Does Public Education Expansion Lead to Trickle-Down Growth?*

Sebastian Böhm[†], Volker Grossmann[‡] and Thomas M. Steger[§]

May 27, 2015

Abstract

The paper revisits the debate on trickle-down growth in view of the widely discussed changes in the distribution of earnings and income that followed a massive expansion of higher education. We propose a dynamic general equilibrium model to dynamically evaluate whether economic growth triggered by an increase in public education expenditure on behalf of those with high learning ability eventually trickles down to low-ability workers and serves them better than redistributive transfers or education policies targeted to the low-skilled. Our results suggest that promoting higher education implies that low-skilled workers first lose in terms of consumption and income but eventually gain. Policies that aim at expanding the skills of low-ability workers make them better off in the shorter run but less so in the longer run. Low-ability workers typically benefit most from redistributive transfers.

Key words: Directed Technological Change; Publicly Financed Education; Redistributive Transfers; Transitional Dynamics; Trickle-Down Growth.

JEL classification: H20, J31, O30.

* *Acknowledgements:* We are grateful to two anonymous referees as well as to Sjak Smulders and Manuel Oechslin for extremely helpful comments and suggestions. We also thank seminar particants at Tilburg University, the University of Leipzig, the Humboldt University of Berlin, the "Poverty and Inequality Workshop" 2014 at Free University of Berlin, the "70th Annual Congress of the International Institute of Public Finance (IIPF)" 2014 in Lugano, and the CESifo Area Conference in Public Sector Economics 2015, particularly Raj Chetty and our discussant Ben Heijdra, for valuable discussions.

† University of Fribourg, Department of Economics, Bd. de Pérolles 90, 1700 Fribourg, Switzerland. E-mail: sebastian.boehm@unifr.ch.

‡ University of Fribourg; CESifo, Munich; Institute for the Study of Labor (IZA), Bonn; Centre for Research and Analysis of Migration (CReAM), University College London. Address: University of Fribourg, Department of Economics, Bd. de Pérolles 90, 1700 Fribourg, Switzerland. E-mail: volker.grossmann@unifr.ch.

§ University of Leipzig; CESifo, Munich; Halle Institute for Economic Research (IWH). Address: University of Leipzig, Institute for Theoretical Economics, Grimmaische Strasse 12, 04109 Leipzig, Germany, Email: steger@wifa.uni-leipzig.de.

"Since 1979, our economy has more than doubled in size, but most of that growth has flowed to a fortunate few." (Barack Obama, December 4, 2013)

1 Introduction

Whether economic growth trickles down to the socially less fortunate has been a key debate for many decades in the US and elsewhere (e.g. Kuznets, 1955; Thornton, Agnello and Link, 1978; Hirsch, 1980; Aghion and Bolton, 1997; Piketty, 1997). In particular, social desirability and choices of growth-promoting policies may critically depend on their expected trickle-down effects. For instance, massive expansion of high school and college education throughout the 20th century has led to a surge in the relative supply of skilled labor (Goldin and Katz, 2008; Gordon, 2013). Goldin and Katz (2008) document the important role of the public sector for this development, particularly between 1950 and 1970.[1] Despite steady economic growth, however, median (full-time equivalent) earnings of males have almost stagnated from the 1970s onwards (e.g. Katz and Murphy, 1992; Acemoglu and Autor, 2012; DeNavas-Walt, Proctor and Smith, 2013). Moreover, earnings of less educated males fell considerably (Acemoglu and Autor, 2011, Tab. 1a).

We propose a suitable dynamic general equilibrium framework with directed technical change, heterogeneous agents and a key role of human capital for economic growth to evaluate the effects of public expenditure reforms on the evolution of living standards over time. We investigate, in particular, whether economic growth triggered by an increase in public education expenditure on behalf of those with high learning ability eventually trickles down to low-ability workers. Moreover, we also wish to assess whether higher education expansion serves low-skill workers better than redistributive transfers or education policies targeted to them specifically. Hence, the following public expenditure policies are comparatively examined: (i) public education finance on behalf

[1] For instance, the fraction of college students in publicly controlled institutions gradually increased between 1900 and 1970. Between 1950 and 1970, it increased from 0.5 to almost 0.7 among students with four years of college attendance (Goldin and Katz, 2008; Fig. 7.7).

of high-ability workers (e.g. post-secondary and tertiary education), (ii) income transfers towards individuals who do not acquire more advanced education (e.g. because of limited ability), and (ii) public education finance targeted to low-ability workers (e.g., qualification and training programs for low-skill workers).[2]

Whether and when growth promoted by education expansion trickles down to low-skilled workers is a key question for at least three reasons. *First*, the evolution of the earnings distribution has recently provoked an intensive policy debate in the US and elsewhere (e.g. Stiglitz, 2012; Deaton, 2013; Mankiw, 2013; Piketty, 2014).[3] For instance, in his maybe most widely received speech of his US presidency (December 4, 2013), Barack Obama referred to it as "the defining challenge of our time", criticizing that "a trickle-down ideology became more prominent".[4] He also urged that "we need to set aside the belief that government cannot do anything about reducing inequality". In fact, the tax-transfer system in the US is rather unsuccessful to improve living standards of the working-poor, compared to other advanced countries (Gould and Wething, 2012). *Second*, upward social mobility has been proved severely limited by intergenerational transmission of learning ability and/or human capital, implying that a significant fraction of individuals may not acquire more than basic education for a long time to come (Corak, 2013).[5] It is thus important to know whether those individuals can profit from stimulating economic growth by promoting human capital expansion of high-ability workers. We focus on the economic situation of low-ability in-

[2] The literature on the effectiveness of programs to promote basic education on behalf of low-income earners is mixed. Some evidence suggests that their success is limited unless governments intervene at a very young age (Cunha et al., 2006). However, the negative view on programs targeted to adolescents and young adults, has been qualified. Schochet et al. (2008) evaluate the Job Corps program, which targets economically disadvantaged youth aged 16 to 24. They find positive effects on skill development but mixed effects on earnings. Osikominu (2013) provides encouraging evidence on long term active labor market policy in Germany. Kautz et al. (2014) argue that education interventions targeted to adolescents and young adults are successful if they involve mentoring and emphasize non-cognitive skills.

[3] The earnings distribution has changed markedly also in Continental Europe, although later than in the US; see e.g. Dustmann, Ludsteck and Schönberg (2009) for evidence on Germany.

[4] See the following URL (retrieved on May 25, 2015): https://www.whitehouse.gov/the-press-office/2013/12/04/remarks-president-economic-mobility.

[5] There is overwhelming evidence for the hypothesis that the education of parents affects the human capital level of children. For instance, Plug and Vijverberg (2003) and Black, Devereux and Salvanes (2005) show that children of high-skilled parents have a higher probability of being high-skilled.

dividuals by deliberately ruling out the possibility of social mobility and comparatively examine economic policy alternatives on behalf of those who will stay disadvantaged. That is, rather than investigating economic inequality per se, we analyze trickle-down growth in a strict sense.[6] *Third*, the literature on directed technological change, initiated by Von Weizsäcker (1966) and advanced by Acemoglu (1998, 2002), allows for the possibility that an increase in the supply of human capital may lead to skill-biased technological change, thus contributing to the differential evolution of living standards across individuals in the first place. Particularly, it is not evident whether and when workers with only basic education benefit from increased public education spending targeted to higher education institutions. It is therefore salient for addressing our research questions to capture the possibility that technological progress does not automatically benefit high-skilled and low-skilled labor symmetrically.[7]

Standard analyses of directed technological change models are inadequate to enter the trickle-down growth debate, because they exclusively focus on the long run and assume that skill supply is exogenous. For instance, as acknowledged by Autor and Acemoglu (2012), such analyses are unsuccessful to explain falling earnings at the bottom of the distribution of income. Rather, our goal is to dynamically evaluate the impact of an increase in public education expenditure that potentially affects both R&D and education decisions, being in line with the observed income dynamics in the last decades, and helps to understand future dynamics.[8]

More specifically, our framework rests on the following features: (i) the government can extend redistributive transfers or promote publicly financed education targeted to

[6] A comprehensive discussion of inequality dynamics should in fact allow for social mobility. This would, however, come at the cost of losing the focus on our research question. Social mobility in the US is indeed quite low, as demonstrated by Chetty et al. (2014). We also abstain from modeling early intervention programs targeted to young children from disadvantaged households (e.g., Kautz et al., 2014).

[7] Che and Zhang (2014) argue that the higher education expansion in China in the late 1990s had a causal positive effect on technological change particularly in human capital intensive industries, suggesting that technical change endogenously benefits primarily high-skilled workers.

[8] We employ the relaxation algorithm proposed by Trimborn, Koch and Steger (2008) to analyze the transitional dynamics of the resulting non-linear, highly dimensional, saddle-point stable, differential-algebraic system. Despite the complexity of our model, the long run equilibrium can be characterized analytically. This is important for understanding the basic mechanisms and facilitates a careful model calibration.

high-ability workers (higher education) or targeted to low-ability workers (e.g., training programs targeted to low-skill adult workers or second-chance programs for high-school drop outs); (ii) low-ability households rely on the public education system, may receive income transfers and potentially benefit from subsidizing higher education via various general equilibrium effects; (iii) there are distortionary taxes on (labor and capital) income and capital gains that are adjusted to finance the respective policy intervention; (iv) growth is endogenously driven by directed technological change that potentially favors different types of skills asymmetrically; (v) only high-ability workers can be employed in R&D or education activities; (vi) the accumulation of physical capital, human capital and R&D-based knowledge capital interact with public policy in determining the evolution of living standards over time.

Our key findings may be summarized as follows. *First*, when the government raises the fraction of tax revenue devoted to publicly finance higher education, earnings, consumption and net income of low-ability workers initially decrease compared to the baseline scenario without policy reform. Consistent with empirical evidence, expansion of higher education is followed by rising inequality and temporarily lower wages at the bottom of the earnings distribution. Only after considerable time elapsed, the economic situation of the least educated improves and they eventually become better off. *Second*, an equally sized increase in public education expenditure targeted to low-ability workers (qualification and training programs) raises their earnings at all times, allowing them to raise consumption in the shorter run. It also triggers off unskilled-biased technological change. However, adverse general equilibrium (growth) effects driven by high opportunity costs in terms of high-skilled labor use, which are rooted in low education returns and tax distortions, imply that low-skilled workers lose in the longer run. *Third*, low-skilled workers are best off by increasing redistributive transfers (equally sized) both in the shorter and in the longer run.

The paper is organized as follows. In Section 2, we briefly discuss the related literature. In Section 3, we set up a comprehensive growth model. Section 4 characterizes its equilibrium analytically. Section 5 presents the calibration strategy. In Section 6 we employ numerical analysis to dynamically evaluate the trickle down dynamics of policy reforms. The last section concludes.

2 Related Literature

We cannot review the vast literature on the interplay between economic growth and economic outcomes for less educated individuals, but rather focus on the most related work. Galor and Zeira (1993) show that human capital investments are suboptimally low under credit constraints. If the wedge between the borrowing and the lending rate is sufficiently large, inequality is not only harmful for growth but may also increase over time, i.e., growth does not trickle down. Aghion and Bolton (1997), Piketty (1997) and Matsuyama (2000) examine the evolution of wealth distribution under imperfect credit market with fixed investment requirements for entrepreneurial projects. They identify conditions under which growth may trickle down and argue that (lump sum) wealth redistribution to the poor may speed up this process by mitigating credit constraints. In contrast to this literature, our focus is on the interplay between physical capital accumulation, human capital accumulation and technological change directed to different types of workers. Most importantly, we stress the role of the public sector for education and redistribution, both financed by distortionary taxation.[9]

Goldin and Katz (2008) argue that the evolution of skill premia can be explained by the pace at which the relative supply of skills keeps track with the relative demand for skills as driven by skill-biased technological change. However, as already pointed out by Acemoglu and Autor (2012), their analysis does not address the possible feedback effect of rising skill supply. Such effects result from education expansion via endogenously biased technological change, altering the relative demand for skills. Closest to our analysis, Acemoglu (1998, 2002) introduces the idea that the relative demand for different types of workers via technological change is endogenous to the supply of human capital. While he focusses on the long run effects of an exogenous increase in human capital, our interest lies in the transitional dynamics when both the formation

[9] Moreover, empirical evidence suggests a minor role of binding credit constraints for education finance in the US (Lochner and Monge-Naranjo, 2011).

of human capital and the extent and direction of technological change are endogenous to public policy reforms. Galor and Moav (2000) examine distributional effects of biased technological change in a dynamic model of endogenous skill supply. There are two main differences to our work. First, whereas Galor and Moav (2000) are interested in the evolution of wage inequality when the rate of (by assumption ability-biased) productivity growth starts below steady state, we evaluate public policy reforms. In particular, we consider the effects on income dynamics of a publicly financed expansion of education on behalf of high-ability individuals versus redistributive transfers and publicly financed promotion of skills on behalf of low-ability individuals. Second, in our model, technological change is based on R&D decisions which are potentially skill-biased endogenously.

Finally, there is a literature on the political economy of public subsidies of higher education. Based on a setup where all individuals would benefit from education and differ in initial endowments, Fernandez and Rogerson (1995) argue that under majority voting incomplete education subsidies emerge that exclude the poor from education because of credit constraints. Since education subsidies are tax-financed as in our model, there thus is redistribution from the poor towards the rich. We do not aim to provide a positive explanation of the policies we consider. Our setup is, however, consistent with the view that the poor may support higher education subsidies because they benefit from it in the longer run.

3 The Model

Consider an infinite-horizon, Ramsey-type growth model in continuous time with three growth engines: (i) physical capital accumulation, (ii) education, and (iii) endogenous, directed technical change. These growth engines interact with each other and are affected by various public policy instruments.

