endogenous growth in which there is heterogeneity in income and in the tax rates. We limit heterogeneity to two types, skilled and unskilled, and posit that the probability of staying or becoming skilled in the subsequent period depends positively on expenses on "teacher" time. In the production sector, we consider two sources of growth. In the first, growth arises as a purely external effect on account of production activities of skilled workers. In the second, a portion of the skilled workforce is used to work in research and other productivity enhancing activities and is compensated for it. Our analysis shows that changes in the progressivity of tax rates can have positive growth effects even in situations where changes in flat rate taxes have no effect. Experiments on a calibrated model indicate that the quantitative effects of moving to a flat rate system are economically significant. The assumption made about the engine of growth has an important effect on the impact of a change in progressivity. Quantitatively, welfare is unambiguously higher in a flat rate system when comparisons are made across balanced growth equlibria; however, when the costs of transition to the higher growth equilibrium is taken into account only the currently rich slightly prefer the flat rate system.

1 Introduction

A justification often given by politicians and policy makers for lowering tax rates is that it would boost economic growth. One of the early formal attempts at studying the growth effects of taxes was made by Lucas (1990), who used an endogenous growth model in which human capital was the engine of growth. His conclusion was that tax changes do alter long-run growth rates, but the effect was "quantitatively trivial." In an effort to generalize Lucas' study and to scrutinize the burgeoning tax and growth literature through the lens of a common framework, Stokey and Rebelo (1995) used a general model of endogenous growth to identify model features and parameters that affect the growth rate in a quantitatively significant way when tax rates are changed. They isolate these parameters (factor shares, depreciation rates, elasticity of intertemporal substitution, and elasticity of labor supply), but ultimately conclude that for empirically relevant values of these parameters, Lucas' conclusion of little or no tax effect on the U.S. growth rate is robust.[1]

These and almost all related studies are done in the context of a representative agent framework, where there is no heterogeneity of income or the rate at which this income is taxed. In this paper, we investigate the relatively unexplored area of the growth effect of changes in the tax *structure*, in particular changes in the progressivity of taxes. Given the prevalence of progressive taxes, considering tax reform in a heterogeneous agent context seems empirically more relevant.[2] Another motivation for invoking heterogeneity and different tax rates can be found in Lucas' conclusion itself. He states that the quantitatively trivial effect of taxes on growth in his human capital accumulation model is "because changes in labor taxation affect equally both the cost and the benefit side of the marginal condition governing the learning decision."

With progressive taxes this will not be true; if one is taxed at a lower rate while one is unskilled and at a higher rate when one becomes skilled, the return to becoming skilled, skill accumulation, and growth will be negatively affected. On the other hand, if the human capital accumulation is done in the presence of liquidity constraints, as some economists believe it is, more progressive taxes will decrease taxes for the poor at the accumulation stage and increase their investment; the effect will be opposite for the rich. The overall effect of progressivity on skill accumulation and growth is therefore

[1] The studies they consider in detail are Lucas (1990), King and Rebelo (1990), Kim (1992), and Jones, Manuelli, and Rossi (1993). Mendoza, Milesi-Ferretti, and Asea (1997) use cross-country panel regressions and numerical simulations to come to similar conclusions on negligible growth effects of taxes.

[2] There has been substantial movement in the progressivity of taxation in the US, at least across decades. For instance, the Statistical Abstracts of the United States report that the lowest and highest marginal tax rates were 0% and 43% in 1960, -10% and 50% in 1975, and 12.5% and 42% in 1984. (These rates are for married couples with two dependents when tax brackets are expressed in 1980 dollars.) Likewise decade-long per capita economic growth has also varied widely, between 1.68% and 2.7%, during this period.

an open theoretical and quantitative issue. The main objective of this paper is to shed analytical and quantitative light on the effects of these forces of decreased progressivity on investments by the rich and the poor, and their impact on aggregate growth when different assumptions are made about its source.

We construct a general equilibrium model of endogenous growth in which there is heterogeneity in income and in the tax rates to study this issue. In the two-period overlapping generation model we use, there are two types of adult workers, skilled and unskilled. They spend resources on educating their children in terms of expenses on "teacher" time. We posit that the probability of becoming skilled in the subsequent period depends positively on these expenses. The fraction of skilled workers in the economy, a key endogenous state variable, and the value to being skilled evolve endogenously based on these human capital accumulation decisions. In the production sector, we consider two sources of growth. In the first, growth arises as a purely external effect on account of production activities of skilled workers. In the second, a portion of the skilled workforce is used to work in research and other productivity enhancing activities and is compensated for it.[3] The model is analytically and computationally tractable, while still being rich enough to facilitate calibration and shed interesting quantitative insights on aggregate growth effects of tax progressivity.

Our analysis indicates that changes in the progressivity of tax rates can have important growth effects even in situations where changes in flat rate taxes do not have any effect. We calibrate the model to the U.S. economy and study the growth and welfare effects of eliminating progressivity. These experiments suggest that the quantitative effects are economically significant, in some experiments as high as 0.52 percentage points. Welfare of either type of agent is unambiguously higher in a balanced growth path with flat rate taxes; however, when the transition to the higher growth equilibrium is taken into account only the currently rich prefer a flat rate tax, and that too only slightly. We also find that the assumption made about the engine of growth has an important effect on the impact of a change in progressivity, with the external growth case exhibiting stronger growth effects than the one in which there is intentional technology adoption.

Our work is related to several strands of literature on endogenous growth. On the modeling front, it is related to studies that feature heterogeneity and growth. Galor and Tsiddon (1997) analyze the connection between economic mobility, inequality, and growth in a model that features heterogeneity in children's ability and parental occupation. They use linearity in the human capital function to simplify aggregation. Our human capital investments are variable as opposed to their pre-specified

[3] By considering, broadly, the main sources of growth that have been extensively discussed in the new growth literature, we are able to study the sensitivity of growth response to the assumption made about its source. In contrast, most existing studies on taxes and growth focus on one engine of growth, say the unlimited accumulation of a factor of production that is not subject to diminishing returns.

sectoral requirements. Instead of linearity, we rely on limited heterogeneity (two types) to simplify aggregation and our focus is different from theirs. Andrade (1998) uses a model similar to ours. He does not focus on progressivity of taxes and growth; neither does he model the cost of education as the wage of the skilled. His analytical approach is also quite different from ours.

Overlapping generation models of endogenous growth have built-in heterogeneity. For example, Uhlig and Yanagawa (1996), consider a two-period overlapping generations model to argue that increasing capital income tax shifts income toward the young who have greater propensity to save and can thus increase the growth rate. Unlike their model, the heterogeneity in income we are capturing is among agents of the same generation. Blackenau and Ingram (1999) consider taxation of skilled and unskilled labor at different rates when there is skill biased technical change. As in Uhlig and Yanagawa, they characterize a life cycle effect of taxing older skilled labor at a higher rate (higher progressivity). This increases saving of the less taxed young who are mostly unskilled, increases the capital stock, and thus the return to skilled labor which is a complement to capital. The overall effect is an increase in the supply of skilled labor. Progressivity thus works through physical capital. In our model where capital accumulation plays a secondary role, the effect of progressivity is felt directly on the return to human capital, and has a negative effect on supply of skill in general. More importantly, the model in Blackenau and Ingram does not feature sustained endogenous growth and therefore cannot be used to address the key question of our paper.

On the substantive front our paper is related to Cassou and Lansing (2000), who also assess the effects of moving to a flat tax regime. Growth in their model is driven by investment in human and physical capital. Lower progressivity increases labor supply, output, investment, and growth. Under their calibration, they find growth effects of moving to a flat rate system range between 0.006 to 0.143 percentage points; they argue untaxed foregone earnings represent the largest input to production of human capital and this accounts for the modest increase. The most important difference between their study and ours is in the modeling of households. They solve a representative agent model and accommodate progressivity by specifying the tax rate as an increasing function of income. The mapping of this problem to the real world presumably involves the assumption that there is a representative family within which there is perfect insurance among family members who are in different tax brackets, which is hard to justify empirically. We take the stance that to understand progressivity we need to model inequality explicitly. Toward this end, we solve a fully specified heterogeneous agent model in which inequality evolves endogenously. Such a model can capture the differential effects of progressivity on the skill premium and welfare of agents in a way a representative agent model cannot. In spite of the significant differences between their study and ours, it is interesting that their higher end estimates of the growth effects are in the ballpark of our lower end estimates.

The rest of the paper is structured in the following way. In section 2 we provide some "back of the envelope" calculations for the growth effects of progressivity using the main marginal condition from Lucas (1990). We argue the effects are significant enough to warrant a more thorough investigation in a fully specified heterogeneous agent model. In Section 3 we describe the economic environment. The balanced growth path equilibrium is analyzed and characterized in Section 4. Section 5 is devoted to issues on calibrating the model to the U.S. economy. Section 6 considers various experiments in tax reform and Section 7 concludes.

2 Motivation and Preliminary Calculations

We start by examining the following marginal condition from Lucas' paper (equation (2.10)), which governs time spent by agents in skill accumulation:

$$w(t)\,h(t) = G'[v(t)] \int_t^{\infty} \exp\left\{-\int_t^s (r(\varsigma) - \lambda)\, d\varsigma\right\} u(s)\, w(s)\, h(s)\, ds. \tag{1}$$

Here, w is the rental rate of human capital, h the stock of human capital, G is a human capital production function that governs the evolution of human capital according to $\dot{h}(t) = h(t)\, G[v(t)]$, $v(t)$ is the time spent in accumulating human capital, $u(t)$ is the time spent working, r is the interest rate, and λ is the effective depreciation rate that includes population growth. In (1), the left hand side is the marginal cost of allocating an extra unit of time to human capital accumulation, which is the wage rate, and the right hand side is the marginal benefit, which is the marginal product weighted present value of future wages earned on account of this accumulation. It is easy to see that the tax rate drops out of this marginal condition because it affects the cost and benefit by the same factor; that is, if τ is the labor income tax, it cancels out of both sides. If human capital accumulation is modeled more realistically for the poor, as the process of acquiring skill when one is in a lower tax bracket and then moving to a higher bracket as happens in a progressive tax system, a tax change will not be neutral with respect to growth. By increasing the wedge between present and future tax rates, progressivity will reduce the return to human capital accumulation and decrease growth (i.e. decrease v and hence the growth rate $G[v]$).

To get a rough idea of the quantitative effect, we manipulated the above equation in Lucas (1990) in the following way. On a balanced growth path (BGP), h grows at a constant rate g, and w, r, u, and v are all constant. Note that $g = G[v]$. We use Lucas' functional form $G[v] = Dv^{\gamma}$. We also abstract from leisure, a conservative assumption in estimating growth effect of taxes, and assume $u + v = 1$. Suppose, at the time of accumulation, a lower tax rate of τ_s applies and future income is taxed at the higher rate of τ_c. It is fairly easy to show that the above condition reduces in such a

BGP to:

$$(1-\tau_s) = G'(v)\ u\left[\frac{1}{(r-\lambda)-g}\right](1-\tau_c)\,.$$

This cavalier slapping down of different tax rates at either side of a marginal condition derived from a representative agent model presumably involves the assumption of perfect consumption insurance among heterogeneously situated agents within a representative family. Using the usual Euler condition that characterizes growth, $g=\frac{r-(\rho+\lambda)}{\sigma}$, where ρ is the rate of time preference and σ is the coefficient of relative risk aversion, and using the functional form for G, this condition can be further reduced to the equation:

$$\{\Theta(\sigma-1)+\gamma\}\,Dv^{\gamma}+\Theta\rho=\frac{\gamma D}{v^{1-\gamma}},$$

where $\Theta\equiv\frac{(1-\tau_s)}{(1-\tau_c)}$ is a measure of progressivity of taxes. Progressivity increases when τ_s decreases or τ_c increases. For a given Θ, the left hand side is a strictly increasing function of v and the right hand side is a strictly decreasing function of v going from ∞ to zero as v goes from zero to one. A unique v and thus a unique growth rate exist. When the progressivity Θ increases, the left side shifts upward and becomes steeper, both of which serve to decrease v and the growth rate. *It is important to note that this growth effect is purely due to the progressivity of taxes;* when $\Theta=1$, there is no effect of tax changes on growth, as in Lucas (1990).

We adopt a benchmark of $\sigma=2$, $\gamma=0.8$, $\rho=0.015$, and $D=0.0595$ (the first three are consistent with the values Lucas (1990) uses, while D is chosen to pin down an annual per capita growth rate of about 1.8% when the progressivity parameter is 1.5). We consider $\Theta=0.75$, 1, 1.5, and 2.[4] A progressivity factor of 2 is not unreasonable in light of the tax rates for 1975 given earlier. More dramatic evidence of progressivity can be found for earlier decades from the *Statistical Abstract of the United States* – the marginal tax rate for single people was 22% at the lowest bracket and 78% at the highest bracket during 1954-63. We also vary, one at a time from the benchmark, the parameters σ, γ, and ρ. The growth rates for the various parameter combinations are given below:

Θ	**Benchmark**	Change σ to 1.5	Change γ to 0.65	Change ρ to 0.01
0.75	2.94%	3.63%	3.11%	3.12%
1.0	2.46%	3.14%	2.69%	2.67%
1.5	**1.81%**	2.40%	2.10%	2.05%
2.0	1.38%	1.87%	1.72%	1.63%

[4] The value of 0.75 implies a regressive tax system. We include it for sake of completeness; we do not expect it to be a politically viable alternative. Also, we will see later that a lower progressivity has an ambiguous effect on welfare.

There are significant growth effects of changes in progressivity – a movement from a progressivity of 1.5 to flat rate taxes increases the annual growth rate by 0.65 percentage points; a movement from a progressivity of 2.0 to flat rate taxes raises growth by 1.08 percentage points. These are large changes as far as growth rates go, especially when one realizes this is over and above any gains that could be realized by a decrease in capital income taxes or by considering an elastic labor supply.

This analysis shows that the issue of progressivity of taxes is worthy of further study in a fully formulated heterogeneous-agent model. Why go beyond this preliminary analysis? 1. The above estimates are likely to be an upper bound given the assumption of perfect consumption insurance and the efficient response to tax changes this causes. 2. There is also a distinct lack of mobility between the poor and the rich here. In reality, one cannot assume that one will end up in the richer group with certainty, an aspect that could dampen the effect of progressivity on the investment of the poor. 3. It is not possible to analyze the effect of liquidity constraints and heterogeneity using the above marginal condition from a representative agent model. 4. Estimates of welfare gains or losses by type of agent are also hard to provide without a fully specified model. The current tax and growth literature rarely provides estimates of welfare gains, especially when transitions are taken into account; as we will see, welfare estimates bring to light several issues that balanced growth rate comparisons alone do not. In the next section we outline the model we use to study this issue.

3 The Model

There are two sectors in the economy. The household sector supplies labor of differing skills and uses the wage for consumption and investment in human capital. This is the only sector that features heterogeneity. The production sector invests in physical capital and in productivity improvements.