3.1 Firms

There is a homogenous final good with price normalized to unity. Following Acemoglu (2002), final output is produced under perfect competition according to

$$Y = \left[(X_H)^{\frac{\varepsilon-1}{\varepsilon}} + (X_L)^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}}, \tag{1}$$

$\varepsilon > 0$. X_L and X_H are composite intermediate inputs. They are also produced under perfect competition, combining capital goods ("machines") with high-skilled and low-skilled labor, respectively. Formally, we have

$$X_H = (H^X)^{1-\alpha} \int_0^{A_H} x_H(i)^\alpha di, \tag{2}$$

$$X_L = (L^X)^{1-\alpha} \int_0^{A_L} x_L(i)^\alpha di, \tag{3}$$

$0 < \alpha < 1$, where $x_H(i)$ and $x_L(i)$ are inputs of machines, indexed by i, which are complementary to the amount of human capital in this sector, H^X, and low-skilled labor, L^X, respectively. The mass ("number") of machines, A_H and A_L, expands through horizontal innovations, as introduced below. The initial number of both types of machines are given and positive; $A_{H,0} > 0$, $A_{L,0} > 0$.

In each machine sector there is one monopoly firm – the innovator or the buyer of a blueprint for a machine. They produce with a "one-to-one" technology by using one unit of final output to produce one machine unit. The total capital stock, K, in terms of the final good, thus reads as

$$K = \int_0^{A_H} x_H(i) di + \int_0^{A_L} x_L(i) di. \tag{4}$$

Machine investments are financed by bonds sold to households. In each machine sector there is a competitive fringe which can produce a perfect substitute for an existing machine (without violating patent rights) but is less productive: input coefficients are higher than that of the incumbents by a factor $\kappa \in (1, \frac{1}{\alpha}]$ in both sectors.[10] Parameter κ determines the price-setting power of firms and allows us to disentangle the price-mark up from output elasticities, which is important for a reasonable calibration of the model. Physical capital depreciates at rate $\delta_K \geq 0$.

There is free entry into two kinds of competitive R&D sectors. In one sector, a representative R&D firm directs human capital to develop blueprints for new machines used to produce the human capital intensive composite input, X_H, the other sector to produce X_L. To each new idea a patent of infinite length is awarded. Following Jones (1995), ideas for new machines in the R&D sectors are generated according to

$$\dot{A}_H = \tilde{\nu}_H (A_H)^{\phi} H_H^A, \ \tilde{\nu}_H = \nu \cdot (H_H^A)^{-\theta}, \tag{5}$$

$$\dot{A}_L = \tilde{\nu}_L (A_L)^{\phi} H_L^A, \ \tilde{\nu}_L = \nu \cdot (H_L^A)^{-\theta}, \tag{6}$$

where H_H^A and H_L^A denote human capital input in the R&D sector directed to the human capital intensive and low-skilled intensive intermediate goods sector, respectively. $\nu > 0$ is a R&D productivity parameter. $\theta \in (0,1)$ captures a negative R&D ("duplication") externality, discussed in Jones and Williams (2000). It measures the gap between privately perceived constant R&D returns of human capital and socially decreasing returns. We assume that $\phi \in (0,1)$. $\phi > 0$ captures a positive ("standing on shoulders") knowledge spillover effect (for empirical support, see e.g. Audretsch and Feldman, 1996).[11]

3.2 Households

There are two types of dynastic households, high-ability "type$-h$" and low-ability "type$-l$" households, who inelastically supply their human capital to a perfect labor market. Their population sizes, N_h and N_l, grow at the same and constant exponential

[10] See Aghion and Howitt (2005), among others, for a similar way of capturing a competitive fringe.

[11] Two remarks are in order: First, Acemoglu (1998, 2002) employs the "lab-equipment" approach with capital investment in R&D. Since empirically R&D costs are mainly salaries for R&D personnel, we prefer specifications (5) and (6). Second, $\phi < 1$ implies that growth is "semi-endogenous" (Jones, 1995), i.e. would cease in the long run if population growth were absent.

rate, $n \geq 0$, i.e. N_l/N_h is time invariant. Type$-l$ individuals can only work as low-skilled workers in the respective machine sector, whereas type$-h$ individuals can be employed in all alternative uses. We rule out social mobility to capture the intergenerational transmission of learning ability in a pointed form and to deliberately model an unfavorable situation for type$-l$ individuals, motivated by our interest in trickle-down dynamics.[12]

Households own machine firms by holding equity and purchasing bonds. Equity finances blueprints for machine producers, whereas bonds provide capital that serves as input for machine producers.

3.2.1 Human Capital Formation

Skill formation depends on the (rival) human capital input of type$-h$ individuals devoted to education ("teachers"). Let h_h^E and h_l^E denote the teaching input in educational production per type$-h$ and type$-l$ individual, respectively. Human capital of type$-h$ and type$-l$ individuals depreciate at the same constant rate $\delta_H > 0$ and evolve according to

$$\dot{h} = \xi (h_h^E)^\beta h^\eta - \delta_H h, \tag{7}$$

$$\dot{l} = \xi (h_l^E)^\gamma l^\eta - \delta_H l, \tag{8}$$

where $\xi > 0$, $\beta, \gamma \in (0, 1)$, $\eta \geq 0$, $\beta + \eta < 1$. Initial levels $h_0 > 0$ and $l_0 > 0$ are given. If $\eta > 0$, there is intergenerational human capital transmission for both types. Our human capital accumulation process of type$-h$ individuals is similar to Lucas (1988). However, Lucas (1988) assumes the change of human capital is proportional to the stock (in our formulation, holding if $\eta = 1$). This would mean that individual human capital levels grow without bounds, a possibility that we rule out.

[12] This simplifying assumption seems in fact empirically plausible: In 2012, the proportion of young students (20-34 year-olds) in tertiary education whose parents have below upper secondary education (12 percent of the total population) was 8 percent. Among those with tertiary educated parents (48 percent of the total population), it is 58 percent (OECD, 2014, Tab. A4.1a).

3.2.2 Preferences

As will become apparent, labor income taxation distorts the human capital investment decision. We abstract, however, from capturing a labor-leisure choice. On the one hand, the elasticity of labor supply with respect to net wages is estimated to be positive in the shorter run. Hours worked have, however, declined over a longer time horizon in many growing economies (e.g. Lee, McCann and Messenger, 2007), suggesting that the long run wage elasticities of labor supply are negative. In view of these conflicting findings, we assume that households do not draw utility from leisure.

Let subscript t on a variable index time (suppressed if not leading to confusion). Preferences of individuals of type $j \in \{h, l\}$ are represented by the standard, dynastic welfare function

$$U_j = \int_0^\infty \frac{(c_{jt})^{1-\sigma} - 1}{1-\sigma} e^{-(\rho-n)t} \mathrm{d}t, \tag{9}$$

$\sigma > 0$, where c_{jt} is consumption of a type$-j$ individual at time t.

3.3 Government

Let w_h and w_l denote the wage rate of type$-h$ and type$-l$ individuals per unit of human capital. We focus throughout on the case where type$-l$ individuals earn (endogenously) less than type$-h$ individuals at all times and have lower marginal tax rates. Formally, suppose that the marginal income tax rate is given by an increasing step-function $\tilde{\tau}(\cdot)$ fulfilling $\tilde{\tau}(w_h h) \equiv \tau_h > \tau_l \equiv \tilde{\tau}(w_l l)$. The step-function $\tilde{\tau}$ is such that τ_h and τ_l are time-invariant for the income ranges we consider.[13] Only type$-l$ individuals earn sufficiently little to be eligible for a transfer payment, denoted by T.

The human capital levels of both type$-h$ and type$-l$ individuals are affected by public education policy. For type$-l$ individuals, teaching input h_l^E is exclusively publicly financed. Examples include government-sponsored qualification and training programs on behalf of adolescents and young adults from disadvantaged backgrounds, such as the JOBSTART and Job Corps (Bloom, 2010), as well as longer-term active

[13] Ensuring this outcome may require that the mapping from income brackets to marginal tax rates is adjusted when income levels grow, i.e. function $\tilde{\tau}(\cdot)$ is adjusted over time.

labor market programs, such as the Employment Training Panel, California, and the Literacy/Basic Skills Program, New Jersey (Crosley and Roberts, 2007).[14] Type$-h$ individuals choose their educational input, h_h^E, and receive a subsidy at rate ϑ on the costs. Examples include post-secondary schooling and college attendance.

The interest rate for bonds is denoted by r. Dividends from equity holdings and bond holdings are taxed by the same constant rate τ_r. Financial assets per dynasty member, a_h and a_l, evolve according to

$$\begin{aligned} \dot{a}_h &= y_h - (1-\vartheta)w_h h_h^E - c_h \quad \text{with } y_h \equiv [(1-\tau_r)r - n]a_h + (1-\tau_h)w_h h \quad (10) \\ \dot{a}_l &= y_l - c_l \quad \text{with } y_l \equiv [(1-\tau_r)r - n]a_l + (1-\tau_l)w_l l + T, \quad (11) \end{aligned}$$

respectively. Initial asset holdings, $a_{h,0} > 0$, $a_{l,0} > 0$, are given. Type$-h$ households maximize U_h with respect to consumption and teaching inputs subject to (7) and (10), whereas type$-l$ households maximize U_l subject to (11).

An increase in τ_h distorts the education decision of type$-h$ individuals. An increase in the tax rate on capital income, τ_r, distorts savings decisions of households and investment decisions of firms. We also allow for taxation of capital gains, taxed with constant rate τ_g. An increase in τ_g distorts the portfolio decision of households (substituting equity for bonds), therefore affecting R&D investments.

Assuming that the economy initially is in steady state, we comparatively investigate how (i) changes in transfers on behalf of type$-l$ individuals, T, (ii) changes in the fraction of human capital devoted to educating type$-l$ individuals, $\mathfrak{h}_l^E \equiv h_l^E/h$, and (iii) changes in the subsidy rate ϑ on education costs of type$-h$ individuals affect the economic situation of type$-l$ individuals. To describe the poor's economic situation, we focus on their consumption level, c_l, their wage income, $W_l \equiv w_l l$, and their total net income, y_l. We assume that the three public expenditure categories are financed via taxation of income and capital gains. For the comparative-static policy analysis, we assume that, realistically, the fractions of tax revenue devoted to financing the three

[14] The effectiveness of these programs is discussed by Cunha et al. (2006), Schochet et al. (2008), and Kautz et al., (2014).

spending categories do not add up to one. That is, there is an additional public spending category which may additively enter the utility function (like public expenditure for defense, the legal system, public order and safety). The government cannot save or incur debt. The specific policy reforms are described in section 6.

4 Equilibrium Analysis

This section provides a number of analytical results that are important to better understand the major implications for wage income of low-skilled workers. The equilibrium definition is standard and relegated to Appendix A.

4.1 Preliminaries

The transversality conditions of the household optimization problems and the requirement of finite intertemporal welfare levels U_h and U_l requires the following parameter restriction to hold

$$\rho - n + (\sigma - 1)g > 0 \text{ with } g \equiv \frac{n(1-\theta)}{1-\phi}. \tag{A1}$$

As will become apparent, g is the long run growth rate of individual consumption levels, individual income components, and knowledge measures A_H, A_L. Thus, in the long run, technological change turns out to be unbiased.

Profit maximization of non-R&D producers implies two intermediate results that remind us on the mechanics of directed technical change.

Lemma 1. *Define $\psi \equiv \alpha + \varepsilon(1-\alpha)$. The relative wage per unit of human capital between type$-h$ and type$-l$ individuals reads as*

$$\frac{w_h}{w_l} = \left(\frac{H^X}{L^X}\right)^{-\frac{1}{\psi}} \left(\frac{A_H}{A_L}\right)^{\frac{\psi-1}{\psi}}. \tag{12}$$

All proofs are relegated to Appendix B. According to (12), ψ is the "derived" elasticity between high-skilled and low-skilled labor in production (Acemoglu, 2002). For given productivity levels, an increase in relative amount of type$-h$ human capital devoted to manufacturing, H^X/L^X, by one percent reduces the relative wage rate, w_h/w_l, by $1/\psi$ percent. Notably, if $\varepsilon > 1$, then $\varepsilon > \psi > 1$; if $\varepsilon < 1$, then $\varepsilon < \psi < 1$.

Let P_H^X and P_L^X denote the price of the high-skilled intensive and low-skilled intensive composite intermediate good used in the final goods sector, respectively. An increase in the relative knowledge stock of the high-skilled intensive sector, A_H/A_L, has two counteracting effects on relative wage rate as given by (12). First, the relative productivity of type$-h$ human capital in the production of composite intermediates rises, w_h/w_l increases for a given relative price of intermediates, $P \equiv P_H^X/P_L^X$. Second, however, since relatively more of the high-skilled intensive composite good is produced when A_H/A_L rises, the relative price of composite goods, P, decreases for given labor inputs. Through this effect, the relative value of the marginal product of type$-h$ human capital declines. If and only if the elasticity of substitution between the composite intermediates is sufficiently high, $\varepsilon > \psi > 1$, the first effect dominates the second one (vice versa if $\varepsilon < \psi < 1$).

The next result provides insights on relative R&D incentives in the two R&D sectors. The respective profits of an intermediate good firms (symmetric within sectors) are denoted by π_H and π_L.

Lemma 2. *The relative instantaneous profit of machine producers reads as*

$$\frac{\pi_H}{\pi_L} = \left(\frac{A_H}{A_L}\right)^{-\frac{1}{\psi}} \left(\frac{H^X}{L^X}\right)^{\frac{\psi-1}{\psi}}. \tag{13}$$

There are counteracting effects of an increase in relative employment in composite input production, H^X/L^X, on relative R&D incentives. First, for a given relative price of the high-skilled intensive good, P, relative profits in the high-skilled intensive sector rise due to the complementarity between labor and machine inputs in (2) and (3) ("market size effect"). Second, however, P falls in response to an increase in relative output of the high-skilled intensive good ("price effect"). In the case where $\varepsilon > \psi > 1$, the first effect dominates the second one, and vice versa if $\varepsilon < \psi < 1$.