3.1 Household Sector

Following Caucutt and Kumar (2000), we develop a model where the heterogeneity is limited to two types of skill levels.[5] Following the lead of Rogerson (1988), we achieve convexification by making the process of skill accumulation probabilistic. At any instant, the economy is populated by two types of adult agents we call "skilled" (subscripted c, in anticipation of the calibration, where these agents are identified with college-educated workers) and "unskilled" (subscripted s, for school-educated workers), with total measure one. There is no population growth. Let n_c be the fraction of skilled agents in the economy. Each adult has a child and can hire a skilled teacher for a fraction of the teacher's time, e, to educate her. With this input, the probability that the child of a type-i agent, $i = c, s$, becomes

[5] The model there is used to study the effect of subsidies in higher education. In addition to adding growth to that model, we abstract from heterogeneity in academic ability and allow varying levels of investment in skill.

skilled is given by $\pi_i(e_i)$; with probability $(1-\pi_i(e_i))$ the child fails and is an unskilled adult in the following period. The probability functions are subscripted because children of skilled agents might have advantages other than just higher education expenditure; they could have better schooling at the earlier levels, better role models, etc. That is, we expect $\pi_c(e) \geqslant \pi_s(e)$, $\forall\, e \in (0,1)$. Additionally, we assume:

$$0 < \pi_i < 1,\ \pi_i' > 0,\ \pi_i'' < 0,\ \pi_i(0) = 0.$$

The increasing nature of π needs little elaboration. We assume concavity in the probability function, in line with the diminishing return to instantaneous investment found in most models of education. When nothing is spent on the child's education, the child remains unskilled for sure as an adult. We use "skilled" ("unskilled"), "rich" ("poor"), and "college-educated" ("school-educated") interchangeably to refer to the two types of agents.[6]

The Bellman equation for a skilled agent, who takes wages as given, is:

$$V_c(n_c) = \max_{e_c}\left\{u\left((1-\tau_c)(1-e_c)w_c(n_c)\right) + \beta\pi_c(e_c)V_c\left(n_c'\right) + \beta(1-\pi_c(e_c))V_s\left(n_c'\right)\right\}. \quad (2)$$

The only state variable is the aggregate state, n_c. The tax rate on labor income of the skilled agents is τ_c, and that on the unskilled agent is τ_s. All agents posit a law of motion for the state variable, $n_c' = \Phi(n_c)$. The Bellman equation for the unskilled agent is similar, except the current return term is given by $u\left((1-\tau_s)(1-e_s p(n_c))w_s(n_c)\right)$. The unskilled agent also needs to hire a skilled person as a teacher, so the cost is $e_s w_c = e_s p w_s$, where $p \equiv \frac{w_c}{w_s}$ is the skill premium.[7]

The first order conditions for skill accumulation for the two types of agent are:

$$\beta\pi_c'(e_c)\Lambda\left(n_c'\right) = (1-\tau_c)w_c(n_c)u'\left((1-\tau_c)(1-e_c)w_c(n_c)\right), \quad (3)$$

$$\beta\pi_s'(e_s)\Lambda\left(n_c'\right) = p(n_c)(1-\tau_s)w_s(n_c)u'\left((1-\tau_s)(1-e_s p(n_c))w_s(n_c)\right). \quad (4)$$

where $\Lambda(n_c) \equiv V_c(n_c) - V_s(n_c)$, can be viewed as the value to being skilled. Inada conditions on the utility and probability functions have been assumed to get $0 < e_i < 1$. The left hand side

[6] Under an alternate interpretation, dynasties are infinitely lived and switch between skilled and unskilled states based on their investments in a stochastic skill accumulation technology. A skilled agent has to keep updating her skills (accumulate human capital) in order to stay skilled. An agent who is originally unskilled, has to accumulate human capital to *become* skilled. One could have only $\pi_s(0) = 0$ in this case. When an unskilled agent spends nothing on skill accumulation she will remain unskilled. However, when a skilled agent spends nothing she might have a small positive probability of staying skilled. The analytical characterization remains the same under either interpretation, but the calibration will depend on the stance taken. While Caucutt and Kumar (2000) connected failure in skill accumulation with college dropout rates, in the present context failure can additionally stand in for idiosyncratic labor productivity shocks, personal and small business bankruptcies, and shocks to the industry for which a particular skill is most suitable.

[7] Eicher (1996) also assumes that the cost of accumulating human capital is the value of skilled time.

is the marginal benefit of investing in human capital – the value to being skilled weighted by the discount factor and the marginal productivity of the investment. The right hand side is the cost of accumulating human capital, weighted by the agent's marginal utility.

Evaluating the Bellman equations for the two types of agents at the optimal policies $e_c(n_c)$ and $e_s(n_c)$ and subtracting one from the other, we get an expression of how the value to being skilled evolves:

$$\Lambda(n_c) = u(c_c(n_c)) - u(c_s(n_c)) + \beta(\pi_c(e_c(n_c)) - \pi_s(e_s(n_c)))\Lambda\left(n_c^{'}\right), \tag{5}$$

where:

$$\begin{aligned} c_c(n_c) &\equiv (1-\tau_c)(1-e_c(n_c))\,w_c(n_c), \\ c_s(n_c) &\equiv (1-\tau_s)(1-e_s(n_c)\,p(n_c))\,w_s(n_c). \end{aligned}$$

The value to being skilled has two parts – a current (potential) increase in utility from being skilled and a greater chance of realizing the future value of being skilled. The quantity $\beta(\pi_c(e_c) - \pi_s(e_s))$ can be interpreted as an endogenous discount factor which increase with the difference in investment between the two types.

The law of motion for the fraction of skilled workers can be written as:

$$\Phi(n_c) \equiv n_c^{'} = n_c\pi_c(e_c(n_c)) + (1-n_c)\,\pi_s(e_s(n_c)). \tag{6}$$

Equations (3) to (6) characterize the dynamics of the household sector through the four functions $e_c(n_c)$, $e_s(n_c)$, $\Lambda(n_c)$, and $\Phi(n_c)$ for any given wage functions $w_c(n_c)$ and $w_s(n_c)$. The matrix that gives the transition probabilities between the skilled and unskilled states is:

$$\begin{pmatrix} & \textbf{skilled} & \textbf{unskilled} \\ \textbf{skilled} & \pi_c(e_c) & 1-\pi_c(e_c) \\ \textbf{unskilled} & \pi_s(e_s) & 1-\pi_s(e_s) \end{pmatrix}.$$

The modeling decisions made regarding human capital accumulation have important implications for tax policy. For this reason, we discuss the more important assumptions below.

1. Agents investing in human capital are liquidity constrained. In condition (1), the only effect of progressivity is an opportunity cost effect – if accumulating poor agents are taxed at a higher rate (lower progressivity), their opportunity cost of not working decreases and their investment in human capital increases. Alternately, the return to being skilled increases, which increases investment. As we will see, this effect is preserved in our model. We have an additional effect coming from the income effect of the liquidity constraint; the lower progressivity is likely to

decrease investment of the poor relative to that of the rich. The economy-wide investment then depends on the relative strengths of investments made by the two types. The liquidity constraint can be used to motivate governmental interest in differential taxation. But, as seen from the analysis of (1) which supposes no such constraint, it is clearly not *required* for progressivity to affect growth; in fact, this is a conservative assumption for quantifying growth effects.

2. Human capital investment is subject to diminishing returns, in line with assumptions in most of the literature. The absence of concavity would lead to the empirically unsavory implication that skilled agents find investment so attractive that they are willing to put up with lower current utility than the unskilled agents. Diminishing returns, in the presence of liquidity constraints, also affects the investment by the two types to different degrees when the tax policy changes.

3. Human capital investment is tax exempt. With tax-exempt investment, changes in flat-rate taxes are growth neutral; hence, this case provides an isolated effect of the progressivity of taxes. This will also allow us to compare our work with those considered by Stokey and Rebelo (1995), who point out that small effects of taxes on growth follow from the empirically justifiable assumption of high factor shares for human capital in production and relatively light taxation of the human capital producing sector.[8] We will later consider the formulation where the tax is levied on the entire wages. With tax-exemption, the effect of tax changes on marginal costs (which we brand "liquidity effect") is more muted. For instance, an increase in tax rate on the rich decreases income, but provides an incentive to get tax exemption by spending more on human capital.

4. Both types of agents need skilled teachers' time to accumulate human capital. Having each type of agent use only her own time to become skilled is not only unappealing *a priori*, but also gives the empirically implausible result of poor agents investing more in human capital than the rich, since the poor have a lower opportunity cost. This is a modeling issue unique to a heterogeneous agent formulation of skill acquisition.[9]

[8] One interpretation of Lucas (1990) is that the human capital producing sector is completely untaxed. Cassou and Lansing (2000) note in their infinite horizon model that untaxed foregone earnings constitute the bulk of education costs.

[9] Of course this anamoly could also be avoided by having $\pi_c \gg \pi_s$. Lucas (1988) assumes that human capital accumulation evolves according to $\dot{h} = Buh$, where u is own time spent in skill acquisition. This is the lead followed by several tax-and-growth studies.

In the infinite horizon view of our model, while an opportunity cost interpretation can be given for those currently skilled, for the currently unskilled the cost is more than just foregone wages since they hire skilled teachers. In fact, the preceding discussion has been couched entirely in terms of teacher time, ignoring time spent by the workers themselves in getting educated. In reality, each type of agent will face a cost that depends both on the value of her own and her

3.2 Production Sector

In addition to the two types of household agents, there is a third type of agent, an infinitely-lived entrepreneur, who carries out production and has preferences identical to the two types of workers. She produces according to the production function:

$$Y = A^{1-\alpha} K^{\alpha} \left[\theta N_c^{\nu} + (1-\theta) N_s^{\nu}\right]^{\frac{1-\alpha}{\nu}}, \tag{7}$$

where N_c is the measure of skilled labor hired and N_s the measure of unskilled labor, and $0 < \nu < 1$. Here K is the physical capital used in production, which we assume is accumulated only by the producer. With this formulation, we have limited the heterogeneity to skill accumulation and kept physical capital accumulation simple.[10]

This entrepreneur consumes according to,

$$c_e = (1-\tau_e)(Y - w_c N_c - w_s N_s) - I. \tag{8}$$

Here, τ_e is the tax rate on the entrepreneur's profits, and I is the investment in physical capital, which evolves according to:

$$K' = I + (1-\delta) K. \tag{9}$$

Unlike human capital investments, physical capital investment is not tax exempt. The Bellman equation for the third type of agent is:

$$W(K, A) = \max_{N_c, N_s, I} \left\{ u(c_e) + \beta W\left(K', A'\right) \right\}, \tag{10}$$

subject to (7), (8), (9), and the law of motion for A to be described below.

We consider two growth specifications that are commonly alluded to in the literature – productivity improvement arising as an externality and arising due to intentional use of human capital by the firm.[11] By broadly considering the main sources of growth that have been extensively discussed in the new growth literature, we are able to study the sensitivity of growth response to the assumption made about its source.

teacher's time. As long as we are willing to accept the necessity for teachers to be skilled, the unit time cost for the unskilled agent will be greater than her own pay.

[10] However, this implies the stance that differences in labor income taxes affect growth more than differences in capital income taxes.

[11] A highly abbreviated list of examples of externality driven models are Romer (1986), Jones and Manuelli (1992), and Stokey (1992), and a few examples of intentional adoption / accumulation models can be found in Romer (1990), Aghion and Howitt (1998), and Grossman and Helpman.(1991). The model in Lucas (1988) features production externality in human capital, though the model can generate growth even without this.

3.2.1 External Growth

In the first specification, growth results due to an external effect which depends on the fraction of skilled workers alone. This is a refinement of human capital externality assumptions in the growth literature that date back at least to Lucas (1988); there, the average human capital of the *entire* economy causes an external effect. We assume that the productivity parameter in the production function evolves according to:

$$A_{t+1} = (1 + \xi (N_c)) A_t, \tag{11}$$

where ξ is the externality function. That is, the mere hiring of skilled employees in the production process is enough to generate productivity improvements; they will not be compensated for it. Assuming competitive labor markets, optimization by the entrepreneur implies that the skill premium is given by:

$$p = \frac{\theta}{1-\theta} \left(\frac{N_s}{N_c} \right)^{1-\nu}. \tag{12}$$

It is easy to see that entrepreneurial before-tax profits are given by αY, and she has to consume, invest, and pay taxes out of it. The first order and envelope conditions for the dynamic program (10) are:

$$\begin{aligned} [I] &: \beta W_1 \left(K', A' \right) = u'(c_e) \\ [ENV_k] &: W_1(K, A) = \alpha (1 - \tau_e) \frac{Y}{K} u'(c_e) + \beta (1 - \delta) W_1 \left(K', A' \right). \end{aligned} \tag{13}$$

3.2.2 Intentional Technology Adoption

A second specification for growth features intentional employment and compensation of skilled workers generate productivity improvements. Each period, the entrepreneur hires a measure N_c of skilled workers, out of which she employs a measure N_{cA} for new technology adoption and productivity improvements. The production function is therefore of the form:

$$Y = A^{1-\alpha} K^{\alpha} \left[\theta (N_c - N_{cA})^{\nu} + (1 - \theta) N_s^{\nu} \right]^{\frac{1-\alpha}{\nu}}. \tag{14}$$

The productivity parameter is then assumed to evolve according to:

$$A_{t+1} = (1 + \xi (N_{cA})) A_t. \tag{15}$$

We use the same ξ for both specifications purely for notational simplicity; they can be different functions.[12] The key difference is that the measure N_{cA} of workers are hired by the firm and are

[12] In fact, the discipline of calibration might dictate a different specification.

compensated for it. By assuming an infinitely-lived entrepreneur, we are sidestepping industrial organization issues that form a central part of most R&D based models of growth. These are quite important, but do not seem to be of first order importance for the question at hand. The skill premium implied by competitive labor markets and optimizing behavior by the entrepreneur is given in this case by:

$$p = \frac{\theta}{1-\theta}\left(\frac{N_s}{N_c - N_{cA}}\right)^{1-\nu}. \tag{16}$$

Here, it is $w_c\left(N_c - N_{cA}\right) + w_s N_s = (1-\alpha)Y$, and out of the remaining αY, the entrepreneur has to invest in technology improvements by paying $w_c N_{cA}$, invest in physical capital, consume, and pay taxes. That is, the entrepreneur has to make two types of investment – in physical capital, and in technology improvements. For simplicity, we assume that adoption involves only skilled labor and no physical capital; the implicit assumption in (8) is that "R&D costs", $w_c n^*_{cA}$, are exempt from the tax on profits.

The first order and envelope conditions for the entrepreneur in this case include (13) and the following additional conditions:

$$\begin{aligned} [N_{cA}] &: \quad \beta A \xi'\left(N_{cA}\right) W_2\left(K', A'\right) = \left(1-\tau_e\right) w_c u'\left(c_e\right) \qquad (17)\\ [ENV_A] &: \quad W_2(K,A) = \left(1-\tau_e\right)(1-\alpha)\frac{Y}{A}u'\left(c_e\right) + \beta\left(1+\xi\left(N_{cA}\right)\right) W_2\left(K', A'\right), \end{aligned}$$

where $w_c = \frac{\partial Y}{\partial N_c} = \theta(1-\alpha)A^{1-\alpha}K^{\alpha}\left[\theta\left(N_c - N_{cA}\right)^{\nu} + (1-\theta)N_s^{\nu}\right]^{\frac{1-\alpha}{\nu}-1}\left(N_c - N_{cA}\right)^{\nu-1}$. The first condition effectively equates the marginal contribution of skilled agents in the two uses it can be put to, technology adoption and production. The second condition states that the benefit of an extra unit of the technology stock is its contribution to current marginal utility via production and its use in future technology improvements arising from (15).

4 Balanced Growth Analysis

In this section, we analyze the effects of a parametric change in tax progressivity on the balanced growth of the economy; we conduct transitional and welfare analyses later. The role of the government is limited to collecting taxes; all collected taxes are spent by the government and do not result in any utility or productivity improvements.

4.1 Definition of Balanced Growth Equilibrium

We use the CRRA utility, $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$, which is the only separable preference specification consistent with balanced growth. We envision all three types of agents situated on a balanced growth equilibrium, making decisions optimally.

Definition: A Balanced Growth Path (BGP) equilibrium is a collection of the quantities $\{K,\ A,\ Y,\ c_e,\ e_c^*,\ e_s^*,\ n_c^*,\ N_c^*,\ N_s^*,\ N_{cA}^*,\ \Lambda,\ W\}$, prices $\{w_c,\ w_s,\ p^*\}$, and a tax policy $\{\tau_c,\ \tau_s,\ \tau_e\}$ such that:

- K, A, Y, c_e, w_c, and w_s all grow at a constant rate g. The skill "return" function Λ, and entrepreneur value function W grow at the gross rate $(1+g)^{1-\sigma}$.[13]

- Human capital investments, e_c^*, e_s^*, the skill premium p^*, and skill attainment n_c^* are all time invariant.

- Given the constant growth paths of w_c, w_s, Λ, the time-invariant p^*, and taxes τ_c, τ_s, the human capital investments e_c^*, and e_s^* solve the household problems given in (2).

- When growth is driven by externality, given τ_e, w_c, and w_s, the quantities N_c^*, N_s^*, c_e, and I, solve (10) subject to (7), (8), (9), and (11). When growth is driven by intentional technology adoption, given τ_e, w_c, and w_s, the quantities N_c^*, N_s^*, N_{cA}^*, c_e, and I, solve (10) subject to (14), (8), (9), and (15).

- The skill return function, Λ, that households posit is consistent with the human capital investment decisions; that is, (5) is satisfied at the BGP quantities.

- The law of motion for the skilled fraction of labor force, Φ, that households posit is consistent with household decisions; that is, (6) is satisfied at the BGP quantities.

- The labor market clears; i.e. $N_c^* = n_c^*$, and $N_s^* = (1 - n_c^*)$.[14]

- In the externality case, $g = \xi(N_c^*)$ and in the technology adoption case, it is given by, $g = \xi(N_{cA}^*)$.

Even though both types of wages grow at the gross rate $(1+g)$, for individuals transiting between the two states, the gross growth rates will be given by:

$$\begin{pmatrix} & \textbf{skilled} & \textbf{unskilled} \\ \textbf{skilled} & (1+g) & (1+g)/p \\ \textbf{unskilled} & (1+g)\,p & (1+g) \end{pmatrix}.$$

[13] The utility function is the same for all agents and is homogeneous of degree $(1-\sigma)$. It follows that $\Lambda\left(n_c'\right) = (1+g)^{1-\sigma}\Lambda(n_c)$, and $W(K', A') = (1+g)^{1-\sigma} W(K, A)$ on the BGP.