Moreover, as already discussed after Lemma 1, an increase in the relative knowledge stock of the high-skilled intensive sector, A_H/A_L, reduces the relative price P. Thus, relative profits π_H/π_L decline. The magnitude of the elasticity of π_H/π_L with respect to A_H/A_L is inversely related to the "derived" elasticity between high-skilled and low-skilled labor in production, ψ.

4.2 Balanced Growth Equilibrium

Empirical estimates suggest that the elasticity of substitution between high-skilled and low-skilled labor is larger than one (Johnson, 1997). Moreover, existence and uniqueness of a balanced growth equilibrium (BGE), in which all variables grow at a constant rate, requires that the derived elasticity of substitution, ψ, is bounded upwards. Formally, we may assume

$$1 < \psi \leq \frac{2 - \phi - \theta}{1 - \theta}. \tag{A3}$$

Let superscript (*) denote long run values (in BGE) throughout and define the human capital per type$-h$ individual devoted to manufacturing by $h^X \equiv H^X/N_h$.

Proposition 1. *Under (A1) and (A2), there exists a unique BGE which can be characterized as follows:*

(i) c_h, c_l, a_h, a_l, A_H, A_L, w_h, w_l, T grow with rate g;

(ii) L^X, H^X, H_H^A, H_L^A, P_H^A, P_L^A grow with rate n;

(iii) X_H, X_L grow with rate $g+n$;

(iv) r, P_H^X, P_L^X are stationary;

(v) the long run fraction of human capital of type$-h$ individuals devoted to education, $\mathfrak{h}_h^{E*}$, and the long run human capital level, h^*, are also stationary and given by

$$\mathfrak{h}_h^{E*} = \frac{1 - \tau_h}{1 - \vartheta} \frac{\beta \delta_H}{\rho - n + (\sigma - 1)g + (1 - \eta)\delta_H}, \tag{14}$$

$$h^* = \left[\frac{\xi (\mathfrak{h}_h^{E*})^\beta}{\delta_H} \right]^{\frac{1}{1-\beta-\eta}}, \tag{15}$$

respectively; thus, both $\mathfrak{h}_h^{E*}$ and h^* are decreasing in labor income tax of high-earners,

τ_h, and increasing in the education subsidy rate ϑ;

(vi) the long run skill level of type$-l$ individuals, l^*, is given by

$$l^* = \left(\frac{\xi (\mathfrak{h}_l^E h^*)^\gamma}{\delta_H} \right)^{\frac{1}{1-\eta}}; \tag{16}$$

thus, l^* is decreasing in τ_h and increasing in both ϑ and the fraction of human capital of type$-h$ individuals devoted to education, $\mathfrak{h}_l^E$;

(vii) changes in τ_h or ϑ exclusively affect the long run fraction of human capital of type$-h$ individuals devoted to manufacturing, h^{X*}, through the effects on $\mathfrak{h}_h^{E*}$; moreover, h^{X*} is increasing in the tax rate of both capital income and capital gains, τ_r and τ_g, respectively; finally, h^{X*} is decreasing in the fraction of human capital devoted to educating type$-l$ individuals, $\mathfrak{h}_l^E$.

According to (1), Proposition 1 implies that also per capita income grows at rate g in steady state. The result parallels the well-known property of semi-endogenous growth models that the economy's long run growth rate is policy-independent (e.g. Jones, 1995, 2005). By contrast, the human capital allocation is affected by policy parameters, with effects on the transitional dynamics. First, according to part (v) of Proposition 1, labor income taxation ($\tau_h > 0$) distorts educational investments of high-ability households. An increase in the rate of the education subsidy, ϑ, mitigates this distortion. Both policy parameters, τ_h and ϑ, affect the long run level and allocation of human capital only through the effect on the long run fraction of human capital of type$-h$ individuals devoted to education, $\mathfrak{h}_h^{E*}$. An increase in $\mathfrak{h}_h^{E*}$ has two counteracting effects on the steady state fraction h^{X*} of human capital of type$-h$ workers employed in manufacturing. On the one hand, h^{X*} rises, as the amount of human capital expands (increase in h^*). On the other hand, employing more teachers to educate type$-h$ individuals means a reallocation of existing human capital of type$-h$ workers away from manufacturing. If h^{X*} rises when higher education expands, low-ability individuals gain via the complementarity of composite inputs in (1). Empirically, expansion of higher education has undoubtedly raised the human capital input in all uses. Finally, promoting skill development of type$-l$ workers by raising $\mathfrak{h}_l^E$, non-surprisingly, raises long run skill level l^* (part (vi) of Proposition 1). It also reallocates human capital of type$-h$ workers from manufacturing to teaching, thus reducing h^{X*} (part (vii) of Proposition 1).

One can show (see the proof of Proposition 1) that in the case where the derived elasticity of substitution between the two types of workers, ψ, is high or the R&D technology parameters ϕ and θ lead to a high steady state growth rate, g, such that the second inequality in (A2) is violated, there may be two interior BGE. To understand why this can happen, suppose again $\psi > 1$, such that the market size effect discussed after Lemma 2 dominates the price effect, and h^X rises. According to (13), relative profits π_H/π_L thus rise, boosting the relative knowledge stock A_H/A_L, all other things being equal ("skill-biased technological change"). This effect is large when the knowledge spillover effect is high (i.e. ϕ is high), the duplication externality is low (i.e. θ is low) and/or ψ is high. If it is sufficiently large, the equilibrium fraction of type$-h$ human capital devoted to R&D targeted to type$-l$ intensive production declines, implying that h^X rises further. In this case, there may be multiple equilibria. Pertubating, by increasing h^X, an interior BGE in which h^X is low to begin with, h^X rises by even more than initially. In other words, the BGE is "unstable". If there are two interior BGE, then one is "stable", characterized by a higher h^X than the "unstable" one (see the Remark in Appendix B and Fig. A.1 in the online-appendix). Property (vii) still holds in the stable BGE. In both types of equilibria, properties (i)-(vi) are satisfied. In the remainder of this section and most of the numerical analysis, we will focus on the case with a unique BGE.

The next result shows the effects of changes in policy instruments affecting the long run human capital input in higher education on the steady state wage income level of low-skilled workers, $W_l^* = w_l^* l^*$.

Proposition 2. *Under (A1) and (A2), wage income of type$-l$ individuals in the long run, W_l^*, is decreasing in τ_h and increasing in ϑ, whenever h^{X*} is decreasing in τ_h and increasing in ϑ.*

Because employment of type$-h$ workers in manufacturing is complementary to low-skilled employment through the imperfect substitutability of composite inputs in final goods production, the long run wage rate per skill unit of type$-l$ workers, w_l^*, critically depends on τ_h and ϑ through its impact on the long run fraction of human capital input in higher education, $\mathfrak{h}_h^{E*}$ (Proposition 1). First, recall from (15) that an increase in subsidy rate ϑ or a decrease in tax rate τ_h raises the long run level of human capital per type$-h$ individual, h^*. If in turn the fraction of human capital devoted to production, h^{X*}, rises, the output level of the human capital intensive composite income, X_H, increases. For given knowledge stocks, because of the complementarity of composite inputs in final goods production, this raises the price of the low-skilled labor intensive composite input, P_L^X. Moreover, as discussed after Lemma 2, for $\psi > 1$, the market size effect of an increase in H^X on profits for high-skilled intensive production, π_H, dominates the price effect. Thus, an increase in ϑ triggers off innovations directed to type$-h$ human capital, i.e. A_H rises. As this also raises relative output X_H of the high-skilled intensive composite good, P_L^X increases through this effect as well. As a result, in the plausible case that raising ϑ or lowering τ_h raises h^{X*}, the value of the marginal product of low-skilled labor, w_l^*, increases, all other things being equal. Moreover, an increase in the level of human capital of type$-h$ individuals raises the skill level of type$-l$ workers, l^* (part (vi) of Proposition 1). This has two counteracting effects on the value of the marginal product of low-skilled labor, w_l^*. First, there is a market size effect of an increase in L^X on profits for low-skilled intensive production, π_L. In turn, R&D activity that makes type$-l$ workers more productive is fostered. Second, however, the price of the low-skilled labor intensive composite input, P_L^X, falls (price effect). One can show that the effect of an (endogenous) increase in skill level l^* unambiguously raises $W_l^* = w_l^* l^*$.

We next turn to the relative wage rate per unit of skill between the two types of workers. Although our focus is on the economic situation of low-ability dynasties, considering wage inequality is interesting to gain further confidence in the empirical relevance of our analysis.

Proposition 3. *Under (A1) and (A2), the following holds for the relative wage per unit of human capital between type−h and type−l individuals in the long run, w_h^*/w_l^*.*

(i) If $\psi = \frac{2-\phi-\theta}{1-\theta}$ (i.e., ψ equals the upper bound in (A2)), w_h^*/w_l^* is independent of policy instruments;

(ii) otherwise (if $\psi < \frac{2-\phi-\theta}{1-\theta}$), w_h^*/w_l^* is decreasing in τ_r and τ_g and increasing in $\mathfrak{h}_l^E$.

Consider an (endogenous) increase in the level of type−h human capital for the production of composite inputs, which raises H^X/L^X in long run equilibrium. For $\psi > 1$, an increase in H^X/L^X spurs innovation directed to type−h human capital relatively more, thus raising A_H/A_L. According to Lemma 1, for $\psi > 1$, an increase in the "relative knowledge stock", A_H/A_L, would raise the relative wage rate per unit of type−h human capital, w_h/w_l, for given H^X/L^X. However, under limited (derived) substitutability between type−h and type−l labor, ψ, as also assumed in (A2), the effect is not large enough to overturn the negative impact of an increase of H^X/L^X on w_h/w_l for a given relative knowledge stock, A_H/A_L (see (12) in Lemma 1). If ψ equals the upper bound in (A2), both effects exactly cancel. This explains part (i) of Proposition 3. Now consider an endogenous increase in $\mathfrak{h}_h^E$ which governs the evolution of human capital of type−h workers, possibly raising the level of high-skilled employment in manufacturing, h^{X*}, in addition to raising l^* (Proposition 1). Thus, an increase in subsidy rate ϑ or a decrease in tax rate τ_h has an ambiguous effect on relative employment in manufacturing, H^X/L^X, and thus an ambiguous effect on relative wage rate w_h^*/w_l^*. By contrast, promoting education of type−l workers (increase in $\mathfrak{h}_l^E$) reduces h^{X*} and raises l^*, thus unambiguously raising w_h^*/w_l^* under limited (derived) substitutability between the two types of labor ($\psi < \frac{2-\phi-\theta}{1-\theta}$). Finally, raising the tax rate of capital income or capital gains, τ_r and τ_g, respectively, implies a reallocation of high-skilled labor from R&D activity to manufacturing (raising h^{X*}), thus lowering w_h^*/w_l^*. This concludes the discussion of Proposition 3.

5 Calibration

The baseline calibration is summarized in Table 1. The parameter values are mostly based on observables (including policy parameters) for the US economy in the 2000s before the financial crisis started in 2007, assuming that the US was in steady state initially (i.e. before the considered policy reforms). We calibrate variables related to type$-h$ individuals by values for the representative individual with at least high school diploma and type$-l$ individuals by values for the representative high-school drop-out. According to OECD (2014), the share of those among the 25-64 year old with less than upper secondary education in the year 2005 is 12 percent, suggesting that the (by assumption time-invariant) relative population size reads as $N_l/N_h = 12/88 = 0.14$.

Parameter	Value	Parameter	Value	Parameter	Value	Variable	Value
n	0.01	α	0.4	τ_g	0.1	N_l/N_h	0.14
g	0.02	ε	1.83	τ_l	0.21	R_h^*/R_l^*	3.7
κ	1.3	δ_K	0.04	τ_h	0.31	$R\&D^*$	0.031
θ	0.5	δ_H	0.03	τ_r	0.17	Ω^*	2.06
ϕ	0.75	β	0.33	s_h^E	0.12	sav^*	0.21
ρ	0.02	γ	0.11	s_l^E	0.016	K^*/Y^*	3
σ	1.91	η	0.2	s^T	0.13	r^*	0.07
ψ	1.5	ξ	0.3	$\bar{\nu}$	0.2	a_h^*/a_l^*	5

$\phi = 1 - \frac{n(1-\theta)}{g}$, $\sigma = \frac{(1-\tau_r)r^*-\rho}{g}$, $\varepsilon = \frac{\psi-\alpha}{1-\alpha}$ and $\delta_K = \frac{sav^*}{K^*/Y^*} - n - g$ are implied.

Table 1: Baseline calibration.

5.1 Policy Parameters

Following the reasoning in Grossmann, Steger and Trimborn (2015) we set the capital gains tax rate to $\tau_g = 0.1$. Tax rates τ_l and τ_h are approximated by the marginal personal tax rate on gross labor income in the year 2005 at 67 and 133 percent of the average labor income, respectively, combining federal and subcentral government taxes and excluding social security contribution rates; this gives us $\tau_l = 0.21$ and

$\tau_h = 0.31$.[15] Moreover, we assume $\tau_r = 0.17$, which coincides with the US net personal capital income tax (equal to the net top statutory rate to be paid at the shareholder level, taking account of all types of reliefs and gross-up provisions at the shareholder level) after the 2003 tax reform (Murray, Singh and Wang, 2012).