[14] Note that we have not accounted for the skilled labor taken by the "teaching". This is done for simplicity; we expect the labor involved in teaching, $e_s(1-n_c) + e_c n_c$, to be a small fraction of the labor force. In data, according to *Education at a Glance: OECD Indicators 1997*, US teaching staff involved in the all levels of education was only 3.2% of the total employed population in 1995, out of which 0.7% was involved in tertiary education. What seems to be of first order importance is to have the cost of teacher's time enter the marginal conditions for human capital accumulation.

4.2 Production Sector on the BGP

Since the production sector has no heterogeneity, it is easiest to start our BGP analysis there. It is convenient to think of the production sector (the entrepreneur) as taking the supply of skill as given, and making production decisions that result in a particular growth rate. This n_c^* versus g schedule can be thought of as the human capital "demand" curve. We will derive a "supply" curve when we analyze the household sector. This way of analyzing the BGP sheds light on the effects of tax policy on each sector.

4.2.1 External Growth

Given the n_c^* that results from the skill acquisition of households, the balanced growth is given in this case by (15) as:

$$g = \xi\left(n_c^*\right). \tag{18}$$

The skill premium on the BGP from (12) is:

$$p^* = \frac{\theta}{1-\theta}\left(\frac{1-n_c^*}{n_c^*}\right)^{1-\nu}. \tag{19}$$

These are the only two conditions needed from the production sector for balanced growth determination. Note that the tax rate on the entrepreneur's profits, τ_e, does not affect the long run growth rate. The "engine" of growth is skill acquisition and a tax policy that does not affect that process will have no effect on long run growth. The tax rate on profits will affect the capital-output ratio, and the levels of profits and wages. On the BGP, given the homogeneity of W (of degree $1-\sigma$), we can write $W_1\left(K', A'\right) = (1+g)^{-\sigma} W_1(K, A)$. Use this and the conditions in (13) to get:

$$(1+\rho)(1+g)^{\sigma} = \alpha(1-\tau_e)\frac{Y}{K} + (1-\delta),$$

where $\beta \equiv \frac{1}{1+\rho}$. This is the analogue of the continuous-time growth condition $\rho + \sigma g = r$, where the interest rate, gross of depreciation, is given as $\alpha(1-\tau_e)\frac{Y}{K}$. But given the growth rate determined by skill acquisition, the interest rate and capital-output ratio merely adjust according to the above condition; they do not determine growth. In particular, the capital-output ratio is given by:

$$\frac{K}{Y} = \frac{\alpha(1-\tau_e)}{(1+\rho)(1+g)^{\sigma} - (1-\delta)}. \tag{20}$$

Higher the taxes on profits, lower this ratio. If investment in physical capital were tax exempt, even this effect of the profit tax disappears, as higher taxes create an incentive to invest and get a "write-off".

The n_c^* vs g demand curve for the external growth case is trivial. It just follows from (18); it does not hinge on any entrepreneurial decision. Higher the availability of skilled labor, higher the spillover and higher the growth. Therefore, independent of σ, the "demand" curve slopes upward in g.

4.2.2 Technology Adoption

For any given n_c^*, the entrepreneur in this case has a decision to make about the fraction of the skilled labor force to devote to technology adoption, n_{cA}^*, which affects the growth rate through the equation, $g = \xi\left(n_{cA}^*\right)$. Again, using $W_1\left(K^{'}, A^{'}\right) = (1+g)^{-\sigma} W_1(K, A)$ and using (17), we can write the equation determining n_{cA}^*, and hence g, as:

$$(1+\rho)\left(1+\xi\left(n_{cA}^*\right)\right)^{\sigma} = (1-\alpha)\frac{Y}{w_c}\xi^{\prime}\left(n_{cA}^*\right) + \left(1+\xi\left(n_{cA}^*\right)\right) \tag{21}$$

where $\beta \equiv \frac{1}{1+\rho}$, and $\frac{Y}{w_c}$ can be backed out from $[n_c]$ as:

$$\frac{Y}{w_c} = \frac{\left(n_c^* - n_{cA}^*\right)^{1-\nu}\left[\theta\left(n_c^* - n_{cA}^*\right)^{\nu} + (1-\theta)\left(1-n_c^*\right)^{\nu}\right]}{\theta(1-\alpha)}.$$

The skill premium on the BGP from (12) is:

$$p^* = \frac{\theta}{1-\theta}\left(\frac{1-n_c^*}{n_c^* - n_{cA}^*}\right)^{1-\nu}. \tag{22}$$

These are the only production conditions needed for balanced growth determination.

Note that even when growth is driven by technology adopted in the production sector, the tax rate on profits, τ_e, does not affect the rate of this growth. In a world in which R&D expenses are tax exempt, an increase in this tax rate decreases the marginal cost of hiring a skilled agent to adopt technology and decreases the marginal benefit arising from the improved technology by the same factor, as is evident from (17). This is similar to the tax neutrality seen in (1).[15] The capital-output ratio continues to be given by (20) and is adversely affected by a tax on profits if physical capital investment is not tax exempt.

What happens when the available pool of skilled labor, n_c^*, increases? Using (15), (21), and (22), and parametrizing $\xi\left(n_{cA}\right) = C n_{cA}^{\varepsilon}$, $0 < \varepsilon < 1$, we prove the following lemma in the appendix.

Lemma 1 *The adoption firm's decision function $n_{cA}^*\left(n_c^*\right)$ is strictly increasing. That is, when the availability of skilled labor increases, the portion of that labor devoted to technology adoption increases in a BGP equilibrium.*

It follows that $g\left(n_c^*\right)$ is a strictly increasing function as it is in the externality case; the increasing nature is independent of σ. And, as in any growth model, the growth rate decreases with σ or ρ for any given n_c^*.

[15] And here too, heterogeneity in profits and progressive taxes could have an effect on the growth rate.

If R&D expenses are not deductible, the entrepreneur's consumption becomes

$$c_e = (1-\tau_e)\left(Y - w_c N_c - w_s N_s\right) - I - \tau_e w_c n_{cA}^*.$$

Now the first term of the right side of (21) will have a factor $(1-\tau_e)$, and the tax rate on profit will have a negative effect on the growth rate.

4.3 Household Sector on the BGP

Analogous to the analysis of production decisions, it is convenient to think of the household sector as anticipating a particular growth rate g, the resulting skill premium p^*, and making human capital investment decisions that imply a supply of skill n_c^*. This n_c^* versus g schedule can then be thought of as the human capital "supply" curve. The intersection of this curve with the "demand" curve gives the balanced growth rate and the stationary skill level on the balanced growth path.

We informally summarize the forces that govern skill acquisition first. An increase in progressivity decreases investment by the rich relative to that of the poor when $\sigma > 1$ (liquidity constraint effect dominates) and increases it if $\sigma < 1$ (tax-exempt effect dominates). An increase in the equilibrium premium decreases investment by the poor relative to the rich as education becomes costlier for them, a "tuition" effect. On the intertemporal front, an increase in the equilibrium premium or a decrease in the progressivity increase the value to being skilled and tends to increase investment by *both* types of agents. An increase in anticipated growth increases the discount factor when $\sigma < 1$ (a substitution effect of increased growth), increasing investment of both types, and works in the opposite direction when $\sigma > 1$ (an income effect of increased growth).

Formally, using (3) and (4), we can get the *intra*temporal condition governing investment of the skilled agents relative to those of the unskilled as:

$$\frac{\pi_c'\left(e_c^*\right)}{\pi_s'\left(e_s^*\right)} = \frac{\Theta^\sigma}{\Theta}\left(\frac{\frac{1}{p^*} - e_s^*}{1 - e_c^*}\right)^\sigma, \tag{23}$$

where $\Theta \equiv \frac{(1-\tau_s)}{(1-\tau_c)}$ is a measure of the progressivity of the tax system. This expression equates the ratio of marginal investment benefit of each type to the ratio of marginal their marginal costs.

It can be seen that an increase in the anticipated skill premium increases the cost of skill acquisition for the unskilled relative to their income and increases e_c^* relative to e_s^*, a "tuition" effect. The Θ^σ in the numerator of the right hand side captures what we earlier termed the liquidity effect. An increase in Θ lowers the income of the skilled relative to that of the unskilled, causing investment by skilled people to go down relative to that of the unskilled; i.e. $\frac{e_s^*}{e_c^*}$ tends to increase given the concavity of π. The Θ in the denominator captures the effect of tax-exemption alluded to earlier. An increase in Θ decreases the marginal cost for the skilled and increases their incentive to increase their investment relative to that of the unskilled to essentially get a tax "write-off". The net effect of Θ in determining the relative investment levels clearly depends on σ. When $\sigma > 1$, the liquidity constraint effect dominates and increase in progressivity lowers the investment by the skilled relative to that of the unskilled.[16] With log utility, both cancel out and the only effect of progressivity is the effect on the return which we will turn to next. Since the effect of the liquidity constraint is diluted

[16] If each agent could use her own time for skill accumulation (i.e. hire an agent of her own type as the teacher), the

by the tax-exempt status, we will also perform computations where human capital investments are not exempt from taxes.

Using the fact, $\Lambda\left(n_c^{'}\right) = (1+g)^{1-\sigma}\Lambda(n_c)$, on the BGP, we can use (5) and write a normalized version of the return to being skilled (an *inter*temporal condition) as:

$$\frac{\Lambda(n_c^*)}{[(1-\tau_s)w_s^*]^{1-\sigma}} = \frac{1}{\left[1-\beta(1+g)^{1-\sigma}(\pi_c(e_c^*)-\pi_s(e_s^*))\right]} \cdot \frac{\left[(1-e_c^*)\frac{p^*}{\Theta}\right]^{1-\sigma} - [1-e_s^* p^*]^{1-\sigma}}{1-\sigma}.$$

The first part of the right side is an effective discount factor and the second is an excess utility term. The discount factor increases with β as well as with increased investment by the skilled. The discount factor increases with anticipated growth, g, if $\sigma < 1$ and decreases if $\sigma > 1$. In the former case, the substitution effect of an increase in anticipated growth dominates, increasing the incentive to invest, while in the latter case, the income effect of increased growth dominates, decreasing the incentive to invest.

The excess utility term, and thus the normalized return, increases with p^* and decreases with the progressivity parameter Θ, no matter what σ is. One can think of $\frac{p^*}{\Theta}$ as the effective premium that determines the returns to human capital. An increase in the return Λ will tend to increase both e_s^* and e_c^* according to (3) and (4).

With flat rate taxes, $\tau_c = \tau_s = \tau$, $\Theta = 1$, and the actual tax rate does not figure in the equations that determine the growth rate. Any effect of tax on growth is because of differences in its *structure*, rather than on its level.

Taking into account the evolution of Λ, we can rewrite (4) as:

$$\frac{\Lambda(n_c^*)}{[(1-\tau_s)w_s^*]^{1-\sigma}} = \frac{1}{\beta(1+g)^{1-\sigma}} \frac{p^*}{\pi_s^{'}(e_s^*)(1-e_s^* p^*)^{1-\sigma}}.$$

Equating the right hand sides of the above two expressions yields the following important equation:

$$\left[\frac{1}{\frac{1}{\beta(1+g)^{1-\sigma}} - (\pi_c(e_c^*)-\pi_s(e_s^*))}\right] \cdot \left[\frac{\left(\frac{1-e_c^*}{\Theta}\right)^{1-\sigma} - \left(\frac{1}{p^*} - e_s^*\right)^{1-\sigma}}{1-\sigma}\right] = \frac{1}{\pi_s^{'}(e_s^*)\left(\frac{1}{p^*} - e_s^*\right)^{\sigma}}. \quad (24)$$

Evaluating (6) at the BGP equilibrium, we get:

$$n_c^* = \frac{\pi_s(e_s^*)}{1-(\pi_c(e_c^*)-\pi_s(e_s^*))}. \quad (25)$$

---

relative investment condition for $\pi(e) = Be^{\gamma}$ becomes:

$$\left(\frac{e_s^*}{e_c^*}\right)^{1-\gamma} = \left(\frac{p^*}{\Theta}\right)^{1-\sigma}\left(\frac{1-e_s^*}{1-e_c^*}\right)^{\sigma}.$$

If the net of tax skill premium $\frac{p^*}{\Theta} > 1$, as we would expect in a BGP equilibrium, $e_c^* < e_s^*$. This outcome is highly counterfactual. Making investment by skilled more productive or assuming the cost of education relative to income is higher for the unskilled is needed to remedy the situation.

As one would intuitively expect, higher the investment in skill by any particular type of agent on the BGP, higher the level of skill attainment; n_c^* is increasing in e_s^*, e_c^*.

Equations (23), (24), and (25) capture the behavior of the household sector on the BGP equilibrium. That is, given the p^* and g arising from production decisions, these three equations determine the investments e_c^* and e_s^*, and thus the skill attainment n_c^*. In practice, to analyze the human capital "supply" curve it is convenient to consider the expression for the premium, (19) or (22), also in this system.

We are now in a position to assess how the above-mentioned forces affect human capital investment response and thus skill attainment to a change in tax policy. First consider the response of investments to a change in progressivity for a given growth rate. An increase in progressivity decreases the value to being skilled, Λ, and tends to decrease investment by both types. But can an increase in investment by the unskilled through the liquidity effect offset this decrease? To investigate this, suppose for a moment that agents do not have foresight about changing wages; that is, ignore the general equilibrium effects of investment decisions of agents. (Alternately one could think of ν being 1; the two types of labor are perfectly substitutable then and the premium stays constant for a given growth rate.) The analysis that ignores general equilibrium is common to both types of growth. We will incorporate the general equilibrium effect, as well as consider what happens when progressivity is fixed but anticipated growth changes, for each type of growth later. In the appendix, using the parametrization $\pi_c = \pi_s = Be^{\gamma}$, $0 < \gamma < 1$, we show:

Lemma 2 *For a given rate of anticipated growth, no matter its source, when general equilibrium effects of changes in the skill premium are ignored, the BGP investments e_c^*, and e_s^*, both decrease with the degree of tax progressivity; the level of skill attainment, n_c^*, thus decreases.* •

This result holds for any $\sigma > 0$. Even though an increase in the progressivity could shift the investment in favor of the unskilled through the liquidity effect, it is the intertemporal effect that ultimately dominates and decreases the investments of both types of agents. Next we consider what happens when we take into account the general equilibrium effects of changes in the skill premium. Since the premium depends on the type of growth, we will analyze the effects separately.

4.3.1 General Equilibrium Effects: External Growth

Use (19) and (25) to get the premium for the chosen parametrization as:

$$p^* = \frac{\theta}{1-\theta}\left(\frac{1 - B\left(e_c^*\right)^{\gamma}}{B\left(e_s^*\right)^{\gamma}}\right)^{1-\nu}. \tag{26}$$

As shown in Lemma 2, an increase in Θ decreases both e_c^* and e_s^*. But from the above equation we can see this tends to increase the premium and thus the value to being skilled. This in turn tends to *increase* the investment by both types now. However, the effect of increased tuition reinforces the original effect and decreases e_s^*. The general equilibrium effect is mildest when ν is high enough (high enough substitution between the two types of labor) and when γ is low enough (there is enough diminishing returns in the human capital accumulation process to keep elasticities of response low). In the appendix, we give sufficient conditions that σ, γ, and ν must satisfy, and show:

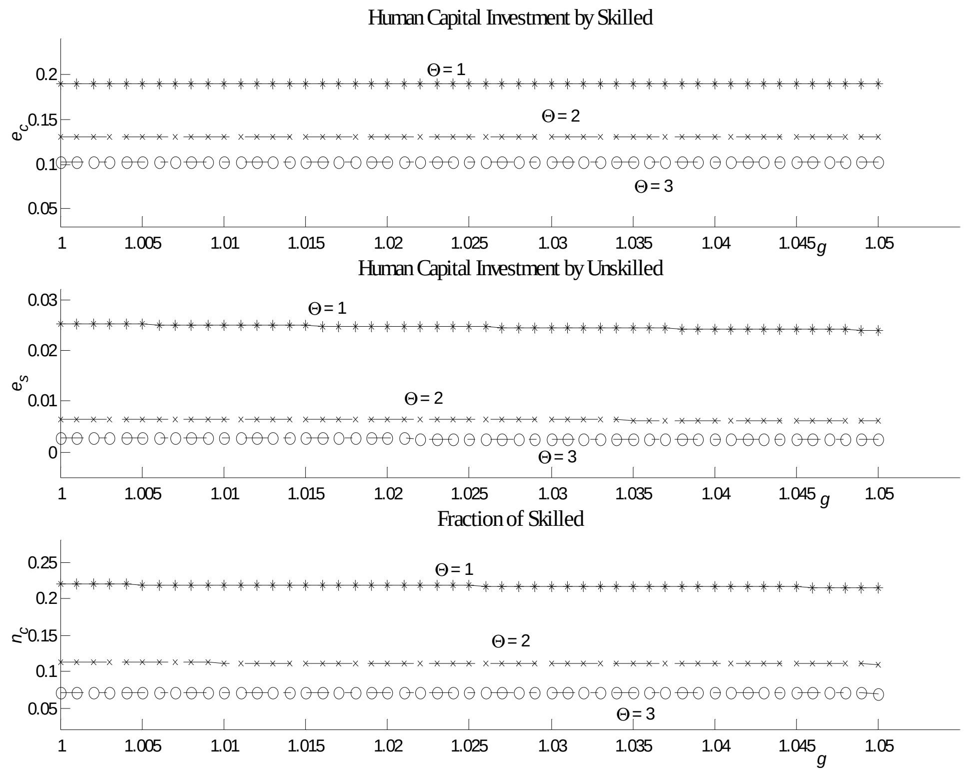
Lemma 3 *For a given rate of anticipated external growth, even when general equilibrium effects of changes in the skill premium are accounted for, the BGP investments e_c^*, and e_s^*, both decrease with the degree of tax progressivity, and therefore, so does the level of skill attainment n_c^*.* •

Next consider the response of human capital investments to the anticipated growth rate g, for a given level of progressivity, Θ. The only expression g enters is the intertemporal condition (24). As mentioned above, the effect of anticipated growth depends on the value of σ. When $\sigma < 1$ the substitution effect of an increase in anticipated growth dominates, increasing the incentive to invest (e_c^* and e_s^* increase), and when $\sigma > 1$, the income effect of increased growth dominates, decreasing the incentive to invest (e_c^* and e_s^* decrease).[17] In other words, the n_c^* versus g schedule shifts down with increased progressivity and is upward or downward sloping, depending on $\sigma < 1$ or $\sigma > 1$.