Let us denote the fraction of tax revenue devoted to redistributive transfers on behalf of type$-l$ individuals by s^T, and that devoted to education of type$-h$ and type$-l$ individuals by s_h^E and s_l^E, respectively, $s^T + s_l^E + s_h^E \leq 1$. We approximate redistributive transfers by total "government social benefits" at the federal, state and local level between 2000-2007 as provided by the Bureau of Economic Analysis,[16] excluding those financed by social insurance contributions (social security, unemployment benefits, medicare), education and training measures and benefits to recipients outside the US. For instance, the measure includes the earned income tax credit and various kinds of public assistance (medical assistance, medicaid, energy assistance etc.). Dividing these expenditures by total public expenditure,[17] we arrive at shares between 12 and 14 percent. We thus set $s^T = 0.13$ in our baseline scenario.

We now come to s_h^E and s_l^E. All US government bodies combined spent 13.6 percent of its total expenditure on education in the year 2011, including "public subsidies to households for living costs (scholarships and grants to students/households and students loans), which are not spent on educational institutions" (OECD, 2014, Tab. B4.1). We assume that education expenditure per head is the same for workers of type$-l$ (including public expenditure for specialized education programs for young adults which are typically more costly than other types of education) and type$-h$ (including public expenditure for tertiary education, which is about 25% of total public education expenditure in the year 2011, according to OECD, 2014). We thus set the government budget share of education for type$-l$ individuals, s_l^E, to $0.12 \times 0.136 \approx$

[15] Data is retrieved from the OECD tax database on May 11, 2015 (http://stats.oecd.org/index.aspx?DataSetCode=TABLE_I4). The respective tax rate at 133 and 167 percent of average labor income is the same.

[16] See www.bea.gov, National Data - National Income and Product Accounts Tables, Tab. 3.12 (Government Social Benefits), retrieved on May 12, 2015.

[17] See www.bea.gov, National Data - National Income and Product Accounts Tables, Tab. 3.1 (Government Current Receipts and Expenditures), retrieved on May 12, 2015.

0.016 and for type$-h$ individuals accordingly to $s_h^E = 0.88 \times 0.136 \approx 0.12$.

With a balanced government budget, the level of transfers on behalf of type$-l$ individuals adjusted for steady state growth, $\tilde{T} \equiv Te^{-gt}$, the fraction of human capital devoted to educating type$-l$ individuals, $h_l^E = h_l^E/h$, and the subsidy rate on education costs of type$-h$ individuals, ϑ, can be endogenously derived given the government expenditure shares s^T, s_l^E and s_h^E, respectively (see online-appendix). Our calibrated parameters imply an education subsidy rate for type$-h$ workers of $\vartheta = 0.29$, which seems reasonable.[18] In the US, the fraction of educational spending financed by public sources is 34.8 percent for tertiary education and 67.9 percent for all levels combined in the year 2011 (OECD, 2014). These figures do not account, however, for private opportunity costs of education which are inherently difficult to estimate.

5.2 Directly Observed Parameters

The per capita income growth rate (g), the population growth rate (n), the mark-up factor (κ), the elasticity of substitution between high-skilled and low-skilled labor (ψ) and the human capital depreciation rate (δ_H) are observed directly.

Recalling $g = \frac{n(1-\theta)}{1-\phi}$, we have $\phi = 1 - \frac{n(1-\theta)}{g}$. Consistent with average values for the period 1990-2006 (thereby averaging out business cycle phenomena) from the Penn World Tables (PWT) 7.1 (Heston, Summers and Aten, 2012), we let the long-run average per capita income growth rate of the US economy, g, be equal to two percent. The average annual population growth was about one percent. With $g = 0.02$ and $n = 0.01$, we have $\phi = 0.5(1 + \theta)$. Assuming an intermediate value $\theta = 0.5$, we arrive at $\phi = 0.75$. Our main conclusions are robust to variations in θ and ϕ which fulfill $\phi = 0.5(1 + \theta)$.

In his survey about skill-biased technological change, Johnson (1997) argues that the elasticity of substitution between high-skilled and low-skilled labor is about 1.5.

[18] According to (14), the education subsidy which offsets the long run distortion of labor income taxation of type$-h$ workers reads as $\vartheta = \tau_h$. Our calibrated levels of ϑ and τ_h thus suggest little distortion of educational choices in the US. For the long run fraction of human capital of the representative high-ability worker devoted to the education, our calibration in Tab. 1 implies $\mathfrak{h}_h^{E*} = 0.18$. Moreover, $\mathfrak{h}_l^E = 0.05$.

We thus take value $\psi = 1.5$ for our baseline calibration.[19] For the output elasticity of capital goods, we choose $\alpha = 0.4$. Our conclusions are rather insensitive to alternative values in the typical range. With $\psi = 1.5$, we obtain an elasticity of substitution between the inputs in final production of $\varepsilon = \frac{\psi-\alpha}{1-\alpha} = 1.83$.

For the mark-up factor on marginal costs of durable goods producers, κ, we take a typical value from the empirical literature, $\kappa = 1.3$ (Norrbin, 1993). The human capital depreciation rates are set within the range of the estimated value in Heckman (1976), who finds that human capital depreciates at a rate in the range between 0.7 and 4.7 percent. We assume $\delta_H = 0.03$.

5.3 Endogenous Observables

We need to calibrate either the long run value of one of the individual asset holdings, a_h^*, a_l^*, or their ratio, a_h^*/a_l^*. As shown in the online-appendix, the long run values of individual asset holdings and consumption levels are indeterminate, i.e. depend on initial values for the number of machines, $A_{H,0}$, $A_{L,0}$, and asset holdings $a_{h,0}$, $a_{l,0}$. Since the equity issued by machine producers to finance blueprints is included in asset holdings (see Appendix A), $A_{H,0}$, $A_{L,0}$, $a_{h,0}$, $a_{l,0}$ are not independent from each other. Thus, for instance, $A_{H,0}$, $A_{L,0}$ and the ratio $a_{h,0}/a_{l,0}$ give us $a_{h,0}$ and $a_{l,0}$. We thus choose a_h^*/a_l^* according to empirical evidence. Survey data for the year 2007 suggests that households headed by someone without a high school diploma (type$-l$ individuals) have, on average, a net worth of US\$ 150,000 (in 2010 dollars). Moreover, the average asset holding of educated households (type$-h$ individuals) is approximately US\$ 750,000.[20] We thus set $a_h^*/a_l^* = 5$.

Also the other parameters are matched to long run values of endogenous observables: the capital-output ratio (K/Y), the interest rate (r), the full-time equivalent of relative wage income of the different types of workers ("skill premium"), Ω, the R&D intensity,

[19] $\theta = 2\phi - 1$ and $\psi = 1.5$ jointly imply that (A3) holds with equality.

[20] Those headed by a college graduate possess about US\$ 1,15 million, whereas those households headed by high-school graduates and educated by some college possess about US\$ 264,000 and US\$ 384,000, respectively. See http://www.federalreserve.gov/econresdata/scf/scf_2010.htm. In the 2000s, about 40 percent of the 25-64 year olds in the US are tertiary-educated (OECD, 2014).

$R\&D$, and the rates of return to education for type$-h$ and type$-l$ individuals, R_h and R_l, respectively.

We assume that the long run interest rate is $r^* = 0.07$ (Mehra and Prescott, 1985). The Keynes-Ramsey rule for consumption growth implies $\sigma = \frac{(1-\tau_r)r^*-\rho}{g}$. Given a typical value for the time preference rate, $\rho = 0.02$, recalling $\tau_r = 0.17$ and $g = 0.02$, we find $\sigma = 1.91$.

Similarly to Grossmann, Steger and Trimborn (2013), we use measures for the investment rate (sav) and the capital-output ratio to calibrate the depreciation rate of physical capital, δ_K. The investment rate is $sav = \frac{\dot{K}+\delta_K K}{Y}$. Using $\dot{K}/K = n + g$ and solving for δ_K yields $\delta_K = \frac{sav}{K/Y} - n - g$. Averaging over the period 1990-2006 suggests that the long run investment rate, sav^*, is equal to about 21 percent, according to PWT 7.1. To find the long capital-output ratio, K^*/Y^*, we take averages over the period 2002-2007 calculated from data of the US Bureau of Economic Analysis. The capital stock is proxied by the amount of total fixed assets (private and public structures, equipment and software). At current prices, this suggests $K^*/Y^* = 3$. Thus, the evidence suggests that $\delta_K = 0.04$, which is a standard value in the literature.

Our calibration of the parameters characterizing the educational production processes, $\xi = 0.3$, $\beta = 1/3$, $\eta = 1/9$, are in line with empirical evidence on observables with theoretically derived long run values of the (pre-tax) skill premium,

$$\Omega \equiv \frac{w_h h}{w_l l}, \tag{17}$$

the R&D intensity (total wage costs for researchers per unit of final output),

$$R\&D \equiv \frac{w_h(H_H^A + H_L^A)}{Y}, \tag{18}$$

and the relative returns to education for type$-h$ and type$-l$ individuals. We define R_h and R_l as the internal rates of return for type$-h$ and type$-l$ individuals from permanently raising teaching inputs h_h^E and h_l^E by one unit, respectively, holding the wage rates w_l and w_h constant and starting in a BGE.

The theoretical long run values of (17), (18) and the internal education returns, Ω^*, $R\&D^*$, R_h^*, R_l^*, are derived in the online-appendix and used to set the remaining parameters. To calibrate Ω^*, we looked at the earnings distribution for those aged 25+ with at least high school diploma and without high school diploma. According to the Bureau of Labor Statistics (2015), the relative median earnings between the two groups is 1.9 and the relative earnings at the 90th percentile about 2.1. We would like to measure relative average earnings to proxy Ω^* which are not available, however. As the earnings dispersion is less pronounced within the group of high school dropouts, it is safe to choose a calibration in line with the notion that the value for relative average earnings is higher than relative median earnings. Our calibration suggests $\Omega^* = 2.06$, which appears reasonable. Moreover, for the R&D intensity, it implies $R\&D^*$ equal to 3.1 percent, which is the value suggested by OECD (2009) for the business R&D intensity (BERD as a percentage of value added in industry) in the US for the year 2007. Finally, we use the theoretically derived relative returns to education, R_h^*/R_l^*. The Mincerian rate of return to education (percentage change of wage income per additional year of schooling) is found to be higher for individuals with at least high-school education than for high-school drop-outs attending special education programs for adolescents and young adults (see e.g. the survey by Kautz et al., 2014). Our baseline calibration suggests $R_h^*/R_l^* = 3.7$, roughly in line with this literature.

It turns out that values of variables in BGE can be written as functions of $\bar{\nu} \equiv \nu(N_{h,0})^{1-\theta}$, where ν is the R&D productivity parameter. The endogenous observables are basically insensitive to changes in $\bar{\nu}$; we choose $\bar{\nu} = 0.2$.

6 Numerical Analysis and Trickle Down Dynamics

This section examines the dynamic implications of policy reforms on gross wage income of type$-l$ individuals, W_l, their consumption level, c_l, and their net total income level, y_l. Starting from a BGE for our baseline calibration, we consider changes in the rate at which education costs of type$-h$ individuals are subsidized, ϑ, in the fraction of human capital devoted to improve skills of type$-l$ individuals, $\mathfrak{h}_l^E$, and in (steady state growth-adjusted) transfers, $\tilde{T}$, triggered by comparable (one percentage point) increases in the government budget shares s_h^E, s_l^E and s^T, respectively. We first focus our discussion on financing rising public expenditure by an increase in the marginal labor income tax rate of high-skilled workers, τ_h, and discuss alternative financing schemes in the aftermath.

To be more precise, for instance, consider a certain percentage point increase in s_h^E which is financed by an increase in τ_h. We let the education subsidy rate ϑ adjust endogenously along with τ_h, such that (i) the government's budget remains balanced, (ii) the other tax rates as well as the other policy instruments which govern the dynamical system (in the example, $\mathfrak{h}_l^E$ and $\tilde{T}$) are held constant, and importantly, (iii) the steady state growth-adjusted level of public spending per capita of the fourth public spending category is held constant at the initial level (it thus continuous to grow at rate g also after a policy reform). The government budget shares (s_h^E, s_l^E, s^T) have been introduced to calibrate the policy instruments $(\vartheta, \mathfrak{h}_l^E, \tilde{T})$ as outlined in section 5 and also allow us to consider three different policy reforms which are comparable to each other. Without the fourth spending category, requirement (iii) would be superfluous (it would trivially hold because the residual expenditure would be zero). However, any reasonable calibration dictates $s^T + s_l^E + s_h^E < 1$.

We apply the relaxation algorithm (Trimborn, Koch and Steger, 2008) which is designed to deal with highly-dimensional and non-linear differential-algebraic equation systems. A favorable feature of the relaxation algorithm is that it does not rely on linearization of the underlying dynamic system. The differential-algebraic system is summarized in the online-appendix.

6.1 Endogenous Adjustment of Income Tax Rate for High-Earners

We first evaluate the dynamic effects of policy reforms under endogenous adjustment of the (marginal) tax rate on labor income of high-skilled workers, τ_h, to keep the government's budget balanced.

6.1.1 Expanding Higher Education

As displayed by the solid lines of Fig. 1-3, an increase in the subsidy rate for higher education, ϑ, leads to a drop on impact and further reduction of W_l/W_l^*, c_l/c_l^* and y_l/y_l^* early in the transition to the new BGE. This primarily reflects a reallocation of high-skilled labor away from manufacturing (decrease in h^X) on impact, lowering the price of the low-skilled intensive composite input, P_L^X.[21] The effect reflects the complementarity of both types of labor, as measured by the derived elasticity of substitution, ψ. It goes along with less R&D directed to machine development targeted to low-skilled labor. In sum, there is a reduction in the wage rate, w_l, relative to the one in the initial steady state. In the longer run, however, human capital of high-ability workers expands despite the distortionary effect of an increase in τ_h. Because of increased availability of teachers, the skill level of low-ability individuals, l, increases. This eventually fosters R&D effort directed to low-ability workers, raising their wage rate. After the initial drop, h^X increases over time, eventually beyond the initial level, in turn raising P_L^X and w_l. In sum, wage income W_l increases in the longer run beyond the initial level, consistent with Proposition 2. R&D-based growth directed to both composite goods sectors is fostered by the human capital formation of both types. Thus, also consumption and net income is boosted unambiguously in the longer run.