To illustrate the effects discussed above with a numerical example, we set annual $\beta = 0.98$, $\pi(e) = e^{0.5}$, $\alpha = 0.5$, $\theta = 0.5$, $\nu = 0.35$, and $\xi(n_c) = 0.08 n_c^{0.95}$. In Figure 1, we plot the investment by the two types of agents and the 'n_c^* versus g' "supply" schedule, when $\sigma = 1.5$. The investment curves are slightly downward sloping, and therefore so is n_c^* versus g.

[17] The fact that this simple intuition carries through when equilibrium effects of a change in premium are taken into account can be shown, analagous to Lemmas 2 and 3. However, we do not present the results here, as this is not our main focus. As we will see the results are insensitive to which effect of an increase in anticipated growth dominates. Kumar (2000) also discusses the two effects of anticipated growth and finds cross-country evidence that the income effect dominates.

Figure 1



The graphs for $\sigma = 0.5$ (not shown) are very similar. The only difference is that the investment curves and n_c^* versus g curve are now all upward sloping. However for any given anticipated growth rate, the investment levels and skill attainment decrease with the progressivity. When these household responses are put together with the upward sloping "demand" curve $n_c^* = \xi^{-1}(g)$, the following proposition emerges.

Proposition 1 *When growth is caused by externalities resulting from activities of skilled labor, an increase in the progressivity of taxes decreases the human capital investment levels of both types of agent on the BGP. The stationary level of skill attainment is lower and thus the growth rate is lower for all* $\sigma > 0$. $\bullet$

For instance, for the numerical example used to get the above graphs, we get growth rates of 1.88%, 1%, and 0.65% for a progressivity parameter, Θ, of 1, 2, and 3, respectively, when $\sigma = 1.5$.

4.3.2 General Equilibrium Effects: Technology Adoption

As mentioned earlier, Lemma 2 continues to hold when growth arises from adoption as well; when the equilibrium effect of investment changes on the premium is ignored, an increase in the progressivity unambiguously decreases investment by both types of agents and the level of skill attainment. However, the expression for the skill premium is different from that under externality, and we will see this alters behavior significantly, since the general equilibrium effect of investment changes are now different. The premium is now given by

$$p^* = \frac{\theta}{1-\theta}\left(\frac{1-B\left(e_c^*\right)^{\gamma}}{B\left(e_s^*\right)^{\gamma}-n_{cA}^*}\right)^{1-\nu}, \tag{27}$$

where $n_{cA}^* = \xi^{-1}(g)$. In addition to entering the discount factor, the anticipated growth rate directly enters the expression for expected premium; higher the fraction of labor in technology adoption, higher the premium. Skilled workers are fully compensated for their role in generating growth.

The direction of investment changes with progressivity, for a given anticipated growth rate, is not as clear-cut as it was with external growth. In the appendix, we show the following.

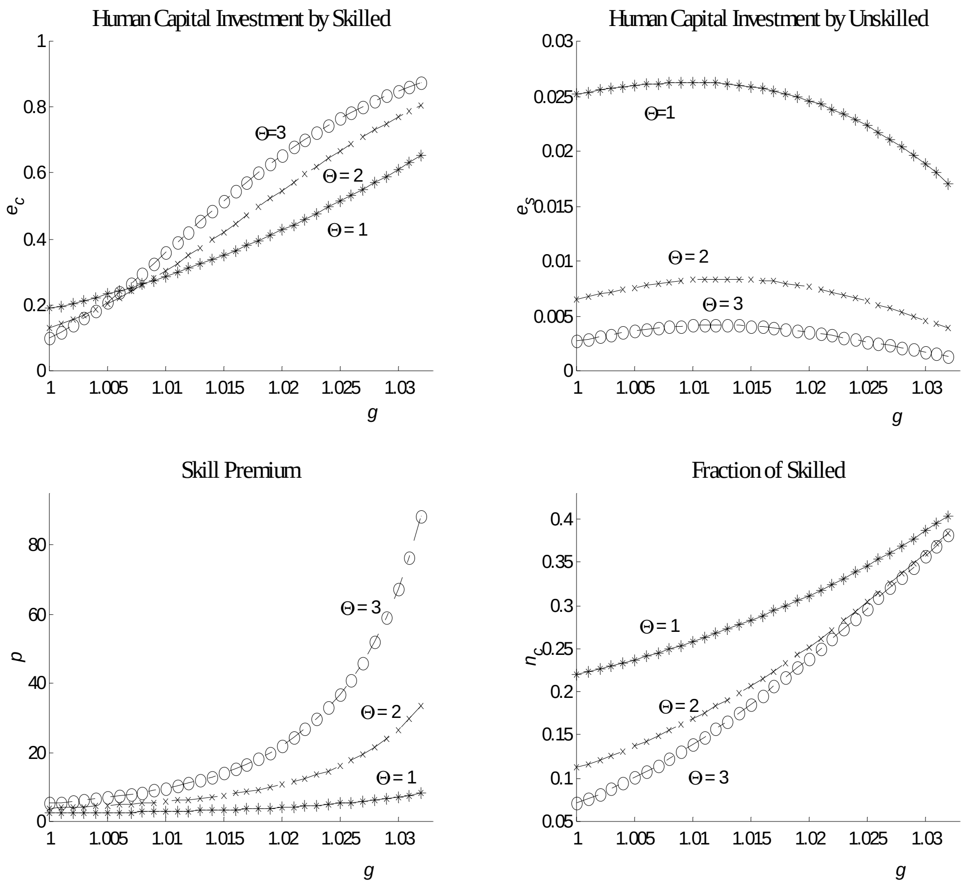
Lemma 4 *When growth is driven by technology adoption, general equilibrium effects of changes in the skill premium imply that the BGP investment of the unskilled, e_s^*, always decreases with the degree of tax progressivity. If the anticipated growth rate is low enough, the investment of the skilled, e_c^*, also decreases, but if the growth rate is high enough, e_c^* can increase with progressivity. Therefore, for low enough growth rates, the level of skill attainment n_c^* decreases with progressivity, but for higher growth rates the effect of progressivity on n_c^* is ambiguous.* •

The equilibrium effect is stronger when the anticipated growth is higher. In the high growth region, the decrease in e_c^*, e_s^*, arising from the forces outlined in Lemma 2 increases p enough to reverse the overall effect on the investment of the skilled while reinforcing it through the tuition effect on the investment of the unskilled.

We use the same parameters as used for the external growth case (interpreting $\xi(n_c)$ as $\xi(n_{cA})$) and illustrate the above-mentioned effects with the following graphs. For a given progressivity, the discount factor effect of an anticipated change in growth is swamped out by the return and tuition effects arising from the increased premium that accompanies increased growth. Therefore, σ plays a minimal role and we show the graphs only for $\sigma = 1.5$; the ones for $\sigma = 0.5$ are very similar.

Figure 2 shows investments done by the two types of agents, the behavior of the premium, and the human capital "supply" curve.

Figure 2



Both investments decrease with the degree of progressivity when the anticipated growth rate is low, similar to the external growth case. But as discussed in Lemma 4, when the anticipated growth rate is high enough, the general equilibrium effect on the premium makes it attractive for the skilled to invest more when progressivity *increases*; the unskilled investment decreases as in the externality case. As can be seen from Figure 2, the premium becomes more sensitive to progressivity at these higher growth rates, which explains the differing effects for high growth rates.

What does this imply for the human capital "supply" curve? At low growth rates, when the investment of both types decreases with progressivity, there is no analytical ambiguity. However at higher rates, when the investment of the two types move in the opposite directions, the analytical result is ambiguous. However, for our numerical example, it appears the decline of investment by the majority of the unskilled outweighs the increase by the skilled so that the overall supply of skill decreases with progressivity, though by smaller amounts.

The effect of anticipated growth rate on skill accumulation for a given level of progressivity is not clear-cut either. Anticipated growth continues to matter through the effective discount factor, $\beta(1+g)^{1-\sigma}$, in (24). Now it also matters through its effect on the premium. Increased anticipated growth increases the premium and therefore the return to education, which increases investment of both types for low growth rates; however, for sufficiently high growth rates the tuition effect of an increased premium dominates and decrease e_s^*.

Taken with the upward sloping "demand" curve resulting from Lemma 1, we get the result that the growth rate decreases with progressivity in this numerical example. The growth rates are 1.42%, 0.54%, and 0.15% for a progressivity parameter, Θ, of 1, 2, and 3, respectively, when $\sigma = 1.5$.

5 Calibration

We now calibrate the above model as a two-period overlapping generation model with parents spending resources on their children's education to U.S. data, and conduct a few policy experiments. The two types of labor can be readily interpreted as school-educated and college-educated labor and the premium as a college premium. The obvious limitations of calibrating a two period economy to match U.S. data are mitigated by the focus on long run growth rates; the present setup affords as much richness in calibration as, if not more than, current tax and growth studies which employ infinite horizon representative agent models.[18]

Our calibration strategy is guided by a desire to make the model economy consistent with the US economy not only along standard dimensions such as the capital-to-output ratio, but also along those particularly relevant to our study, such as the fraction of labor force that is college educated, the annual per capita growth rate, the share of GDP devoted to education, and the two independent entries in the long-term earnings mobility matrix. We set the capital's share of income α to 0.36. We set ν to 0.35 which implies an elasticity of substitution between skilled and unskilled labor of 1.54. Autor, Katz, and Krueger (1998) report that most estimates for this elasticity fall between 1.4 and 1.5. We use this value of ν, *anticipate* a college attainment of 35% and a college premium of 1.75 (both of which are consistent with values reported in the labor literature), and use (19) to get θ as 0.527. We set σ to 2.0, a standard value for the utility curvature parameter.[19] The tax on profits, τ_e, is set to 20%, and the annual physical capital depreciation rate is set to 4.4%.[20] We assume

[18] The alternative suggested earlier, of infinitely lived dynasties that transit between the rich and poor states, while appealing because it is open to interpretations of skill accumulation that are broader than just college education, is also much more challenging to calibrate. We leave this strategy for the future and proceed with the OLG interpretation for now.

[19] Lucas (1990), for instance, uses this value.

[20] See İmrohoroğlu, İmrohoroğlu, and Joines (1999) for the latter.

a thirty year generation. Given these values for σ and τ_e, and an anticipated annual growth rate of 1.8%, the intergenerational discount parameter is set to 0.74 (0.99 in annual terms), so that the capital-to-output ratio given by (20) is close to 3 for annual GDP flows.

As our benchmark progressivity, we use $\Theta = 1.5$; for example, model tax rates of $\tau_s = 10\%$ and $\tau_c = 40\%$ on the poor will yield this level of progressivity. We will see later that this parametrization of progressivity is conservative, since the ratio of tax payments made by the rich to those made by the poor in equilibrium is lower than it is in the data; so we present results for $\Theta = 2.0$ (for example, model tax rates of $\tau_s = 10\%$ and $\tau_c = 50\%$), which is closer to the U.S. reality.

We do not restrict π_c and π_s to be the same as we did in the theoretical analysis. Analogous to the earlier parametrization, we use $\pi_c(e) = B_c e^{\gamma_c}$ and $\pi_s(e) = B_s e^{\gamma_s}$. Clearly, B_c, B_s, γ_c, and γ_s are the parameters that are most specific to our model. Of these four parameters, we normalize B_c to 1. To summarize, the following are the parameters we fix:

$$\begin{aligned} \text{Technology} &: \ \theta = 0.527,\ \nu = 0.35,\ \alpha = 0.36,\ \delta = 4.4\% \\ \text{Preference} &: \ \beta = 0.74,\ \sigma = 2, \\ \text{Education} &: \ B_c = 1 \\ \text{Others} &: \ \tau_e = 20\%. \end{aligned}$$

We then calibrate the remaining parameters, those in the human capital production functions and the ξ specification, so that the model outcomes of the key quantities mentioned above match observed values.[21] For the case of external growth, we used the specification $\xi(n_c) = C n_c^{\varepsilon}$. For the benchmark progressivity, our calibration yields: $\gamma_c = 0.1$, $B_s = 0.5$, $\gamma_s = 0.1$; $C = 1.82$, $\varepsilon = 0.9$. For the intentional adoption case, given that the fraction of skilled labor force involved in "R&D" is likely to be small, we use the specification $\xi(n_{cA}) = C + \varepsilon n_{cA}$, so that the constant term can pick up growth arising from other causes. While the education parameters are the same as those given above, the technology adoption parameters are: $C = 0.495$, and $\varepsilon = 3.36$.[22] The model outcomes are compared with their empirical counterparts in Table 1.[23]

[21] An alternate strategy involves the use of scant micro evidence to calibrate the function parameters directly, making the goal of matching other empirical outcomes a secondary one. Given the lack of direct connection between available empirics and our model, this process necessarily entails making heroic assumptions to go from regression coefficients to model parameters; therefore, we present results from this approach, based on empirical work by Kahn and Lim (1998), in footnotes.

[22] The annual growth rate, g, can be calculated from ξ, using the relation $(1+g)^{30} = 1 + \xi$.

[23] The estimates for the US were obtained from the following sources: fraction of college educated is the full-time equivalent figure for 1990 from Autor, Katz, and Krueger (1998), as is the range for the skill premium. The share devoted to education is for the early 90s from the Digest of Education Statistics (1997); the share of adult population with Master's, Professional, and Doctorate degrees is from the same source for the year 1996. The transition probabilities

Table 1

Comparison of Model Outcomes ($\Theta = 1.5$) *with U.S. Data*

Model Quantity	**Interpretation**	**U.S. Data**	**Externality**	**Adoption**
n_c	fraction of college educated	38.6%	35%	37.7%
g	per cap. annual growth rate	1.8%	1.8%	1.81%
$(n_c e_c + (1-n_c)e_s)\frac{w_c}{Y}$	share devoted to education	college: 2.9%; K-12: 7.3%	3.7%	6.90%
$b_s(e_s)^{\gamma_s}$	prob(rich|poor)	0.23 (17 year mobility)	0.24	0.25
$b_c(e_c)^{\gamma_c}$	prob(rich|rich)	0.65 (17 year mobility)	0.56	0.59
n_{cA}	fraction in "R&D"	6.8% PhD+Masters	N/A	6.4%
$n_{cA}\frac{w_c}{Y}$	share devoted to "R&D"	2.75%	N/A	6.16%
$p = \frac{w_c}{\Theta w_s}$	post tax skill premium	1.66-1.73	1.66	1.75
K/Y	capital-output ratio	2.5-3	2.88	2.88
$\tau_c n_c w_c / (\tau_c n_c w_c + \tau_s n_s w_s)$	% tax payments by rich	90.7%	70.4%	74.3%

The model does well in matching our targets, which are listed in the first five rows. It also produces related outcomes that are broadly consistent with the data; these are listed below the targeted outcomes.[24]

6 Policy Experiments

We now turn to quantifying the growth effects of moving from a progressivity of $\Theta = 1.5$, and from $\Theta = 2$, to a flat rate system.

6.1 Progressivity Changes – The Exempt Case

In the first experiment, we consider a change in progressivity when investments in human capital are tax exempt. Recall that this is the case where the actual level of taxes do not matter for long-run growth; only the ratio of retention rates matters. For this reason we do not worry about revenue neutrality while discussing the growth effects; taxing the rich and the poor at different rates in order

are from Gottschalk and Moffitt (1994). The share of tax payments by the rich is from the tax tables of the Statistical Abstract of the United States (1995); we calculate the percentage of taxes paid by roughly the top 38% to match up with the n_c value.

[24] While we compare the fraction of labor in "R&D" in the model with the fraction of the labor force in the data that has more than a bachelor's degree, we compare the model's GDP share of R&D with actual R&D expenses in the data. This explains the overstatement of the "R&D" expenses in the model; the wage share of this highly educated group might be a better benchmark.

to meet a given level of government expenditure would still affect growth only through the effect the change has on the retention ratio. However, revenue neutrality will matter for welfare comparisons.

6.1.1 Externality-driven Growth

Table 2 summarizes the effects of moving to a flat rate system when growth is driven by an externality.[25]

Table 2

Change in Progressivity – Externality

Variable	**Interpretation**	$\Theta = 1.0$	$\Theta = 1.5$	$\Theta = 2.0$
n_c	fraction of college educated	42.6%	35.2%	28.1%
g	growth rate (annualized)	2.06%	1.80%	1.54%
$(n_c e_c + (1 - n_c) e_s) \frac{w_c}{Y}$	share devoted to education	12.8%	3.7%	0.8%
$b_s(e_s)^{\gamma_s}$	prob(rich|poor)	0.27	0.24	0.20
$b_c(e_c)^{\gamma_c}$	prob(rich|rich)	0.63	0.56	0.48
$p = \frac{w_c}{\Theta w_s}$	post tax skill premium	1.35	1.66	2.05
K/Y	capital-output ratio	2.37	2.88	3.56
$\tau_c n_c w_c / (\tau_c n_c w_c + \tau_s n_s w_s)$	% tax payments by rich	50.1%	70.4%	81.6%

A move from a progressivity level of 1.5 to a flat rate system results in an increase of 0.26 percentage points in the long-run growth rate. As argued in the theoretical section, this is caused by an increase in the level of skill attainment, here from about 35% of the labor force to nearly 43%. This brings down the skill premium. In other words, a move to a less progressive system both increases growth and decreases inequality in the long run. The share of GDP devoted to human capital accumulation increases sharply. The increased level of investment translates to increased mobility for the poor (0.24 to 0.27) and greater persistence for the rich (0.56 to 0.63). The capital-output ratio will drop with an increased growth rate by raising the interest rate; here it goes from 2.88 in the benchmark to 2.37 under a flat rate system.

The increase in the growth rate and the drop in the premium is more dramatic when we move from the high level of progressivity of 2 to a flat rate system. Growth increases by 0.52 percentage points and the premium drops from 2.05 to 1.35. Education investment by the poor increases by a

[25] Here, and in Table 4, the actual tax rates matter for calculating the percent of taxes paid by the rich; we use the same rates that we later use in the welfare analysis.

higher factor than education investment by the rich.[26]

In assessing all these growth rate changes it is important to realize that they are brought about by the effect of progressivity on human capital investments alone. Potential channels for decreased progressivity to increase the labor supply and physical capital investment have been shut down. As a result, the above-mentioned increases seem quite significant, amounting to a 14.4% increase from the current growth rate when the transition is from $\Theta = 1.5$, and a 33.8% increase when the transition is from $\Theta = 2.0$.[27]

How does the transition from a progressive system to a flat-rate system look? Studying the transition paths will give us insights into the welfare analysis that follows. In the appendix we derive the system of difference equations that characterize the dynamics in terms of n_c, the share of skilled labor, z, the value of capital normalized by the technology level, and $\frac{\Lambda}{[(1-\tau_s)w_s]^{1-\sigma}}$, a normalized value to being skilled. Figure 3 shows the transition paths for selected variables from both levels of progressivity to a flat-rate system.

The endogenous state variables, n_c and z, do not jump at the time of the unexpected policy change. The share of skilled labor increases monotonically to its final level in response to a decrease in Θ, since the wedge in return to skill created by progressivity now vanishes and more people acquire skills. The increase in skilled labor is preferentially beneficial to technological change; therefore, the ratio of capital to technology decreases steadily to its final level. The (normalized) value to being skilled jumps at the time of the policy change. In fact, it overshoots its final value on account of the initial fixed nature of the stock of skilled labor. It takes time for the households to respond to the increase in value to being skilled by acquiring skills; as n_c increases over time, we see that the value to being skilled drops to its final level. This is reflected in the skill premium, p, which does not jump given that n_c does not jump, but declines over time as the fraction of skilled increases. The investment in education by both types jump at time zero, with a prominent overshooting observed

[26] Individual investments are not reported in the tables for brevity and are available from the authors; these numbers are mainly useful for understanding the mechanics of the model rather than in understanding macro level outcomes.

[27] Kahn and Lim (1998) regress industry level TFP growth on skilled labor's share of income, among other variables, which allows them to provide estimates of how much TFP growth would increase when that share increases. We used their regression coefficient of 0.1969 reported in Table V for the 1974-91 time period for the second calibration strategy mentioned in Section 5.

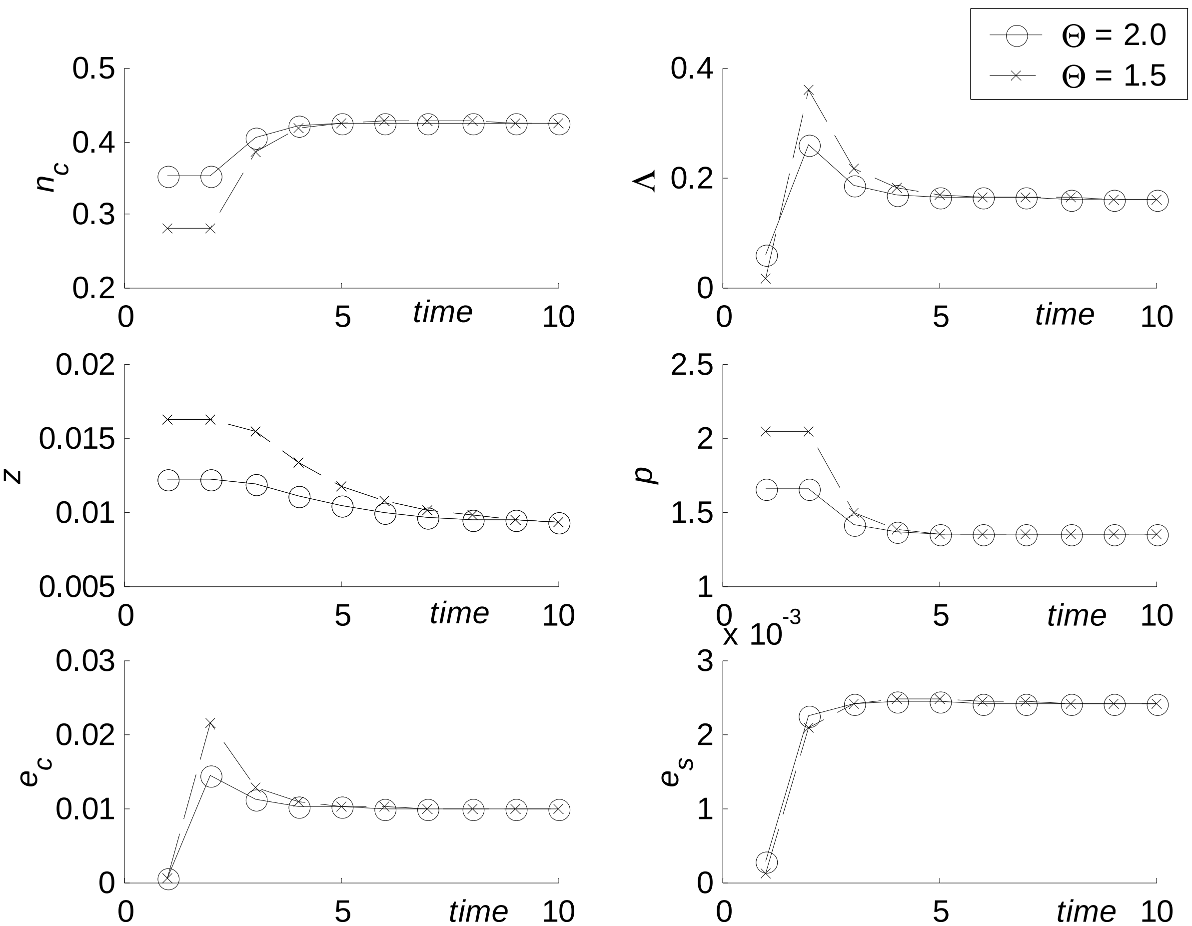
In our model the share corresponds to $\frac{w_c n_c}{Y}$. Under the simplifying (but false) assumption that $\frac{w_c}{Y}$ is a constant, independent of n_c, and estimating it from the data in Table 1, we can convert this regression coefficient to a corresponding one on n_c. We start with a model specification of $\xi = (B + Cn_c)^{\varepsilon}$, linearize it, and connect the coefficient on n_c to the value obtained above, to derive a restriction connecting the three parameters, B, C, and ε. The two free parameters are then calibrated to obtain outcomes as close as possible to the US data.

Doing this, we obtain the increase in growth rate to be 0.16 percentage points when going from a progressivity of 1.5 and an increase of 0.33 percentage points when going from a progressivity of 2.0.

for the “rich”. The interest rate increases over time to its final value consistent with a decreasing z.

Figure 3

Transition to Flat-rate Tax System – Externality



In Table 3, we present results on welfare gains from eliminating progressivity. Details on welfare computation are provided in the appendix. In order to conduct revenue neutral experiments, we set the tax rate in the flat rate regime is set to $\tau_c = \tau_s = 30\%$, which yields labor tax revenues 19.2% of the GDP, “reasonable” expectations for a flat-rate system.[28] The tax rates in the progressive regimes that yield the same tax revenues as a share of GDP are, $\tau_c = 44.6\%$, $\tau_s = 16.9\%$, when $\Theta = 1.5$, and $\tau_c = 55\%$, $\tau_s = 10\%$, when $\Theta = 2.0$.

[28] The actual tax rate chosen will not affect the *qualitative* implications of welfare discussed below.

Table 3

Welfare Gain from Eliminating Progressivity – Externality

Variable	$\Theta = 1.5$ **to** $\Theta = 1.0$	$\Theta = 2.0$ **to** $\Theta = 1.0$
BGP (Aggregate)	0.55%	1.31%
BGP utility of c	0.82%	1.68%
BGP utility of s	0.35%	1.04%
Including Transition (Aggregate)	-0.047%	-0.057%
Utility of c including transition	0.002%	0.061%
Utility of s including transition	-0.088%	-0.10%

The first row in Table 3 presents the equivalent yearly increase in consumption each agent would have to be given in a BGP indexed by $\Theta > 1$, in order to make an equally weighted aggregate welfare measure the same as that in a flat rate BGP. If the move to flat rate taxes is from a progressive system with parameter Θ of 1.5, each consumer has to be given more than 0.5% of her current consumption annually to make aggregate welfare the same across regimes; if the move is from a Θ of 2.0, each consumer has to be given more than 1.3%. As seen in Table 2, the growth rate increases when progressivity decreases, which increases aggregate welfare in the BGP of the flat rate regime and households in the progressive regimes have to be compensated to equate welfare.

The second and third rows attempt to address the following question. "Is it better to be rich (poor) in the BGP of the progressive regime or the flat rate regime?" From the equivalent variations given above, it is clear that the welfare gain for the rich is higher than that of the poor in going to a flat rate system. In addition to the increased growth rate in a flat rate system, the rich are being taxed less. The growth gains of the poor are partly, but not completely, negated by their increased taxes. The aggregate numbers lie in between the individual numbers. Thus, looking across balanced growth paths alone, it is clearly beneficial to everyone in the model economy to move to a flat rate system.

The last three rows, which provide the equivalent variation taking the transition into account, tell a different story. From the fourth row, we can see each consumer will be willing to pay to stay in the progressive regime in order to equate aggregate welfare across regimes. To understand this result, we look at the welfare of the rich and the poor at the time the regime change occurs given in the last two rows. The college-educated rich prefer the move to the flat rate system even when the transition is accounted for, but their gain is smaller than the one indicated by the BGP comparisons. While they are helped by the decreased tax rate, there are two other forces that work to decrease their consumption and temper the gain – first is their increased investment in education and a drop in the equilibrium skill premium; second is the discounting of the gain from increased long run growth.

For the school-educated poor, the decrease in consumption from the increased tax rate and increased investment early on seem to outweigh the later gains from a decreased skill premium, and they prefer the progressive system. Their willingness to pay to stay in the progressive system, and the fact there are nearly twice as many of them, account for the overall (slightly) negative effect on aggregate welfare.

Two points are worth noting in this regard. The liquidity constraint on households clearly contributes to the initial consumption loss they endure when investment increases. If in the real economy there are households that are not constrained, the above welfare measures will underestimate the gains in going to the flat rate regime. Second, it might be possible to design a scheme in which the government issues debt to finance the transition costs of households, especially the poor, and start repaying it when the growth gains are realized, thereby making the transition more palatable. Or one could phase out the progressive system gradually to soften the direct effect of increased taxes for the poor. The exploration of such schemes are topics for future research.

6.1.2 Adoption-driven Growth

Table 4 summarizes the effects of changes in progressivity in the adoption case.

Table 4

Change in Progressivity – Adoption

Variable	**Interpretation**	$\Theta = 1.0$	$\Theta = 1.5$	$\Theta = 2.0$
n_c	fraction of college educated	44.6%	37.7%	30.6%
g	growth rate (annualized)	1.93%	1.81%	1.66%
$(n_c e_c + (1 - n_c) e_s) \frac{w_c}{Y}$	share devoted to education	20.8%	6.9%	1.7%
$b_s(e_s)^{\gamma_s}$	prob(rich|poor)	0.28	0.25	0.22
$b_c(e_c)^{\gamma_c}$	prob(rich|rich)	0.65	0.59	0.51
n_{cA}	fraction in "R&D"	8.4%	6.4%	4.3%
$n_{cA} \frac{w_c}{Y}$	share devoted to "R&D"	7.2%	6.2%	4.7%
$p = \frac{w_c}{\Theta w_s}$	post tax skill premium	1.47	1.75	2.09
K/Y	capital-output ratio	2.61	2.88	3.22
$\tau_c n_c w_c / (\tau_c n_c w_c + \tau_s n_s w_s)$	% tax payments by rich	54.1%	74.3%	84.4%

The increase in the growth rate is more modest when we move from the benchmark progressivity to a flat rate system, amounting to 0.12 percentage points. Here growth does not increase automatically as more people become skilled; the adopting firm has to absorb a higher fraction of the labor force into R&D, and given the cost of doing this it treads lightly. And as discussed in the theoretical section, the increase in growth rate exerts an upward pressure on the premium and negatively affects the investment of the poor; this numbs the positive effect of an increase in the return to being skilled. Indeed, unlike the externality case, here we found that the investment of the rich increased by higher factors than did those for the poor.[29] The share to education is driven up nevertheless. The level of skill attainment increases, from 37.7% to 44.6% which brings the premium down from 1.75 to 1.47. The percentage of labor involved in adoption goes up (as argued in Lemma 1) from 6.4% to 8.4%. And, as in the externality case, the mobility of the poor increases.