[21] In the online-appendix (Fig. A.2-A.4) we graphically display the transitional dynamics of all variables in response to policy reforms.

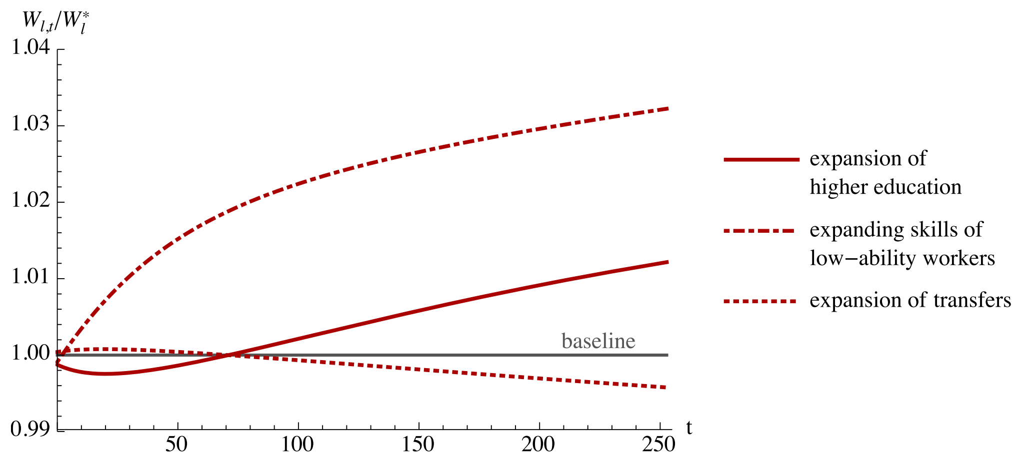

Figure 1: Time paths of normalized wage income of type-l individuals, W_l/W_l^*, in response to three policy reforms under endogenous adjustment of τ_h: Government budget shares s_h^E, s_l^E and s^T are raised by one percentage point to expand higher education (increase in ϑ), skills of low-ability workers (increase in h_l^E) and transfers towards low-skilled workers (increase in $\tilde{T}$), respectively. Set of parameters as in Table 1.

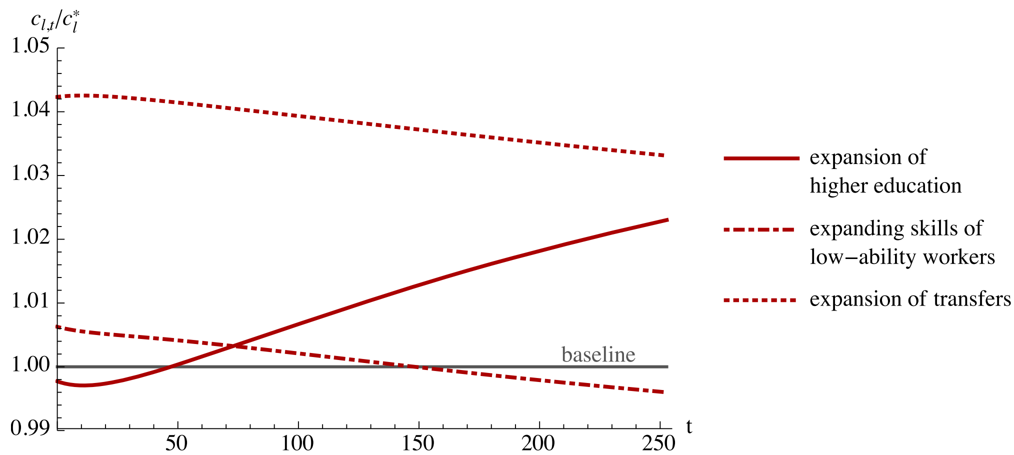

Figure 2: Time paths of normalized consumption of type-l individuals, c_l/c_l^*, in response to same policy reforms as in Fig. 1. Set of parameters as in Table 1.

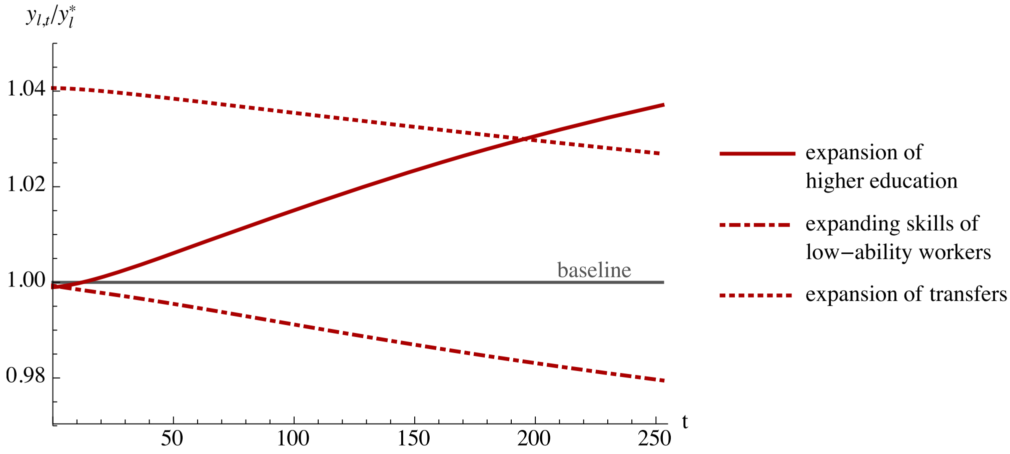

Figure 3: Time paths of normalized income of type-l individuals, y_l/y_l^*, in response to same policy reforms as in Fig. 1. Set of parameters as in Table 1.

6.1.2 Comparison to Direct Policy Measures On Behalf of Low-Skilled Workers

Now consider the dashed lines of Fig. 1-3. Expanding skills of type$-l$ individuals (increase in $\mathfrak{h}_l^E$ that raises l) unambiguously raises $W_l = w_l l$ early in the transition and increasingly so over time. This occurs although wage rate w_l declines along with an increasing output level of the low-skilled intensive composite good that lowers the composite input price P_L^X. This adverse general equilibrium effect is dampened by unskilled-biased technological change triggered off by the increase in skill level l. Using human capital of high-ability workers for teaching low-ability ones has rather high opportunity costs for the economy, however. First, given the model calibration, the return to education for low-ability types is relatively low (recall the discussion in section 5.3). Second, increased education costs are financed by an increase in distortionary tax rate τ_h, implying that human capital of type$-h$ workers, h, declines over time. In turn, R&D effort directed to complementing high-skilled labor is lowered and thus physical capital accumulation is depressed. That is, in the transition to the new BGE, low-ability workers choose to accumulate asset holdings, a_l, at a lower rate than g, implying that net income and consumption falls over time relative to initial levels.

From the set of policy reforms under consideration, raising transfers is most effective for boosting consumption of low-skilled workers at least for a long time after the policy reform (see the dotted line of Fig. 2). Because of the distortionary way transfers are financed, human capital of high-ability workers and therefore also of low-ability workers decline over time. Thus, the policy measure has general equilibrium effects on wage income that are eventually negative (slightly positive early in the transition, because on impact the reduced teaching input goes along with an increase in h^X). The human capital decumulation over time reduces R&D effort in the longer run (albeit not in the shorter run), in turn depressing capital accumulation. The adverse income effects occur rather late in the transition and are too small, however, to overturn the positive effect of transfers on net income, and therefore on consumption of low-skilled workers. From the viewpoint of low-skilled workers, the transfer policy dominates the policy to improve their skills and, at least for a relevant time horizon, also the policy of expanding higher education.[22] Our comparative policy conclusions are robust to alternative, reasonable parameter sets that are consistent with the observable data.

6.2 Trickle-down Growth From Higher Education: Discussion

Our analysis suggests that expanding higher education has a dismal consumption and income effect for low-ability workers early in the transition and an eventual trickle-down growth effect. It takes about five decades until consumption, c_l, becomes higher compared to the initial steady state level, c_l^*.

[22] In the online-appendix, to isolate the role of tax distortions, we display the implications of the three policy shocks for consumption of low-ability individuals, except that additional public spending is financed in a non-distortionary way at the expense of the fourth spending category (Fig. A.5). That is, all tax rates are kept constant. The dynamic effects of higher education expansion look rather similar to Fig. 2, whereas the other two policies become more beneficial: c_l is now raised through the entire transition also when expanding skills of low-ability individuals (increase in $\mathfrak{h}_l^E$); raising transfers boosts c_l the most also in the long run.

6.2.1 Robustness

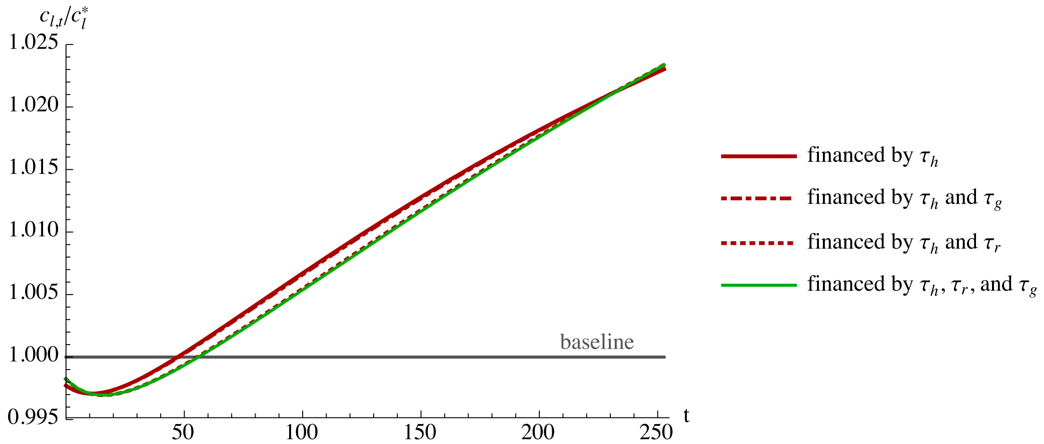

Figure 4: Time paths of normalized consumption of type-l individuals, c_l/c_l^*, under the policy reform "expanding higher education" (s_h^E is raised by one percentage point), assuming alternative tax rate adjustments. Set of parameters as in Table 1.

How robust is this result? First, we consider alternative ways how the increase in higher education subsidy rate, ϑ, is financed. Rather than solely adjusting labor income tax rate τ_h, consider an adjustment of τ_h along with an adjustment of the tax rate on capital income, τ_r, and the capital gains tax rate, τ_g, such that ratios τ_h/τ_r and τ_h/τ_g remain constant, respectively. Also consider an adjustment of τ_h along with both τ_r and the capital gains tax, τ_g, such that ratios of tax rates remain constant. In our model, taxation of both capital income and capital gains slows down the accumulation process of physical capital and knowledge capital, by giving disincentives for households to save and for firms to invest in R&D. Labor income taxation, however, has adverse growth effects by distorting human capital accumulation. As shown in Fig. 4, the considered alternatives to finance an increase in ϑ changes the evolution of consumption in a minor way.

Second, there is an intensive discussion on the returns to education particularly regarding programs aiming to improve skills of high-school drop-outs. It turns out that the curvature parameters β and γ, capturing the effectiveness of teaching high-ability and low-ability workers, according to (7) and (8), critically determine the relative education return, R_h/R_l, while playing little role for other endogenous observables. Fig. 5 displays that varying R_h^*/R_l^* by changing β (increase to 0.35 to match $R_h^*/R_l^* = 5.4$) and γ (increase to 0.15 to match $R_h^*/R_l^* = 1.5$) has little effect on the time elapsing until an increase in ϑ becomes beneficial for low-skilled workers.

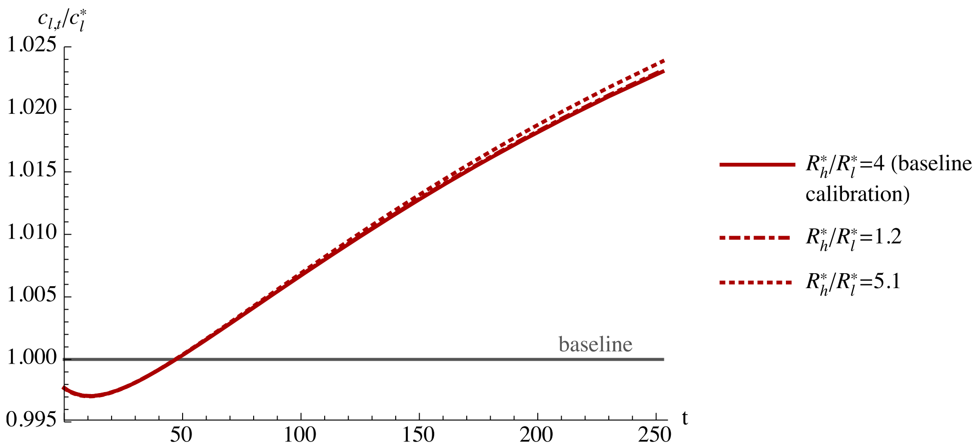

Figure 5: Time paths of normalized consumption of type-l individuals, c_l/c_l^*, under policy reform "expanding higher education" (s_h^E is raised by one percentage point), assuming endogenous adjustment of τ_h and alternative relative rates of return to education. Set of parameters as in Table 1.