The increase in the growth rate is higher when one goes to a flat rate system from a higher progressivity, amounting to 0.27 percentage points. Changes along other dimensions are also higher.[30]

As in the externality case, we examine the transition paths for selected variables in Figure 4 to give us insights into the welfare analysis that follows. As we argue in the appendix, there is an extra variable, n_{cA}, in the dynamic system. As in the externality case, the endogenous state variables, n_c and z, do not jump at the time of the unexpected policy change. The share of skilled labor increases to its final level in response to a decrease in progressivity and the ratio of capital to technology decreases to its final level after a slight initial increase. The (normalized) value to being skilled jumps at the time of the policy change and overshoots its final value, as earlier. The more interesting question here is why the share of labor devoted to "R&D" does not jump by more than it does. The intuition for this can be gleaned by examining (22). If n_c is unchanged initially, a steep increase in n_{cA} will cause the premium to shoot up; the firm will have to pay more for its skilled labor. As households respond to the decreased progressivity, become more educated and cause n_c to increase, the firm can afford to increase n_{cA}. In (22) as n_c increases, n_{cA} can increase without causing the premium to shoot up. Indeed, the time path for n_{cA} is remarkably similar to that for n_c. The fact that increasing A is now costly, and it is optimal for firms to increase it sluggishly means that K will adjust faster than A, causing z to initially increase slightly before it settles down to its final value. Corresponding to this, the interest rate initially drops and increases over time to its final value on the BGP. Given that increased growth exerts an upward pressure on the premium, the premium drops less dramatically

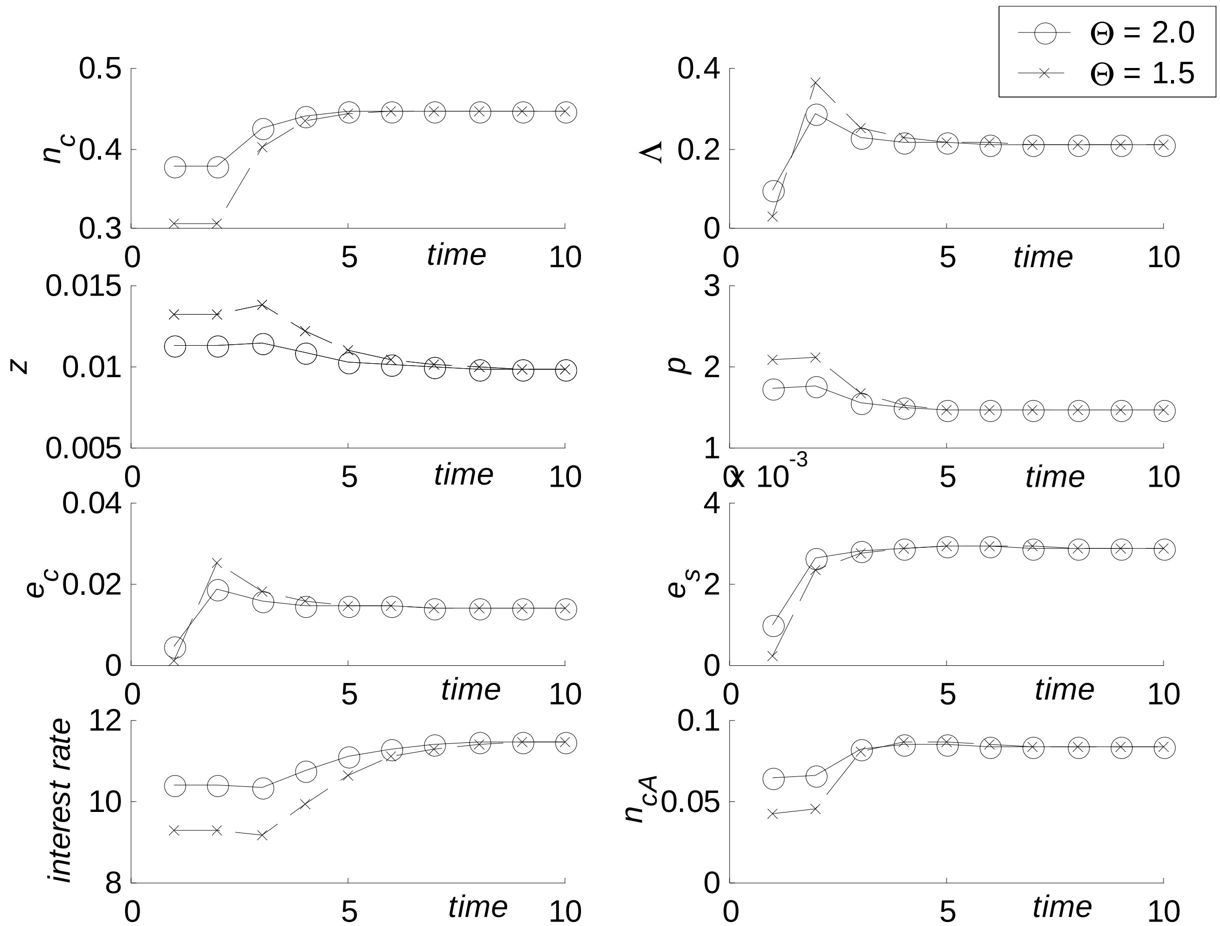
[29] If we had made the empirically implausible assumption that each type of agent needs only her own time to accumulate human capital, the growth effect would have only been higher. Therefore, our assumption that skilled time is required for human capital accumulation is a conservative one.

[30] We attempted the calibration using Kahn and Lim (1998), as outlined in a footnote in the previous subsection, here too. There is an extra step involved since the regression coefficient has to be converted to a corresponding one on n_{cA} instead of on n_c. We write $n_c = fn_{cA}$ and estimate f from data in Table 1; this again invloves a useful but heroic assumption that f is a constant. We use the model specification of $\xi = (B + Cn_{cA})^{\varepsilon}$, but otherwise proceed as in the externality case. Doing this, we obtain the increase in growth rate to be 0.07 percentage points when going from a progressivity of 1.5 and an increase of 0.13 percentage points when going from a progressivity of 2.0.

than it does in Figure 3.

Figure 4

Transition to Flat-rate Tax System – Adoption



We present the welfare estimates in Table 5. In these revenue neutral experiments, the tax rate in the flat rate regime is set to $\tau_c = \tau_s = 30\%$, which yields labor tax revenues that are 21.4% of GDP. The tax rates in the progressive regimes that yield the same tax revenues as a share of GDP are, $\tau_c = 44.1\%$, $\tau_s = 16.1\%$, when $\Theta = 1.5$ and $\tau_c = 54.7\%$, $\tau_s = 9.3\%$, when $\Theta = 2.0$.

Table 5

Welfare Gain from Eliminating Progressivity – Adoption

Variable	$\Theta = 1.5$ **to** $\Theta = 1.0$	$\Theta = 2.0$ **to** $\Theta = 1.0$
BGP (Aggregate)	0.38%	0.97%
BGP utility of c	0.67%	1.41%
BGP utility of s	0.15%	0.64%
Including Transition (Aggregate)	-0.052%	-0.053%
Utility of c including transition	0.026%	0.072%
Utility of s including transition	-0.094%	-0.105%

These numbers are very similar to those for the externality case. However, the gains are smaller reflecting the smaller increases in growth rates.

In summary, a decrease in progressivity increases growth, mobility for the poor, and decreases inequality as measured by the skill premium. The increase in growth rates is non-trivial, ranging from 0.13 to 0.53 percentage points depending on the experiment considered. The assumption made about the engine of growth matters, both qualitatively and quantitatively. There are significant welfare gains across BGPs, but transitional costs are large resulting in little change in aggregate welfare.

6.2 Progressivity Changes – The Non-Exempt Case

When investments in human capital are not tax exempt, the actual tax levels matter even for studying the long-run growth effects; it is no longer the case that progressivity affects allocations only through the ratio of the retention rates, Θ. This can be seen from the analogues of (23) and (24), the only conditions that will change for the non-exempt case:

$$\frac{\left[((1-\tau_c)-e_c^*)^{1-\sigma}(p^*)^{1-\sigma}-((1-\tau_s)-e_s^*p^*)^{1-\sigma}\right]}{(1-\sigma)\left[1-\beta(1+g)^{1-\sigma}(\pi_c(e_c^*)-\pi_s(e_s^*))\right]} = \frac{p^*((1-\tau_s)-e_s^*p^*)^{-\sigma}}{\beta(1+g)^{1-\sigma}\pi_s'(e_s^*)}, \tag{28}$$

$$\frac{\pi_c'(e_c^*)}{\pi_s'(e_s^*)} = \frac{((1-\tau_s)-e_s^*p^*)^{\sigma}}{((1-\tau_c)-e_c^*)^{\sigma}(p^*)^{\sigma}}. \tag{29}$$

The non-exempt case also allows us to examine the issue of whether changes in the level of flat rate taxes indeed matter less than changes in the progressivity of taxes, one of the motivations which started this paper. We consider flat rate tax change of 30% to 10%, as well as changes from a progressive system to a flat rate system. The tax rates are set in the progressive system so as to match the pre-specified progressivity parameter Θ, and raise the same tax revenue as a fraction of GDP as the flat rate system with 30% taxes does.

6.2.1 Externality-driven Growth

The results for growth driven by externality are given in Table 6. The tax rate in the flat rate regime is set to 30%, which yields labor tax revenues that are 19.2% of GDP.[31] The tax rates in the progressive regimes that yield the same tax revenues as a share of GDP are, $\tau_c = 44.7\%$, $\tau_s = 17\%$, which corresponds to $\Theta = 1.5$ and $\tau_c = 55\%$, $\tau_s = 10\%$, which corresponds to $\Theta = 2.0$.

Table 6

Change in Progressivity (non-exempt) – Externality

Variable	**Interpretation**	$\Theta = 1.0$	$\Theta = 1.0$	$\Theta = 1.5$	$\Theta = 2.0$
n_c	fraction of college educated	42.1%	40.9%	34.0%	27.6%
g	growth rate (annualized)	2.04%	2.01%	1.76%	1.52%
$(n_c e_c + (1-n_c)e_s)\frac{w_c}{Y}$	share devoted to education	11.79%	9.99%	2.83%	0.68%
$b_s(e_s)^{\gamma_s}$	prob(rich|poor)	0.27	0.27	0.24	0.21
$b_c(e_c)^{\gamma_c}$	prob(rich|rich)	0.63	0.62	0.53	0.44
$p = \frac{w_c}{\Theta w_s}$	post tax skill premium	1.37	1.41	1.71	2.09
K/Y	capital-output ratio	2.40	2.47	2.98	3.63
$\tau_c n_c w_c / (\tau_c n_c w_c + \tau_s n_s w_s)$	% tax payments by rich	49.8%	49.5%	69.9%	81.3%

The are two reasons for presenting this table. The first purpose is to reiterate the point that studying tax reform in a representative agent model with changes in flat rate taxes is likely to miss out important effects arising from changes in the structure of taxes.[32] The first two columns show the effect of change in a flat rate system. As we can see, there is little change in going from a 30% to a 10% system from the point of view of growth or other variables, in line with the findings of Lucas (1990) and Stokey and Rebelo (1995). There is a small positive effect due to increased income of the liquidity constrained agents but the effect of lower progressivity in eliminating the wedge in human capital return is absent. As can be seen from the last three columns and discussed below, the change in the degree of progressivity has more telling effects.

The second purpose is to check the robustness of the results presented in Table 2. When the baseline progressivity is 1.5 going to a flat rate system increases growth by 0.25 percentage points; when the progressivity is 2.0, the increase is 0.49 percentage points. The quantities, n_c, n_{cA}, education

[31] While we go from the 30% flat rate system to the 10% flat rate system, the taxes collected as a fraction of income drops; it is no longer 19.2%. One way to make this a revenue neutral change would be to increase the capital tax rate by the appropriate amount; in this model, such a change would be growth neutral.

[32] Note this experiment was not possible in the exempt case where flat rate changes had no long run growth effect.

and "R&D" share of GDP, mobility of the poor, and persistence of the rich all increase and the premium decreases. These results are similar to those in Table 2. Evidently, the strengthening of the liquidity constraint effect mentioned in Section 4 is countered by the fact that the effect of a decrease in progressivity on the value to being skilled is stronger when human capital investments are exempt from taxes.

For the sake of brevity, we do not report the welfare figures for this experiment, but they are very comparable to those in Table 3.

6.2.2 Adoption-driven Growth

For completeness, the results for growth driven by adoption in the non-exempt case are given in Table 7. The tax rate in the flat rate regime is again set to 30%, which yields labor tax revenues that are 21.3% of GDP. The tax rates in the progressive regimes that yield the same tax revenues as a share of GDP are, $\tau_c = 44.2\%$, $\tau_s = 16.3\%$, which corresponds to $\Theta = 1.5$ and $\tau_c = 54.8\%$, $\tau_s = 9.7\%$, which corresponds to $\Theta = 2.0$.

Table 7

Change in Progressivity (non-exempt) – Adoption

Variable	**Interpretation**	$\Theta = 1.0$	$\Theta = 1.0$	$\Theta = 1.5$	$\Theta = 2.0$
n_c	fraction of college educated	44.0%	42.7%	36.1%	29.6%
g	growth rate (annualized)	1.92%	1.90%	1.78%	1.64%
$(n_c e_c + (1 - n_c) e_s) \frac{w_c}{Y}$	share devoted to education	19.2%	15.9%	5.0%	1.2%
$b_s(e_s)^{\gamma_s}$	prob(rich|poor)	0.28	0.27	0.25	0.22
$b_c(e_c)^{\gamma_c}$	prob(rich|rich)	0.65	0.64	0.56	0.46
n_{cA}	fraction in "R&D"	8.2%	7.8%	6.0%	4.0%
$n_{cA} \frac{w_c}{Y}$	share devoted to "R&D"	7.2%	7.0%	5.9%	4.4%
$p = \frac{w_c}{\Theta w_s}$	post tax skill premium	1.49	1.54	1.82	2.15
K/Y	capital-output ratio	2.63	2.68	2.94	3.28
$\tau_c n_c w_c / (\tau_c n_c w_c + \tau_s n_s w_s)$	% tax payments by rich	53.9%	53.4%	73.6%	83.7%

The points discussed under Table 6 are applicable here too – reform in the tax progressivity has more of an effect than a decrease in flat rate taxes, and the results are very similar to the exempt case.

7 Conclusions

The heterogeneous-agent, endogenous growth framework developed in this paper is a step toward analyzing the structural linkages between growth and inequality. We find it interesting that a less progressive tax system, which is rarely perceived as an egalitarian measure, gives rise to increased growth, decreased inequality, and greater mobility for the poor in the long run, especially in light of contradicting claims in the literature regarding the connection between growth and inequality.[33]

The finding that the progressivity of taxes has a non-neutral (negative) effect on growth seems theoretically robust; experiments on a calibrated model indicate the quantitative effects are economically significant, ranging from 0.13 to 0.53 percentage points. Reform in the *structure* of taxes has more of an effect than reform in the *level* of taxes alone. We also find that the assumption made about the engine of growth matters, both qualitatively and quantitatively.

When progressivity decreases, welfare unequivocally increases across balanced growth paths; however, once transition is taken into account, only the currently rich slightly prefer the flat-rate system. While the long-run welfare gains of moving to a flat rate system is high, so are the costs of transition, resulting in little change in aggregate welfare; the effect, if anything, is slightly negative. The exploration of debt-based schemes and gradual phase out of progressivity to soften the blow for the initially poor and make them "buy into" the flat rate system seem useful avenues for future research.

A Appendix

A.1 Lemma 1 ($n_{cA}^*(n_c^*)$ is strictly increasing for the adoption firm)

The left hand side of (21), which is the analogue of $\rho + \sigma g$ in continuous time growth models, is increasing in n_{cA}^* for any increasing ξ function and any $\sigma > 0$. It is a marginal cost term that captures the utility costs of lost production by diverting labor to technology adoption on the BGP. The right hand side captures the marginal benefit on the BGP of increasing labor in adoption; the first term reflects an increase in current output arising from an increased productivity level and the second term reflects an increase in the level of future productivity improvements given the law of motion (15). The first term decreases with n_{cA}^* for any concave adoption function ξ; an increase in R&D labor is less effective in increasing A on the margin, and is further compounded by a diminishing effect on output due to a reduction in the skilled labor force available for production ($\frac{Y}{w_c}$ is decreasing). The second term is clearly increasing in n_{cA}^*. Therefore, to determine the sign of the right hand side, we parametrize the adoption function to be $\xi(n_{cA}) = Cn_{cA}^{\varepsilon}$, $0 < \varepsilon < 1$. Some algebra

[33] For instance, in a recent study, Forbes (2000) concludes that a reassessment of the linkages between inequality, growth, and their determinants is warranted. Our study points to the progressivity of taxes as one such structural linkage.

now shows that the right side hand decreases unambiguously (noting $\nu < 1$). In summary, the left hand side when plotted against n_{cA}^* is increasing, starting from $(1+\rho)$. The right hand side is decreasing from ∞. And it is possible to find parameters for which a unique intersection exists.

For any given n_{cA}^*, the only term in (21) that is affected by n_c^* is the $\frac{Y}{w_c}$ term. This term increases in n_c^*, provided the p^* in (22) is greater than one. We expect the post-tax premium on a BGP equilibrium, $\frac{p^*}{\Theta}$, to be greater than one; otherwise there will be no incentive for anyone to accumulate skills. For a given level of adoption activity, higher the skilled labor force, higher the output will be (though increasing at a diminishing rate) and lower will the skilled wages be, both of which increase $\frac{Y}{w_c}$ and the production benefit of increasing adoption. In terms of the graphical analysis outlined in the preceding paragraph, the right side shifts up and the intersection point, n_{cA}^*, increases ■ .