6.2.2 Wage Inequality

Is the preceding dynamic policy evaluation consistent with rising skill premia along with expansion of higher education, as observed in many advanced countries? Fig. 6 (a) and Fig. 6 (b) display the impact of an increase in ϑ on the skill premium, Ω, and the relative skill level, h/l, under the alternative ways of tax rate adjustments as for Fig. 4. Acemoglu (2002) has analyzed the impact of an *exogenous* increase in the supply ratio of high-skilled to low-skilled labor on the long run skill premium. His analysis requires that the (derived) elasticity of skilled labor and unskilled labor must be larger than two to explain a rising skill premium, which is higher than most empirical studies indicate. Our model, in contrast, suggests a gradual increase in Ω during the transition towards the steady state along with an endogenous increase in the relative skill level triggered off by higher education expansion for the empirically plausible calibration $\psi = 1.5$.[23]

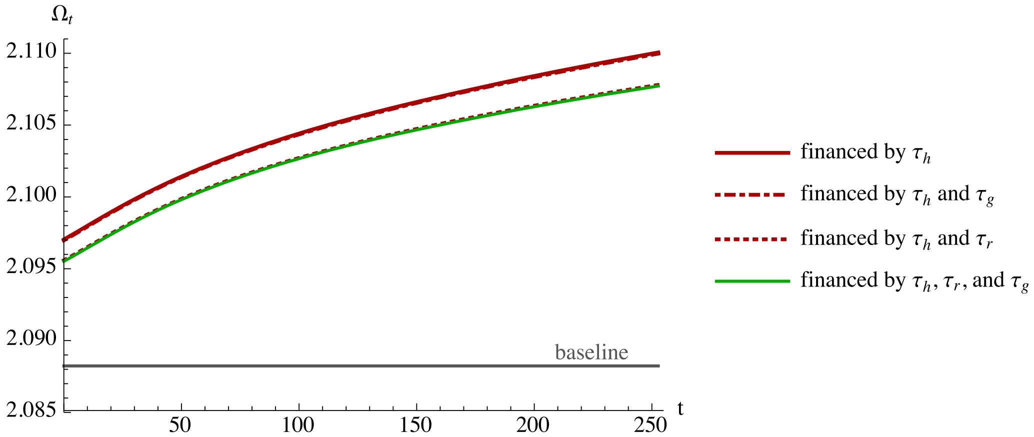

Figure 6 (a): Time paths of the skill premium $\Omega = w_h h / w_l l$, under the policy reform "expanding higher education" (s_h^E is raised by one percentage point), assuming alternative tax rate adjustments. Set of parameters as in Table 1.

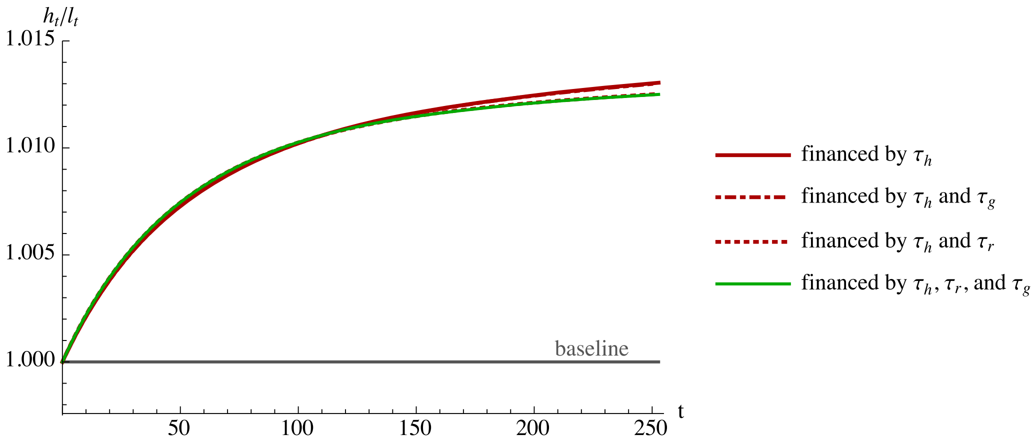

Figure 6 (b): Time paths of the skill ratio, h/l, under the policy reform "expanding higher education" (s_h^E is raised by one percentage point), assuming alternative tax rate adjustments. Set of parameters as in Table 1.

[23] As in our baseline calibration part (i) of Proposition 3 applies, the change in the long run skill premium, Ω^*, is entirely driven by the change in h^*/l^*.

7 Conclusion

The first goal of this paper was to understand whether and, if so, when economic growth caused by an increase in public education expenditure on behalf of high-ability individuals trickles down to low-ability workers who do not acquire higher education. We contrasted the dynamic effects of higher education expansion with those of an equally sized increase in redistributive transfers and of skill formation targeted to low-ability workers. In our dynamic general equilibrium framework, (changes in) public expenditures are financed by (changes in) various distortionary income taxes, human capital accumulation is endogenous, and R&D-based technical change could be directed to complement high-skilled or low-skilled labor (or both).

In the shorter run, the poor lose from expanding higher education for an extended period relative to the status quo. Consistent with empirical evidence for the US from the 1970s onwards, our analysis suggests that human capital accumulation is accompanied by falling or stagnating earnings of low-skilled individuals early in the transition phase and rising skill premia. In the longer run, however, low-ability workers benefit from promoting education of high-ability workers. The trickle-down effect is driven by the (static) complementarity of different types of human capital in goods production, higher availability of potential teachers for educating low-ability individuals and an eventual increase in the level of human capital devoted towards R&D for producing low-skilled labor intensive goods. At least for a relevant time horizon, low-ability workers are worse off than under the transfer policy, however.

Skill promotion targeted to low-ability workers is more effective than higher education expansion to raise their wage income also in the longer run and triggers off unskilled-biased technological change. However, the policy may be economically costly for two reasons. First, when the internal rate of education return is relatively low for low-ability workers, there are high opportunity costs of allocating high-skilled workers to teach low-ability ones. Second, the necessary increase in tax rates for financing such education is distortionary. Thus, the policy effects are growth-reducing during the transition to the new balanced growth equilibrium. As a result, although the policy reform allows low-skilled workers to raise consumption in the short run, it raises their well-being less than increasing transfers through the entire transition.

As a caveat, although our analysis suggests that moderately increasing transfers works best to raise well-being of low-ability workers, there are obvious limits to redistribution resulting from growth-reducing tax distortions to finance them. Moreover, while demonstrating that the evaluation of skill formation programs on behalf of the socially disadvantaged should account for general equilibrium effects, its limited role to raise living standards of the poor critically hinges on exclusion of social mobility in the model. The modeling choice served to highlight the comparative role of policy options if potentially high-ability children from disadvantaged households do not benefit from interventions preventing them to end up as high-school drop-outs or similar. This is, unfortunately, not unrealistic for the time-being. It would be interesting, however, to direct future research on education programs (possibly early in the childhood) which promotes social mobility and comparatively dynamically evaluate them vis-à-vis other policy interventions from a general equilibrium perspective.

Appendix

Appendix A. Definition of Equilibrium

Denote sizes of "type$-h$" and "type$-l$" households by $N_h > 0$ and $N_l > 0$, respectively. Let P_H^X and P_L^X denote the price of the high-skilled intensive and low-skilled intensive composite intermediate good used in the final goods sector, respectively, and $p_H(i)$, $p_L(i)$ the prices of machine i in the respective composite input sector. Moreover, let P_H^A and P_L^A denote the present discounted value of the profit stream generated by an innovation in the low-skilled and high-skilled intensive sector, respectively. These are equal to equity prices. There are no arbitrage possibilities in the financial market; thus, the after-tax returns from equity (capital gains and dividends) in both sectors and bonds and must be equal:

$$(1-\tau_g)\frac{\dot{P}_H^A}{P_H^A}+(1-\tau_r)\frac{\pi_H}{P_H^A}=(1-\tau_g)\frac{\dot{P}_L^A}{P_L^A}+(1-\tau_r)\frac{\pi_L}{P_L^A}=(1-\tau_r)r. \tag{19}$$

For given policy parameters $(\tau_g, \tau_r, \tau_h, \tau_l, T, \mathfrak{h}_l^E, \vartheta)$, an equilibrium consists of time paths for quantities $\{H_t^X, L_t^X, H_{Ht}^A,$

$H_{Lt}^A, h_t, h_{ht}^E, X_{Ht}, X_{Lt}, \{x_{Ht}(i)\}_{i\in[0,A_{Ht}]}, \{x_{Lt}(i)\}_{i\in[0,A_{Lt}]}, A_{Ht}, A_{Lt}, c_{ht}, c_{lt}, a_{ht}, a_{lt}\}$ and prices $\{P_{Ht}^X, P_{Lt}^X, \{p_{Ht}(i)\}_{i\in[0,A_{Ht}]}, \{p_{Lt}(i)\}_{i\in[0,A_{Lt}]}, P_{Ht}^A, P_{Lt}^A, w_{ht}, w_{lt}, r_t\}$ such that

1. R&D firms and producers of the final good, the composite intermediate goods, and machines maximize profits, taking prices as given;[24]

2. taking factor prices as given, type$-h$ households choose the consumption path $\{c_{ht}\}_{t=0}^{\infty}$ and teaching inputs $\{h_{ht}^E\}_{t=0}^{\infty}$ to maximize utility U_h s.t. (7) and (10); type$-l$ households choose the consumption path $\{c_{lt}\}_{t=0}^{\infty}$ to maximize U_l s.t. (11);[25]

3. the no-arbitrage conditions (19) in the financial market hold;

4. the total value of assets (owned by households) fulfills

$$N_h a_h + N_l a_l = K + P_H^A A_H + P_L^A A_L, \tag{20}$$

where K is given by (4).

5. the labor markets for type–h and type–l workers clear:

$$\begin{aligned} H^X + H_H^A + H_L^A + N_h h_h^E + N_l h_l^E &= N_h h, && (21)\\ L^X &= N_l l. && (22)\end{aligned}$$

Appendix B. Proofs

Proof of Lemma 1. According to (1), inverse demand functions in the composite

[24] Condition 1 implies that the composite intermediate goods markets and the market for machines clear.

[25] Households also observe standard non-negativity constraints which lead to transversality conditions (see the proof of Proposition 1).

input sectors are given by

$$P_H^X = \frac{\partial Y}{\partial X_H} = \left(\frac{Y}{X_H}\right)^{\frac{1}{\varepsilon}}, \; P_L^X = \frac{\partial Y}{\partial X_L} = \left(\frac{Y}{X_L}\right)^{\frac{1}{\varepsilon}}. \tag{23}$$

Thus, relative intermediate goods demand is given by

$$\frac{X_H}{X_L} = \left(\frac{P_H^X}{P_L^X}\right)^{-\varepsilon} \tag{24}$$

According to (3), the inverse demand for machine i in the human capital intensive sector is $p_H(i) = \alpha P_H^X (H^X/x_H(i))^{\alpha-1}$. Machine producers, being able to transform one unit of the final good to one unit of output, have marginal production costs equal to the sum of the interest rate and the capital depreciation rate, $r + \delta_K$. In absence of a competitive fringe, the incumbent's profit-maximizing price would be $(r + \delta_K)/\alpha$. A price equal to $\kappa(r + \delta_K)$ (the marginal cost of the competitive fringe) is the maximal price, however, a producer can set without losing the entire demand. Since $\kappa \leq 1/\alpha$, it is also the optimal price. Thus, with $p_H(i) = p_L(i) = \kappa(r + \delta_K)$ for all i,

$$x_H(i) = x_H = \left(\frac{\alpha P_H^X}{\kappa(r + \delta_K)}\right)^{\frac{1}{1-\alpha}} H^X \overset{(2)}{\Longrightarrow} X_H = A_H H^X \left(\frac{\alpha P_H^X}{\kappa(r + \delta_K)}\right)^{\frac{\alpha}{1-\alpha}}, \tag{25}$$

$$x_L(i) = x_L = \left(\frac{\alpha P_L^X}{\kappa(r + \delta_K)}\right)^{\frac{1}{1-\alpha}} L^X \overset{(3)}{\Longrightarrow} X_L = A_L L^X \left(\frac{\alpha P_L^X}{\kappa(r + \delta_K)}\right)^{\frac{\alpha}{1-\alpha}}, \tag{26}$$

Hence, relative supply of composite inputs is

$$\frac{X_H}{X_L} = \frac{A_H H^X}{A_L L^X} \left(\frac{P_H^X}{P_L^X}\right)^{\frac{\alpha}{1-\alpha}}. \tag{27}$$

Equating the right-hand sides of (24) and (27) and using $\psi = \alpha + \varepsilon(1 - \alpha)$ leads to an expression for the relative price of the composite inputs,

$$P \equiv \frac{P_H^X}{P_L^X} = \left(\frac{A_H H^X}{A_L L^X}\right)^{-\frac{1-\alpha}{\psi}}, \tag{28}$$

which is inversely related to the relative "efficiency units" of high-skilled to low-skilled labor in production activities, $\frac{A_H H^X}{A_L L^X}$.

According to (2) and (3), wage rates per unit of high-skilled and low-skilled labor are given by $w_h = P_H^X(1-\alpha)X_H/H^X$ and $w_l = P_L^X(1-\alpha)X_L/L^X$, respectively. Dividing both equations and using both (27) and (28) confirms (12). ■

Proof of Lemma 2: According to (25) and (26), the instantaneous profits of machine producers, $\pi_H = (\kappa-1)(r+\delta_K)x_H$ and $\pi_L = (\kappa-1)(r+\delta_K)x_L$, read as

$$\pi_H = (\kappa-1)\left(\frac{\alpha}{\kappa}P_H^X\right)^{\frac{1}{1-\alpha}}(r+\delta_K)^{-\frac{\alpha}{1-\alpha}}H^X, \tag{29}$$

$$\pi_L = (\kappa-1)\left(\frac{\alpha}{\kappa}P_L^X\right)^{\frac{1}{1-\alpha}}(r+\delta_K)^{-\frac{\alpha}{1-\alpha}}L^X. \tag{30}$$

Dividing both expressions, substituting (28) and noting from the definition of ψ that $\frac{\alpha}{1-\alpha} = \frac{\varepsilon-\psi}{\psi-1}$ confirms (13). ■

Proof of Proposition 1: First, we define $l^X \equiv L^X/N_h$, $h^X \equiv H^X/N_h$, and the relative population size $\chi \equiv N_l/N_h$. We also define $h_k^A \equiv H_k^A/N_h$, $p_k^A \equiv P_k^A/N_h$, $k \in \{H, L\}$. With these definitions we can rewrite labor market clearing conditions (21) and (22) as

$$\begin{aligned} h^X + h_H^A + h_L^A + h_h^E + \chi h_l^E &= h, & (31)\\ l^X &= \chi l. & (32) \end{aligned}$$

Moreover, let $\tilde{z}_t \equiv z_t e^{-gt}$ for $z \in \{T, c_h, c_l, a_h, a_l, w_h, w_l, A_h, A_l\}$. That is, if a variable z grows with rate g in the long run, then $\tilde{z}$ is stationary. Combining (4) and (20) and substituting both (25) and (26), we then have

$$\begin{aligned} \tilde{a}_h + \chi\tilde{a}_l \;=\; & \tilde{A}_H\left(\frac{\alpha}{\kappa(r+\delta_K)}P_H^X\right)^{\frac{1}{1-\alpha}}h^X + \\ & \tilde{A}_L\left(\frac{\alpha}{\kappa(r+\delta_K)}P_L^X\right)^{\frac{1}{1-\alpha}}l^X + p_H^A\tilde{A}_H + p_L^A\tilde{A}_L. \qquad (33) \end{aligned}$$

The representative R&D firm which directs R&D effort to the human capital intensive sector maximizes

$$P_H^A \dot{A}_H - w_h H_H^A = P_H^A \tilde{\nu}_H (A_H)^{\phi} H_H^A - w_h H_H^A, \tag{34}$$

taking A_H and $\tilde{\nu}_H$ as given. Analogously for the R&D sector targeted to machines which are complementary to low-skilled labor. Thus, using (5) and (6), we have

$$P_H^A \nu \left(A_H\right)^{\phi} (H_H^A)^{-\theta} = P_L^A \nu \left(A_L\right)^{\phi} (H_L^A)^{-\theta} = w_h. \tag{35}$$

Define $\bar{\nu} \equiv \nu(N_{h,0})^{1-\theta}$ and recall $g = \frac{(1-\theta)n}{1-\phi}$. According to (35), we can then write

$$p_H^A \bar{\nu} \left(\tilde{A}_H\right)^{\phi-1} (h_H^A)^{-\theta} = \frac{\tilde{w}_h}{\tilde{A}_H}, \tag{36}$$

$$p_L^A \bar{\nu} \left(\tilde{A}_L\right)^{\phi-1} (h_L^A)^{-\theta} = \frac{\tilde{w}_h}{\tilde{A}_L}. \tag{37}$$

We turn next to composite input prices. Combining (23) with (1) implies

$$P_H^X = \left[1 + \left(\frac{X_H}{X_L}\right)^{-\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{1}{\varepsilon-1}}, \tag{38}$$

$$P_L^X = \left[1 + \left(\frac{X_H}{X_L}\right)^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{1}{\varepsilon-1}}. \tag{39}$$

Substituting (28) into (24) we find

$$\frac{X_H}{X_L} = \left(\frac{A_H H^X}{A_L L^X}\right)^{\frac{\varepsilon(1-\alpha)}{\varepsilon(1-\alpha)+\alpha}}. \tag{40}$$

Substituting (40) into (38) and (39), and using $A_H/A_L = \tilde{A}_H/\tilde{A}_L$, $H^X/L^X = h^X/l^X$

and $\psi = \varepsilon(1-\alpha)+\alpha$, we obtain

$$P_H^X = \left[1+\left(\frac{\tilde{A}_H h^X}{\tilde{A}_L l^X}\right)^{-\frac{\psi-1}{\psi}}\right]^{\frac{1}{\varepsilon-1}}, \tag{41}$$

$$P_L^X = \left[1+\left(\frac{\tilde{A}_H h^X}{\tilde{A}_L l^X}\right)^{\frac{\psi-1}{\psi}}\right]^{\frac{1}{\varepsilon-1}}. \tag{42}$$

The current-value Hamiltonian which corresponds to the optimization problem of a type$-h$ household (see Definition 1) is given by

$$\begin{aligned}\mathcal{H}_h = {} & \frac{(c_h)^{1-\sigma}-1}{1-\sigma} + \mu\left(\xi(h_h^E)^\beta h^\eta - \delta_H h\right) + \\ & \lambda\left([(1-\tau_r)r-n]a_h + (1-\tau_h)w_h h - (1-\vartheta)w_h h_h^E - c_h\right),\end{aligned} \tag{43}$$

where μ and λ are multipliers (co-state variables) associated with constraints (7) and (10), respectively. Necessary optimality conditions are $\partial\mathcal{H}_h/\partial c_h = \partial\mathcal{H}_h/\partial h_h^E = 0$ (control variables), $\dot{\mu} = (\rho-n)\mu - \partial\mathcal{H}_h/\partial h$, $\dot{\lambda} = (\rho-n)\lambda - \partial\mathcal{H}_h/\partial a_h$ (state variables), and the corresponding transversality conditions. Thus,

$$\lambda = (c_h)^{-\sigma}, \tag{44}$$

$$\mu\beta\xi(h_h^E)^{\beta-1}h^\eta = \lambda(1-\vartheta)w_h, \tag{45}$$

$$\frac{\dot{\mu}}{\mu} = \rho - n - \eta\xi_h\left(h_h^E\right)^\beta h^{\eta-1} + \delta_H - \frac{\lambda}{\mu}(1-\tau_h)w_h, \tag{46}$$

$$\frac{\dot{\lambda}}{\lambda} = \rho - (1-\tau_r)r, \tag{47}$$

$$\lim_{t\to\infty}\mu_t e^{-(\rho-n)t}h_t = 0, \tag{48}$$

$$\lim_{t\to\infty}\lambda_t e^{-(\rho-n)t}a_{ht} = 0. \tag{49}$$

Differentiating (44) with respect to time and using (47) as well as $\tilde{c}_h = c_h e^{-gt}$, we obtain Euler equation

$$\frac{\dot{\tilde{c}}_h}{\tilde{c}_h} = \frac{(1-\tau_r)r-\rho}{\sigma} - g. \tag{50}$$

Define $m_t \equiv \mu_t e^{(\sigma-1)gt}$ and $\mathfrak{h}_h^E \equiv h_h^E/h$. Combining (44) and (45) we can then write

$$m\xi\beta(\mathfrak{h}_h^E)^{\beta-1}h^{\eta+\beta-1} = (\tilde{c}_h)^{-\sigma}(1-\vartheta)\tilde{w}_h \tag{51}$$

(recall $\tilde{w}_h = w_h e^{-gt}$). Moreover, combining (45) and (46) and making use of (44) and (51),

$$\frac{\dot{m}}{m} = \delta_H + \rho - n + (\sigma-1)g - \left(\eta\mathfrak{h}_h^E + \beta\frac{1-\tau_h}{1-\vartheta}\right)\xi(\mathfrak{h}_h^E)^{\beta-1}h^{\eta+\beta-1}. \tag{52}$$

Moreover, (10) can be written as

$$\frac{\dot{\tilde{a}}_h}{\tilde{a}_h} = (1-\tau_r)r - n + (1-\tau_h)\frac{\tilde{w}_h h}{\tilde{a}_h} - (1-\vartheta)\frac{\tilde{w}_h h_h^E}{\tilde{a}_h} - \frac{\tilde{c}_h}{\tilde{a}_h} - g. \tag{53}$$

For low-skilled individuals (who decide about their consumption profile only), we find analogously to (50) that

$$\frac{\dot{\tilde{c}}_l}{\tilde{c}_l} = \frac{(1-\tau_r)r-\rho}{\sigma} - g. \tag{54}$$

By using (11) we also obtain

$$\frac{\dot{\tilde{a}}_l}{\tilde{a}_l} = (1-\tau_r)r - n + (1-\tau_l)\frac{\tilde{w}_l l}{\tilde{a}_l} - \frac{\tilde{c}_l}{\tilde{a}_l} + \frac{\tilde{T}}{\tilde{a}_l} - g. \tag{55}$$

Using $A_k = \tilde{A}_k e^{gt}$, $H_k^A = N_h h_k^A$, $k \in \{H, L\}$, as well as $N_{h,t} = N_{h,0}e^{nt}$, $\bar{\nu} = \nu(N_{h,0})^{1-\theta}$ and $g = \frac{(1-\theta)n}{1-\phi}$ we can rewrite (5) and (6) as

$$\frac{\dot{\tilde{A}}_H}{\tilde{A}_H} = \bar{\nu}(\tilde{A}_H)^{\phi-1}\left(h_H^A\right)^{1-\theta} - g, \tag{56}$$

$$\frac{\dot{\tilde{A}}_L}{\tilde{A}_L} = \bar{\nu}(\tilde{A}_L)^{\phi-1}(h_L^A)^{1-\theta} - g. \tag{57}$$

Recall that competitive wage rates read as $w_h = P_H^X(1-\alpha)X_H/H^X$ and $w_l = P_L^X(1-\alpha)X_L/L^X$. Combining these expressions with (25) and (26), respectively, we find for adjusted wage rates:

$$\tilde{w}_h = (1-\alpha)\left(\frac{\alpha}{\kappa(r+\delta_K)}\right)^{\frac{\alpha}{1-\alpha}} \tilde{A}_H \left(P_H^X\right)^{\frac{1}{1-\alpha}}, \tag{58}$$

$$\tilde{w}_l = (1-\alpha)\left(\frac{\alpha}{\kappa(r+\delta_K)}\right)^{\frac{\alpha}{1-\alpha}} \tilde{A}_L \left(P_L^X\right)^{\frac{1}{1-\alpha}}. \tag{59}$$

Substituting (29) and (30) into (19) implies

$$\dot{p}_H^A + np_H^A = \frac{1-\tau_r}{1-\tau_g}\left(rp_H^A - \frac{(\kappa-1)\left(\frac{\alpha}{\kappa}P_H^X\right)^{\frac{1}{1-\alpha}} h^X}{(r+\delta_K)^{\frac{\alpha}{1-\alpha}}}\right), \tag{60}$$

$$\dot{p}_L^A + np_L^A = \frac{1-\tau_r}{1-\tau_g}\left(rp_L^A - \frac{(\kappa-1)\left(\frac{\alpha}{\kappa}P_L^X\right)^{\frac{1}{1-\alpha}} l^X}{(r+\delta_K)^{\frac{\alpha}{1-\alpha}}}\right). \tag{61}$$

In sum, the dynamical system is given by (7), (31)-(33), (36), (37), (41), (42) and (50)-(61).

To prove that a steady state with the properties stated in Proposition 1 exists, we need to show that h_k^A, p_k^A, P_k^X, $\tilde{A}_k$ ($k \in \{H, L\}$), l^X, h^X, h_h^E, h, $\tilde{T}$, $\tilde{c}_j$, $\tilde{a}_j$, $\tilde{w}_j$ ($j \in \{l, h\}$), r and m are stationary in the long run. To see this, we next derive steady state values of the just derived dynamical system.

First, set $\dot{h} = 0$ and use $h_h^E = \mathfrak{h}_h^E h$ in (7) to find $\xi(\mathfrak{h}_h^E)^\beta h^{\beta+\eta-1} = \delta_H$, which can also be rewritten as

$$h = \left[\frac{\xi(\mathfrak{h}_h^E)^\beta}{\delta_H}\right]^{\frac{1}{1-\beta-\eta}} \equiv \bar{h}(\mathfrak{h}_h^E). \tag{62}$$

Using (62) in (52) and setting $\dot{m} = 0$ confirms (14). Evaluating (62) at $\mathfrak{h}_h^E = \mathfrak{h}_h^{E*}$ then gives us $h^* = \bar{h}(\mathfrak{h}_h^{E*})$, as stated in (15). This confirms part (v). Note that h^* and $\mathfrak{h}_h^{E*}$ are indeed time-invariant (i.e., $\dot{h} = 0$ for $t \to \infty$), as claimed in part (iv). Moreover, the long run teaching input for educating type$-h$ workers, $h_h^{E*} \equiv \mathfrak{h}_h^{E*} h^*$ reads as

$$h_h^{E*} = \left[\frac{\xi(\mathfrak{h}_h^{E*})^{1-\eta}}{\delta_H}\right]^{\frac{1}{1-\beta-\eta}} \equiv \bar{h}_h^E(\mathfrak{h}_h^{E*}), \tag{63}$$

according to (14) and (15).