A.2 Lemma 2 (Ignoring equilibrium effects for both types of growth)

Consider the parametrization $\pi_c = \pi_s = Be^{\gamma}$, $0 < \gamma < 1$, to characterize the household sector on the BGP. For this parametrization, (23) becomes:

$$\left(\frac{e_s^*}{e_c^*}\right)^{1-\gamma} = \frac{\Theta^{\sigma}}{\Theta}\left(\frac{\frac{1}{p^*} - e_s^*}{1 - e_c^*}\right)^{\sigma}. \tag{30}$$

Eliminate $\Theta^{1-\sigma}$ from (30) in (24), use the parametrization chosen for π, and simplify to get:

$$\frac{\left[\frac{1-e_c^*}{(e_c^*)^{1-\gamma}} - \frac{\frac{1}{p^*}-e_s^*}{(e_s^*)^{1-\gamma}}\right]}{1-\sigma} = \frac{1}{\gamma B}\left[\frac{1}{\beta(1+g)^{1-\sigma}} - B\left((e_c^*)^{\gamma} - (e_s^*)^{\gamma}\right)\right]. \tag{31}$$

Take logs and differentiate (30) with respect to Θ, and simplify to get:

$$\left[\frac{1-\gamma}{e_s} + \frac{\sigma}{\frac{1}{p} - e_s}\right]\frac{\partial e_s}{\partial \Theta} - \left[\frac{1-\gamma}{e_c} + \frac{\sigma}{1-e_c}\right]\frac{\partial e_c}{\partial \Theta} = -\frac{(1-\sigma)}{\Theta} - \frac{\frac{\sigma}{p^2}}{\frac{1}{p} - e_s}\frac{\partial p}{\partial \Theta}, \tag{32}$$

where we have dropped asterisks from all variables for simplicity of notation. Differentiate (31) with respect to Θ, and wade through some algebra to get:

$$\left[\sigma + (1-\gamma)\frac{\left(\frac{1}{p} - e_s\right)}{e_s}\right]\frac{\partial e_s}{\partial \Theta} - \left(\frac{e_s}{e_c}\right)^{1-\gamma}\left[\sigma + (1-\gamma)\frac{(1-e_c)}{e_c}\right]\frac{\partial e_c}{\partial \Theta} = -\frac{1}{p^2}\frac{\partial p}{\partial \Theta}. \tag{33}$$

When we ignore the effects on the skill premium, $\frac{\partial p}{\partial \Theta} = 0$. In this case, we can see from (33) $\frac{\partial e_c}{\partial \Theta}$ and $\frac{\partial e_s}{\partial \Theta}$ have the same signs. This is important in understanding the final result; the intertemporal effect of an increase in Θ is same for both agents' investments. And given that we expect $e_c > e_s$, the above equation indicates the effect is stronger for e_c. Substitute for $\frac{\partial e_s}{\partial \Theta}$ from (33) into (32), using (30), and wading through more algebra, one can get:

$$\left\{\frac{\left[\frac{1-e_c}{\Theta}\right]^{1-\sigma} - \left(\frac{1}{p} - e_s\right)^{1-\sigma}}{1-\sigma}\right\}\frac{1}{\left(\frac{1}{p} - e_s\right)^{1-\sigma}}\left[\frac{1-\gamma}{e_c} + \frac{\sigma}{1-e_c}\right]\frac{\partial e_c}{\partial \Theta} = -\frac{1}{\Theta}. \tag{34}$$

The term within curly braces is an excess utility term which needs to be positive in any BGP equilibrium with positive investment; otherwise there will be no incentive for agents to become skilled (Λ will become negative). This means $\frac{\partial e_c}{\partial \Theta} < 0$ and as discussed above $\frac{\partial e_s}{\partial \Theta} < 0$. From (25) it follows that n_c decreases with Θ.

Even though an increase in the progressivity could shift the investment in favor of the unskilled through the liquidity effect, it is the intertemporal effect that ultimately dominates and decreases the investments of both types of agents. ■

A.3 Lemma 3 (With equilibrium effects for external growth)

For the external growth case, we can take logs and differentiate (26) to get:

$$\frac{1}{p}\frac{\partial p}{\partial \Theta} = -\frac{\gamma B(1-\nu)}{e_c^{1-\gamma}(1-Be_c^{\gamma})}\frac{\partial e_c}{\partial \Theta} - \frac{\gamma(1-\nu)}{e_s}\frac{\partial e_s}{\partial \Theta}.$$

Using this in (33) we get:

$$\left\{\sigma + (1-\gamma)\frac{\left(\frac{1}{p}-e_s\right)}{e_s} - \frac{\gamma(1-\nu)}{pe_s}\right\}\frac{\partial e_s}{\partial \Theta} = \left\{\left(\frac{e_s}{e_c}\right)^{1-\gamma}\left[\sigma + (1-\gamma)\frac{(1-e_c)}{e_c}\right] + \frac{\gamma B(1-\nu)}{pe_c^{1-\gamma}(1-Be_c^{\gamma})}\right\}\frac{\partial e_c}{\partial \Theta}. \tag{35}$$

The result is that $\frac{\partial e_s}{\partial \Theta}$ is made stronger relative to $\frac{\partial e_c}{\partial \Theta}$, than when the effect of tuition increase is ignored as in Lemma 2. Doing the same for (32), using (35) and (30), and after some messy algebra we can get:

$$\Bigg[\frac{1}{\left(\frac{1}{p}-e_s\right)^{1-\sigma}}\left[\sigma + (1-\gamma)\frac{(1-e_c)}{e_c}\right]\left\{\frac{\Psi\left[\frac{1-e_c}{\Theta}\right]^{1-\sigma} - \left(\frac{1}{p}-e_s\right)^{1-\sigma}}{1-\sigma}\right\} +$$
$$\left[\frac{\gamma B(1-\nu)}{\left(\frac{1}{p}-e_s\right)pe_c^{1-\gamma}(1-Be_c^{\gamma})}\right]\left\{\frac{\sigma + (1-\gamma)\frac{\left(\frac{1}{p}-e_s\right)}{e_s}}{\sigma + (1-\gamma)\frac{\left(\frac{1}{p}-e_s\right)}{e_s} - \frac{\gamma(1-\nu)}{pe_s}}\right\}\Bigg]\cdot\frac{\partial e_c}{\partial \Theta} = -\frac{1}{\Theta},$$

where,

$$\Psi \equiv \frac{\left[\sigma + (1-\gamma)\frac{\left(\frac{1}{p}-e_s\right)}{e_s} - \frac{\sigma\gamma(1-\nu)}{pe_s}\right]}{\left[\sigma + (1-\gamma)\frac{\left(\frac{1}{p}-e_s\right)}{e_s} - \frac{\gamma(1-\nu)}{pe_s}\right]}.$$

When compared to (34), the factor Ψ within the first set of curly braces and the second term are new. It is straightforward to show that the first term within curly braces is positive no matter what σ is, as long as $\frac{1-e_c}{\Theta} > \frac{1}{p} - e_s$, which as we argued in Lemma 1 needs to be the case on the BGP for Λ to be positive. So, we can unambiguously sign $\frac{\partial e_c}{\partial \Theta}$ provided the numerator and denominator of Ψ are positive. The sufficient conditions that guarantee this are:

$$(1-\sigma) < \gamma < \min\left\{\frac{1}{2-\nu}, \frac{1}{1+\sigma(1-\nu)}\right\}.$$

The second part of the inequality is automatically satisfied if $\nu = 1$. In other words, as alluded to in the main text, we want high enough elasticity of substitution in production, sufficient diminishing returns in human capital accumulation, and high enough σ so that elasticities of investment to premium changes are low and general equilibrium effect does not overturn the result in Lemma 1. These seem relatively mild conditions. For instance, $\gamma = \nu = 1/2$, will work for $\sigma = 1/2$, and $\nu = 1/2$, $\gamma = 1/4$ will work for $\sigma = 2$. The same sufficient conditions guarantee the sign of $\frac{\partial e_s}{\partial \Theta}$ is the same as $\frac{\partial e_c}{\partial \Theta}$. ■

A.4 Lemma 4 (With equilibrium effects for adoption-driven growth)

For the external growth case, we can take logs and differentiate (27) to get:

$$\frac{1}{p}\frac{\partial p}{\partial \Theta} = -\frac{\gamma B(1-\nu)}{e_c^{1-\gamma}\left(1 - Be_c^{\gamma}\right)}\frac{\partial e_c}{\partial \Theta} - \frac{\gamma B(1-\nu)}{e_s^{1-\gamma}\left(Be_s^{\gamma} - n_{cA}\right)}\frac{\partial e_s}{\partial \Theta}.$$

The second term is different from the analogous expression in Lemma 3. One can derive an expression analogous to (35) in the previous lemma:

$$\left\{\sigma + (1-\gamma)\frac{\left(\frac{1}{p} - e_s\right)}{e_s} - \frac{\gamma B(1-\nu)}{pe_s^{1-\gamma}\left(Be_s^{\gamma} - n_{cA}\right)}\right\}\frac{\partial e_s}{\partial \Theta} = \left\{\left(\frac{e_s}{e_c}\right)^{1-\gamma}\left[\sigma + (1-\gamma)\frac{(1-e_c)}{e_c}\right] + \frac{\gamma B(1-\nu)}{pe_c^{1-\gamma}\left(1 - Be_c^{\gamma}\right)}\right\}\frac{\partial e_c}{\partial \Theta}.$$

Again, only the second term within the curly braces in the left side is different. However the earlier sufficient conditions to guarantee e_s and e_c move in the same direction will not be enough here. In particular, for high enough anticipated growth (high enough n_{cA}), the second term can become negative enough to cause the two types of investment to move in the opposite direction. Proceeding as we did in the previous lemma, we can get:

$$[\frac{1}{\left(\frac{1}{p} - e_s\right)^{1-\sigma}}\left[\sigma + (1-\gamma)\frac{(1-e_c)}{e_c}\right]\left\{\frac{\widehat{\Psi}\left[\frac{1-e_c}{\Theta}\right]^{1-\sigma} - \left(\frac{1}{p} - e_s\right)^{1-\sigma}}{1-\sigma}\right\} +$$
$$\left[\frac{\gamma B(1-\nu)}{\left(\frac{1}{p} - e_s\right)pe_c^{1-\gamma}\left(1 - Be_c^{\gamma}\right)}\right]\left\{\frac{\sigma + (1-\gamma)\frac{\left(\frac{1}{p}-e_s\right)}{e_s}}{\sigma + (1-\gamma)\frac{\left(\frac{1}{p}-e_s\right)}{e_s} - \frac{\gamma B(1-\nu)}{pe_s^{1-\gamma}(Be_s^{\gamma}-n_{cA})}}\right\}]\cdot\frac{\partial e_c}{\partial \Theta} = -\frac{1}{\Theta},$$

where,

$$\widehat{\Psi} \equiv \frac{\left[\sigma + (1-\gamma)\frac{\left(\frac{1}{p}-e_s\right)}{e_s}\right] - \frac{\sigma\gamma B(1-\nu)}{pe_s^{1-\gamma}(Be_s^{\gamma}-n_{cA})}}{\left[\sigma + (1-\gamma)\frac{\left(\frac{1}{p}-e_s\right)}{e_s}\right] - \frac{\gamma B(1-\nu)}{pe_s^{1-\gamma}(Be_s^{\gamma}-n_{cA})}}.$$

When $n_{cA} \longrightarrow 0$, these expressions reduce to the ones in the previous lemma and the same results hold. The numerator and denominator of $\widehat{\Psi}$ can be negative for large enough growth, leaving $\widehat{\Psi}$ positive. But the more important point is when the term $\frac{\gamma B(1-\nu)}{pe_s^{1-\gamma}(Be_s^{\gamma}-n_{cA})}$ is large enough to make the denominator positive, it makes $\frac{\partial e_c}{\partial \Theta} > 0$. This *same* force makes $\frac{\partial e_s}{\partial \Theta}$ and $\frac{\partial e_c}{\partial \Theta}$ moves in opposite directions and causes $\frac{\partial e_s}{\partial \Theta} < 0$. In other words, while e_c can increase with progressivity for large enough growth rates, e_s always decreases with progressivity.

■

A.5 Transitions

We seek a system of difference equations in the following variables when growth is driven by externality: n_{ct}, the share of skilled labor in the economy, $\frac{\Lambda_{t+1}}{[(1-\tau_s)w_{st}]^{1-\sigma}}$, a normalized version of the value to being skilled, and $z_t \equiv \frac{K_t}{A_t}$, the capital stock normalized by the productivity level. These normalizations help in analyzing the system in terms of variables that settle down to stationary values on the BGP; on the BGP, Λ grows at the gross rate $(1+g)^{1-\sigma}$, w_s, K, and A grow at the gross rate $(1+g)$, so the last two variables will indeed settle down.

The steady state values of the normalized quantities in terms of the four BGP quantities determined in the main text, p^*, e_s^*, g, and n_c^*, are:

$$\begin{aligned}\left(\frac{\Lambda}{[(1-\tau_s)w_s]^{1-\sigma}}\right)^* &= \frac{p^*}{\beta\gamma_s B_s}\frac{(e_s^*)^{1-\gamma_s}}{(1-p^*e_s^*)^{\sigma}},\\ z^* &= \left[\theta\left(n_c^*\right)^{\nu}+(1-\theta)\left(1-n_c^*\right)^{\nu}\right]^{\frac{1}{\nu}}\left(\frac{K}{Y}\right)_*^{\frac{1}{1-\alpha}},\end{aligned}$$

where the capital-to-output ratio is given on the BGP as before: $\left(\frac{K}{Y}\right)_* = \frac{\alpha(1-\tau_e)}{(1+\rho)(1+g)^{\sigma}-(1-\delta)}$. Note we use the parametrization $\pi_c(e) = B_c e^{\gamma_c}$; $\pi_s(e) = B_s e^{\gamma_s}$

The law of motion for n_c is:

$$n_{ct+1} = n_{ct}B_c e_{ct}^{\gamma_c} + (1-n_{ct})B_s e_{st}^{\gamma_s}. \tag{36}$$

The difference equation for the value to being skilled can be derived similar to the equations in the main text as:

$$\begin{aligned}\frac{\Lambda_t}{[(1-\tau_s)w_{st-1}]^{1-\sigma}}\left(\frac{w_{st-1}}{w_{st}}\right)^{1-\sigma} &= \frac{(1-e_{ct})^{1-\sigma}\left(\frac{p_t}{\Theta}\right)^{1-\sigma}-(1-p_t e_{st})^{1-\sigma}}{1-\sigma}\\ &\quad+\beta\left(B_c e_{ct}^{\gamma_c}-B_s e_{st}^{\gamma_s}\right)\frac{\Lambda_{t+1}}{[(1-\tau_s)w_{st}]^{1-\sigma}}.\end{aligned} \tag{37}$$

To get the wage growth term in this equation in terms of the system variables, one needs to take derivatives of the production function, use the definition of z, and use the law of motion for A to get:

$$\frac{w_{st-1}}{w_{st}} = \frac{1}{1+\xi(n_{ct-1})}\left(\frac{z_{t-1}}{z_t}\right)^{\alpha}\left[\frac{\theta(n_{ct-1})^{\nu}+(1-\theta)(1-n_{ct-1})^{\nu}}{\theta(n_{ct})^{\nu}+(1-\theta)(1-n_{ct})^{\nu}}\right]^{\frac{1-\alpha}{\nu}-1}\left(\frac{1-n_{ct}}{1-n_{ct-1}}\right)^{1-\nu}.$$

To get the difference equation for z we start with the entrepreneur's Euler equation:

$$\left(\frac{c_{et+1}}{c_{et}}\right)^{\sigma} = \beta\left[1+(1-\tau_e)\alpha\frac{Y_{t+1}}{K_{t+1}}-\delta\right],$$

where we can write $c_{et} = (1-\tau_e)\alpha Y_t - K_{t+1} + (1-\delta)K_t$. Divide throughout by A_t, use the definition of z to get:

$$\frac{c_{et}}{A_t} = (1-\tau_e)\alpha\frac{Y_t}{A_t} - z_{t+1}(1+\xi(n_{ct})) + (1-\delta)z_t.$$

One can therefore rewrite the above Euler equation and get the equation for z as:

$$\frac{(1-\tau_e)\alpha\frac{Y_{t+1}}{K_{t+1}}z_{t+1} - z_{t+2}(1+\xi(n_{ct+1})) + (1-\delta)z_{t+1}}{(1-\tau_e)\alpha\frac{Y_t}{K_t}z_t - z_{t+1}(1+\xi(n_{ct})) + (1-\delta)z_t}(1+\xi(n_{ct})) = \beta^{\frac{1}{\sigma}}\left[1+(1-\tau_e)\alpha\frac{Y_{t+1}}{K_{t+1}}-\delta\right]^{\frac{1}{\sigma}}. \tag{38}$$

In terms of the system variables, the Y/K variable used above is given as:

$$\frac{Y_t}{K_t} = \frac{\left[\theta\left(n_{ct}\right)^{\nu} + (1-\theta)\left(1-n_{ct}\right)^{\nu}\right]^{\frac{1-\alpha}{\nu}}}{z_t^{1-\alpha}}.$$

Equations (36) through (38) completely characterize the dynamics of the system when growth is driven by externality.

The transition path we are interested in will be from the BGP of one tax regime to the BGP of the new tax regime. The starting and ending BGP values can be computed in a straightforward way, given those the above difference equations can be used to compute the transition paths using a standard relaxation method.

When growth is driven by intentional adoption, one has an additional state variable in $n_{cA,t}$. This gives rise to a fourth difference equation; this is the main change when compared to the earlier transitions. The steady state quantities are n_{cA}^*, n_c^*, and:

$$\begin{aligned}
\left(\frac{\Lambda}{\left[(1-\tau_s) w_s\right]^{1-\sigma}}\right)^* &= \frac{p^*}{\beta\gamma_s B_s}\frac{\left(e_s^*\right)^{1-\gamma_s}}{\left(1-p^* e_s^*\right)^{\sigma}}, \\
z^* &= \left[\theta\left(n_c^* - n_{cA}^*\right)^{\nu} + (1-\theta)\left(1-n_c^*\right)^{\nu}\right]^{\frac{1}{\nu}}\left(\frac{K}{Y}\right)_*^{\frac{1}{1-\alpha}},
\end{aligned}$$

where the capital-to-output ratio is given on the BGP as before: $\left(\frac{K}{Y}\right)_* = \frac{\alpha(1-\tau_e)}{(1+\rho)(1+g)^{\sigma}-(1-\delta)}$.To derive the difference equation for $n_{cA,t}$, use the equations in (17) to get:

$$\beta\left[(1-\alpha)\frac{Y_t}{w_{ct}}\xi'\left(n_{cA,t}\right) + \left(1+\xi\left(n_{cA,t}\right)\right)\right]\frac{w_{ct}u'\left(c_{et}\right)}{A_t\xi'\left(n_{cA,t}\right)} = \frac{w_{ct-1}u'\left(c_{et-1}\right)}{A_{t-1}\xi'\left(n_{cA,t-1}\right)}.$$

Use the entrepreneur's Euler equation to eliminate consumption and get:

$$\left(\frac{w_{ct}}{w_{ct-1}}\right)\left(\frac{A_{t-1}}{A_t}\right)\frac{\xi'\left(n_{cA,t-1}\right)}{\xi'\left(n_{cA,t}\right)}\left[(1-\alpha)\frac{Y_t}{w_{ct}}\xi'\left(n_{cA,t}\right) + \left(1+\xi\left(n_{cA,t}\right)\right)\right] = 1 + (1-\tau_e)\alpha\frac{Y_t}{K_t} - \delta.$$

Using the law of motion for A and simplifying, we get the difference equation in n_{cA} as:

$$\left(\frac{w_{ct}}{w_{ct-1}}\right)\frac{\xi'\left(n_{cA,t-1}\right)}{1+\xi\left(n_{cA,t-1}\right)}\left[(1-\alpha)\frac{Y_t}{w_{ct}} + \frac{\left(1+\xi\left(n_{cA,t}\right)\right)}{\xi'\left(n_{cA,t}\right)}\right] = 1 + (1-\tau_e)\alpha\frac{Y_t}{K_t} - \delta. \tag{39}$$

Note that this will reduce to condition (21) in the paper on the BGP. In the above expression:

$$\begin{aligned}
\frac{Y_t}{w_{ct}} &= \frac{\left(n_{ct} - n_{cA,t}\right)^{1-\nu}\left[\theta\left(n_{ct} - n_{cA,t}\right)^{\nu} + (1-\theta)\left(1-n_{ct}\right)^{\nu}\right]}{\theta(1-\alpha)} \\
\frac{w_{ct}}{w_{ct-1}} &= \left(1+\xi\left(n_{cA,t-1}\right)\right)\left(\frac{z_t}{z_{t-1}}\right)^{\alpha}\left[\frac{\theta\left(n_{ct} - n_{cA,t}\right)^{\nu} + (1-\theta)\left(1-n_{ct}\right)^{\nu}}{\theta\left(n_{ct-1} - n_{cA,t-1}\right)^{\nu} + (1-\theta)\left(1-n_{ct-1}\right)^{\nu}}\right]^{\frac{1-\alpha}{\nu}-1}\left(\frac{n_{ct-1} - n_{cA,t-1}}{n_{ct} - n_{cA,t}}\right)^{1-\nu}.
\end{aligned}$$

Most of the household equations do not change. The only other change of any significance is now:

$$c_{et} = (1-\tau_e)\left(\alpha Y_t - n_{cA,t}w_{ct}\right) - K_{t+1} + (1-\delta)K_t.$$

This will be reflected in the difference equation in z which is:

$$\frac{(1-\tau_e)\frac{Y_{t+1}}{K_{t+1}}z_{t+1}\left(\alpha - n_{cA,t+1}\frac{w_{ct+1}}{Y_{t+1}}\right) - z_{t+2}\left(1+\xi\left(n_{cA,t+1}\right)\right) + (1-\delta)z_{t+1}}{(1-\tau_e)\frac{Y_t}{K_t}z_t\left(\alpha - n_{cA,t}\frac{w_{ct}}{Y_t}\right) - z_{t+1}\left(1+\xi\left(n_{cA,t}\right)\right) + (1-\delta)z_t}\left(1+\xi\left(n_{cA,t}\right)\right) = \beta^{\frac{1}{\sigma}}\left[1 + (1-\tau_e)\alpha\frac{Y_{t+1}}{K_{t+1}} - \delta\right] \tag{40}$$

The inverse of $\frac{w_c}{Y}$ was given earlier, and $\frac{Y}{K}$ is given as:

$$\frac{Y_t}{K_t}=\frac{\left[\theta\left(n_{ct}-n_{cA,t}\right)^{\nu}+\left(1-\theta\right)\left(1-n_{ct}\right)^{\nu}\right]^{\frac{1-\alpha}{\nu}}}{z_t^{1-\alpha}}.$$

Equations (36), (37), (39), and (40) are the four difference equations that govern the dynamics of the economy when growth is driven by technology adoption.

The equations governing the dynamics when education investments are non-exempt are identical except for (37), which will now be derived from (28) and (29).

A.6 Welfare Measures

We can compute two types of welfare measures to evaluate progressivity policies – one an equally weighted aggregate measure, which will help us see in an overall sense how a policy fares, and individual welfare measures for those initially skilled or unskilled to study the distributional implications. We compute these measures only for the households and not the entrepreneur – there is no heterogeneity in the latter, making the welfare computation of the entrepreneur a bit uninteresting. In all cases we will assume that the BGP and transition quantities have been obtained.

A.6.1 Aggregate welfare

This measure is fairly straightforward to compute. First denote the individual consumptions as:

$$\begin{aligned}c_{ct} &= \left(1-\tau_c\right)\left(1-e_{ct}\right)w_{ct}\\ c_{st} &= \left(1-\tau_s\right)\left(\frac{1}{p_t}-e_{st}\right)w_{ct}.\end{aligned}$$

At time t aggregate welfare is $W_t=n_{ct}\frac{c_{ct}^{1-\sigma}}{1-\sigma}+\left(1-n_{ct}\right)\frac{c_{st}^{1-\sigma}}{1-\sigma}$, which amounts to:

$$W_t=\left\{n_{ct}\left[\left(1-\tau_c\right)\left(1-e_{ct}\right)\right]^{1-\sigma}+\left(1-n_{ct}\right)\left[\left(1-\tau_s\right)\left(\frac{1}{p_t}-e_{st}\right)\right]^{1-\sigma}\right\}\frac{w_{ct}^{1-\sigma}}{1-\sigma}.$$

The aggregate welfare on a BGP is given by:

$$\begin{aligned}W^{BGP,0} &= \sum_{t=0}^{\infty}\beta^t\left(1+g\right)^{(1-\sigma)t}\left\{n_c^*\left[\left(1-\tau_c\right)\left(1-e_c^*\right)\right]^{1-\sigma}+\left(1-n_c^*\right)\left[\left(1-\tau_s\right)\left(\frac{1}{p^*}-e_s^*\right)\right]^{1-\sigma}\right\}\frac{w_{c0}^{1-\sigma}}{1-\sigma}\\ &= \frac{\left\{n_c^*\left[\left(1-\tau_c\right)\left(1-e_c^*\right)\right]^{1-\sigma}+\left(1-n_c^*\right)\left[\left(1-\tau_s\right)\left(\frac{1}{p^*}-e_s^*\right)\right]^{1-\sigma}\right\}}{\left[1-\beta\left(1+g\right)^{(1-\sigma)}\right]}\cdot\frac{w_{c0}^{1-\sigma}}{1-\sigma}.\end{aligned}\tag{41}$$

This welfare can be computed only relative to a starting wage; hence the superscript zero in $W^{BGP,0}$ to indicate w_{c0} is being used. If we use $W^{BGP,T}$ it means w_{cT} is being used. To compare two regimes the same starting wage needs to be used; we normalize w_{c0} to one.

Suppose transition from one BGP to another is almost completed at time T. Then the aggregate welfare including transition is given by:

$$W^{new}=\sum_{t=0}^{T-1}\beta^t W_t+\beta^T W^{BGP,T}.$$

This assumes we know $w_{c0} \cdots w_{cT}$, which is where the normalization $w_{c0} = 1$ is helpful. Use that and use expressions for $\frac{w_{ct}}{w_{ct-1}}$ from the earlier section, to get the sequence of wages, including the w_{cT} needed to get $W^{BGP,T}$.

The aggregate welfare on the BGP of the pre-policy change regime, W^{old}, can be computed using (41); all the starred quantities will correspond to the old BGP and $w_{c0} = 1$. The consumption equivalence ω can be obtained from:

$$(1+\omega)^{1-\sigma} W^{old} = W^{new}.$$

A.6.2 Individual welfare

To first get individual welfare on the BGP at an arbitrary time zero, start with the maximized Bellman equations:

$$\begin{aligned} V_c &= u(c_{c0}) + \beta\pi_c(e_c^*)(1+g)^{1-\sigma} V_c + \beta(1-\pi_c(e_c^*))(1+g)^{1-\sigma} V_s \\ V_s &= u(c_{s0}) + \beta\pi_s(e_s^*)(1+g)^{1-\sigma} V_c + \beta(1-\pi_s(e_s^*))(1+g)^{1-\sigma} V_s. \end{aligned}$$

Simplifying, we get the following linear system:

$$\begin{bmatrix} 1-\beta\pi_c(e_c^*)(1+g)^{1-\sigma} & -\beta(1-\pi_c(e_c^*))(1+g)^{1-\sigma} \\ -\beta\pi_s(e_s^*)(1+g)^{1-\sigma} & 1-\beta(1-\pi_s(e_s^*))(1+g)^{1-\sigma} \end{bmatrix} \begin{bmatrix} V_c^{BGP,0} \\ V_s^{BGP,0} \end{bmatrix} = \begin{bmatrix} \left[(1-\tau_c)(1-e_c^*)\right]^{1-\sigma} \\ \left[(1-\tau_s)\left(\frac{1}{p^*}-e_s^*\right)\right]^{1-\sigma} \end{bmatrix} \frac{w_{c0}^{1-\sigma}}{1-\sigma},$$

which can be solved for the BGP welfare given all the other quantities.

Since the time T, when the economy is almost on the new BGP, is usually small in our case, we can "draw" the event tree for each type of agent starting at time zero, when the policy change occurs. The probabilities of each branch, the discount factor weighted utility at each node, are all known; the terminal nodes at time T will have value $\beta^T V_c^{BGP,T}$ or $\beta^T V_s^{BGP,T}$. One can compute welfare at time zero as a probability weighted sum of node utilities. In other words, we do not have to simulate the economy to compute the individual welfare measures, given the discreteness of types and rapid convergence to steady states.

We can then compute the consumption equivalences as above:

$$\begin{aligned} (1+\omega_s)^{1-\sigma} V_s^{BGP,old} &= V_s^{BGP,old} \\ (1+\omega_c)^{1-\sigma} V_c^{BGP,new} &= V_c^{BGP,new}. \end{aligned}$$

References

[1] Aghion, Philippe and Peter Howitt. (1998). *Endogenous Growth Theory.* Cambridge, Mass.: MIT Press.

[2] Andrade, Eduardo De Carvalho. (1998). "Growth, Distribution, and School Policy," Ph.D. dissertation, University of Chicago.

[3] Autor, David H., Lawrence F. Katz, and Alan B. Krueger (1998). "Computing Inequality: Have Computers Changed the Labor Market?" *The Quarterly Journal of Economics*, November 1998, 1169-1213.

[4] Blackenau, William F. and Beth F. Ingram. (1999). "The Welfare Implications of Factor Taxation With Rising Wage Inequality," Working Paper, University of Iowa.

[5] Cassou, Steven P. and Kevin J. Lansing. (2000). "Growth Effects of a Flat Tax." Working Paper, Kansas State University, August 23, 2000.

[6] Caucutt, Elizabeth M. and Krishna B. Kumar. (2000). "Higher Education Subsidies and Heterogeneity: A Dynamic Analysis," Working Paper, University of Southern California.

[7] Eicher, Theo S. "Interaction Between Endogenous Human Capital and Technological Change." *Review of Economic Studies* 63, no 1 (January 1996): 127-144.

[8] Forbes, Kristin J. (2000). "A Reassessment of the Relationship Between Inequality and Growth." *American Economic Review*, vol 90 no 4, 869-887.

[9] Galor, Oded, and Daniel Tsiddon. (1997). "Technological Progress, Mobility, and Economic Growth." *American Economic Review*, vol 87 no 3, 363-382.

[10] Grossman, Gene M. and Elhanan Helpman. (1991). *Innovation and Growth in the Global Economy.* Cambridge, Mass.: MIT Press.

[11] Gottschalk, Peter, and Robert Moffitt. (1994). "The Growth of Earnings Instability in the U.S. Labor Market," Brookings Papers on Economic Activity, vol 2, 217-272.

[12] Imrohoroglu, Ayse, Selahattin Imrohoroglu, and Douglas H. Joines. (1999). "Social Security in an Overlapping Generations Economy with Land," *Review of Economic Dynamics*, vol2, 638-665.

[13] Jones, Larry E. and Rodolfo E. Manuelli. (1992). "Finite Lifetimes and Growth," *Journal of Economic Theory*, vol 58, 171-197.

[14] Jones, Larry E., Rodolfo E. Manuelli, and Peter E. Rossi. (1993). "Optimal Taxation in Models of Endogenous Growth," *Journal of Political Economy*, vol 101, (June 1993), 485-517.

[15] Kahn, James A., and Jong-Soo Lim. (1998). "Skilled Labor-Augmenting Technical Progress in U.S. Manufacturing," *Quarterly Journal of Economics*, vol 108, no 4 (November 1998), 1281-1308.

[16] Kim, Se-Jik. (1992). "Taxes, growth and Welfare in an Endogenous Growth Model," Ph.D. dissertation, University of Chicago, 1992.

[17] King, Robert G., and Sergio Rebelo. (1990). "Public Policy and Economic Growth: Developing Neoclassical Implications," *Journal of Political Economy*, vol 98, no 5, pt. 2 (October 1990), S126-S150.

[18] Kumar, Krishna B. (2000). "Education, Openness, and Technical Change," Working Paper, University of Southern California.

[19] Lucas, Robert E., Jr. (1988). "On the Mechanics of Economic Development." *Journal of Monetary Economics*, vol 22, 3-42.

[20] Lucas, Robert E., Jr. (1990). "Supply-Side Economics: An Analytical Review," *Oxford Economic Papers*, vol 42, 293-316.

[21] Mendoza, Enrique G., Gian Maria Milesi-Firretti, and Patrick Asea. (1997). "On the ineffectiveness of tax policy in altering long-run growth: Harberger's superneutrality conjecture." *Journal of Public Economics*, vol 66, 99-126.

[22] Rogerson, Richard. (1988). "Indivisible Labor, Lotteries and Equilibrium," *Journal of Monetary Economics*, vol 21, 3-16.

[23] Romer, Paul M. (1986). "Increasing Returns and Long Run Growth," *Journal of Political Economy*, vol 94, no 5, 1002-1037.

[24] Romer, Paul M. (1990). "Endogenous Technological Change," *Journal of Political Economy*, vol 98, no 5, part 2, 71-102.

[25] Stokey, Nancy L. (1991). "Human Capital, Product Quality, and Growth." *Quarterly Journal of Economics,* vol 106, no.2, 587-616.

[26] Stokey, Nancy L. and Sergio Rebelo. (1995). "Growth Effects of Flat-Rate Taxes," *Journal of Political Economy*, vol 103, 519-550.

[27] Uhlig, Harald, and Noriyuki Yanagawa. (1996). "Increasing the capital income tax may lead to faster growth," European Economic Review, vol 40, 1521-1540.