Setting $\dot{l} = 0$ in (8) and using $h_l^E = \mathfrak{h}_l^E h^*$ confirms (16) in part (vi). Substituting $h^* = \bar{h}(\mathfrak{h}_h^{E*})$ into (16) implies that l^*, can be written as

$$l^* = \left(\frac{\xi(\mathfrak{h}_l^E)^\gamma (\bar{h}(\mathfrak{h}_h^{E*}))^\gamma}{\delta_H} \right)^{\frac{1}{1-\eta}} \equiv \bar{l}(\mathfrak{h}_l^E, \mathfrak{h}_h^{E*}). \tag{64}$$

Next, set $\dot{\tilde{c}}_h = 0$ in (50) to find that the long run interest rate, r^*, is given by

$$r^* = \frac{\rho + \sigma g}{1 - \tau_r}. \tag{65}$$

Thus, also $\dot{\tilde{c}}_l = 0$ holds, according to (54). Next, set $\dot{\tilde{A}}_H = \dot{\tilde{A}}_L = 0$ in (56) and (57) to obtain

$$\tilde{A}_H = \left(\frac{\bar{\nu} \left(h_H^A\right)^{1-\theta}}{g} \right)^{\frac{1}{1-\phi}}, \tag{66}$$

$$\tilde{A}_L = \left(\frac{\bar{\nu} \left(h_L^A\right)^{1-\theta}}{g} \right)^{\frac{1}{1-\phi}}, \tag{67}$$

respectively, or $\bar{\nu}(\tilde{A}_k)^{\phi-1} = \left(h_k^A\right)^{\theta-1} g$, $k \in \{H, L\}$. Using the latter together with (58) in (36) and (37) yields

$$p_H^A = (1-\alpha) \left(\frac{\alpha}{\kappa(r+\delta_K)} \right)^{\frac{\alpha}{1-\alpha}} \left(P_H^X\right)^{\frac{1}{1-\alpha}} \frac{h_H^A}{g}, \tag{68}$$

$$p_L^A = (1-\alpha) \left(\frac{\alpha}{\kappa(r+\delta_K)} \right)^{\frac{\alpha}{1-\alpha}} \left(P_H^X\right)^{\frac{1}{1-\alpha}} \frac{h_L^A}{g} \frac{\tilde{A}_H}{\tilde{A}_L}, \tag{69}$$

respectively. Now substitute (65) and (68) into (60) and set $\dot{p}_H^A = 0$ to find

$$h_H^A = \Gamma(\tau_r, \tau_g) h^X, \tag{70}$$

where

$$\Gamma(\tau_r, \tau_g) \equiv \frac{1 - \frac{1}{\kappa}}{\frac{1}{\alpha} - 1} \frac{(1-\tau_r) g}{\rho + \sigma g - (1-\tau_g) n}. \tag{71}$$

Note that $\Gamma > 0$ under (A1). Similarly, substituting (65) and (69) into (61) and setting $\dot{p}_L^A = 0$ we obtain

$$h_L^A = \frac{\Gamma(\tau_r, \tau_g) l^X}{P^{\frac{1}{1-\alpha}} \frac{\tilde{A}_H}{\tilde{A}_L}}. \tag{72}$$

From (70) and (72) we get

$$\frac{h_H^A}{h_L^A} = \frac{\tilde{A}_H}{\tilde{A}_L} \frac{h^X}{l^X} P^{\frac{1}{1-\alpha}}. \tag{73}$$

Moreover, (66) and (67) imply that

$$\frac{\tilde{A}_H}{\tilde{A}_L} = \left(\frac{h_H^A}{h_L^A}\right)^{\frac{1-\theta}{1-\phi}} = \left(\frac{h^X}{l^X} P^{\frac{1}{1-\alpha}}\right)^{-\frac{1-\theta}{\phi-\theta}}, \tag{74}$$

where the latter equation follows after substituting (73).

Next, substitute $A_H/A_L = \tilde{A}_H/\tilde{A}_L$ as given by (74) into (28), and use $H^X/L^X = h^X/l^X$ to obtain

$$P^{\frac{1}{1-\alpha}} = \left(\frac{h^X}{l^X}\right)^{-\frac{1-\phi}{1-\theta-\psi(\phi-\theta)}}. \tag{75}$$

According to assumption (A2), $\psi > 1$ and $2 - \phi - \psi + \theta(\psi - 1) = 1 - \theta - \psi(\phi - \theta) - (\psi - 1)(1 - \phi) > 0$, implying $1 - \theta - \psi(\phi - \theta) > 0$. Hence, the relative composite input price, P, and relative employment in input production, h^X/l^X, are negatively related under (A2).

Substituting (74) and into (72) and using (75) then leads to

$$h_L^A = \Gamma(\tau_r, \tau_g) \left(h^X\right)^{\varrho} (l^X)^{1-\varrho}, \text{ where} \tag{76}$$

$$\varrho \equiv \frac{2 - \phi - \psi + \theta(\psi - 1)}{1 - \theta - \psi(\phi - \theta)} \tag{77}$$

(thus, $1 - \varrho = \frac{(\psi-1)(1-\phi)}{1-\theta-\psi(\phi-\theta)}$). According to assumption (A2), $0 \leq \varrho < 1$.

Using (70), (76) and $h_l^E = \mathfrak{h}_l^E h$ in (31), and substituting $h_h^E = \bar{h}_h^E(\mathfrak{h}_h^{E*})$, $h = \bar{h}(\mathfrak{h}_h^{E*})$ and, from (32) and (64), $l^X = \chi \bar{l}(\mathfrak{h}_l^E, \mathfrak{h}_h^{E*})$, the long run value of h^X is implicitly defined as

$$[1+\Gamma(\tau_r,\tau_g)]\,h^X = (1-\chi\mathfrak{h}_l^E)\bar{h}(\mathfrak{h}_h^{E*}) - \bar{h}_h^E(\mathfrak{h}_h^{E*}) - \Gamma(\tau_r,\tau_g)\left(h^X\right)^{\varrho}(\chi\bar{l}^*(\mathfrak{h}_l^E,\mathfrak{h}_h^{E*}))^{1-\varrho}. \tag{78}$$

We write $h^{X*} = \bar{h}^X(\tau_r,\tau_g,\mathfrak{h}_l^E,\mathfrak{h}_h^{E*})$. The left-hand side of (78) as a function of h^X is an increasing line through the origin. For $h^X = 0$, the right-hand side of (78) is positive. If $\varrho > 0$, it is monotonically decreasing in h^X and eventually becomes negative. If $\varrho = 0$, it is a (positive) constant. Thus, whenever $\varrho \geq 0$, h^{X*} is unique.[26] To prove part (vii), first note that $\Gamma(\tau_r,\tau_g)$ is decreasing in both τ_r and τ_g, according to (71). Moreover, the right-hand side of (78) is decreasing in $\mathfrak{h}_l^E$ (as $\bar{l}^*$ is increasing in $\mathfrak{h}_l^E$). Finally, policy instruments τ_h and ϑ enter (78) only through their impact of $\mathfrak{h}_h^{E*}$. An increase in $\mathfrak{h}_h^{E*}$, however, has an ambiguous effect on the right-hand side of (78).

Remark: *If $\varrho < 0$, meaning that assumption (A2) is violated, the right-hand side of (78) is strictly increasing and concave in h^X, goes to $-\infty$ for $h^X \to 0$ and approaches a strictly positive value for $h^X \to \infty$. Thus, in this case, either two solutions or no interior solution for h^X as given by (78) exist. Two solutions means that two interior BGE exist, one stable and one unstable (see online-appendix for a numerical example). If no solution to (78) exists, then $h^X = 0$ would hold in BGE.*

It is easy to check that (25), (26), (31), (41), (42), (50), (52), (53), (54), (55), (58), (59), (62)-(70), (76) and (78) are consistent with parts (i)-(iv) of Proposition 1.

Finally, it remains to be shown that the transversality conditions (48) and (49) hold under assumption (A1). Differentiating (45) with respect to time and using that $\dot{h} = 0$ as well as $\dot{w}_h/w_h = g$ for $t \to \infty$ implies that, along a balanced growth path, $\dot{\mu}/\mu = \dot{\lambda}/\lambda + g$. From (44) and $\dot{c}_h/c_h = g$ for $t \to \infty$ we find $\dot{\lambda}/\lambda = -\sigma g$ and thus $\dot{\mu}/\mu = (1-\sigma)g$. As h becomes stationary, (48) holds iff $\lim_{t\to\infty} e^{[(1-\sigma)g+n-\rho]t} = 0$, i.e., iff (A1) holds. Similarly, using $\dot{\lambda}/\lambda = -\sigma g$ and the fact that a_h grows with rate g in the long run, we find that also (49) holds under (A1). The same is analogously true

[26] If $\varrho < 0$, meaning that assumption (A3) is violated, the right-hand side of (78) is strictly increasing and concave in h^X, goes to $-\infty$ for $h^X \to 0$ and approches a strictly positive value for $h^X \to \infty$. Thus, in this case, either two solutions or no solution for h^X as given by (78) exist. Two solutions means that two interior BGE exist, one stable and one unstable (see online-appendix for a numerical example). If no solution to (78) exists, then $h^X = 0$ would hold in BGE.

for the transversality condition associated with a_l. This concludes the proof. ■

Proof of Proposition 2. Using (70) and (76) we have

$$\frac{H_H^A}{H_L^A}=\frac{h_H^A}{h_L^A}=\left(\frac{h^X}{l^X}\right)^{1-\varrho}. \tag{79}$$

Substituting (79) into $\tilde{A}_H/\tilde{A}_L=\left(h_H^A/h_L^A\right)^{\frac{1-\theta}{1-\phi}}$ (recall (74)) and using $1-\varrho=\frac{(\psi-1)(1-\phi)}{1-\theta-\psi(\phi-\theta)}$, according to (77), we obtain

$$\frac{A_H}{A_L}=\frac{\tilde{A}_H}{\tilde{A}_L}=\left(\frac{h^X}{l^X}\right)^{\frac{(\psi-1)(1-\theta)}{1-\theta-\psi(\phi-\theta)}}. \tag{80}$$

Substituting (42) and (67) into (59) and using $\frac{1}{1-\alpha}=\frac{\varepsilon-1}{\psi-1}$, we find

$$\tilde{w}_l=(1-\alpha)\left(\frac{\alpha}{\kappa(r+\delta_K)}\right)^{\frac{\alpha}{1-\alpha}}\left(\frac{\bar{\nu}}{g}\right)^{\frac{1}{1-\phi}}\left(h_L^A\right)^{\frac{1-\theta}{1-\phi}}\left(1+\left(\frac{\tilde{A}_H h^X}{\tilde{A}_L l^X}\right)^{\frac{\psi-1}{\psi}}\right)^{\frac{1}{\psi-1}}. \tag{81}$$

Substituting (76), (80) and $l^X=\chi l^*$ into (81) implies that the long run level of $\tilde{w}_l$, denoted by $\tilde{w}_l^*$, is given by

$$\begin{aligned}\tilde{w}_l^* &= (1-\alpha)\left(\frac{\alpha}{\kappa(r^*+\delta_K)}\right)^{\frac{\alpha}{1-\alpha}}\left(\frac{\bar{\nu}}{g}\right)^{\frac{1}{1-\phi}}\Gamma(\tau_r,\tau_g)^{\frac{1-\theta}{1-\phi}}\times \\ &\quad \left(h^{X*}\right)^{\frac{(1-\theta)\varrho}{1-\phi}}\left(\chi l^*\right)^{\frac{(1-\theta)(1-\varrho)}{1-\phi}}\left(1+\left(\frac{h^{X*}}{\chi l^*}\right)^{1-\varrho}\right)^{\frac{1}{\psi-1}},\end{aligned} \tag{82}$$

where we used $1-\varrho=\frac{(\psi-1)(1-\phi)}{1-\theta-\psi(\phi-\theta)}$. Defining $\tilde{W}_l^*\equiv\tilde{w}_l^* l^*$ and recalling the definition of ϱ in (77), we find

$$\begin{aligned}\tilde{W}_l^* &= \tilde{w}_l^* l^*=(1-\alpha)\left(\frac{\alpha}{\kappa(r^*+\delta_K)}\right)^{\frac{\alpha}{1-\alpha}}\left(\frac{\bar{\nu}}{g}\right)^{\frac{1}{1-\phi}}\Gamma(\tau_r,\tau_g)^{\frac{1-\theta}{1-\phi}}\left(h^{X*}\right)^{\frac{(1-\theta)\varrho}{1-\phi}}\times \\ &\quad \chi^{-\varrho}\left(l^*\right)^{1-\varrho}\left[\left(\chi l^*\right)^{1-\varrho}+\left(h^{X*}\right)^{1-\varrho}\right]^{\frac{1}{\psi-1}}.\end{aligned} \tag{83}$$

Recalling both $0\le\varrho<1$ and $\psi>1$, according to assumption (A2), we see that the right-hand side of (83) is increasing in both $h^{X*} = \bar{h}^X(\tau_r, \tau_g, \mathfrak{h}_l^E, \mathfrak{h}_h^{E*})$ and $l^* = \bar{l}(\mathfrak{h}_l^E, \mathfrak{h}_h^{E*})$. According to (62) and (64), $\bar{l}^*$ is an increasing function of $\mathfrak{h}_h^{E*}$. Moreover, an increase in ϑ and a decrease in τ_h affect $\bar{h}^X$ and $\bar{l}$ only through raising $\mathfrak{h}_h^{E*}$. Thus, $\tilde{W}_l^*$ is decreasing in τ_h and increasing in ϑ if and only if $\bar{h}^X(\cdot, \mathfrak{h}_h^{E*})$ is an increasing function of $\mathfrak{h}_h^{E*}$. This concludes the proof. ■

Proof of Proposition 3. Substituting $H^X/L^X = h^X/l^X$ and (80) into (12) we obtain $w_h/w_l = \left(h^X/l^X\right)^{-\varrho}$. Thus, the long run relative wage rate, w_h^*/w_l^*, can be written as

$$\frac{w_h^*}{w_l^*} = \left(\frac{\chi \bar{l}^*(\mathfrak{h}_l^E, \mathfrak{h}_h^{E*})}{\bar{h}^{X*}(\tau_r, \tau_g, \mathfrak{h}_l^E, \mathfrak{h}_h^{E*})}\right)^{\varrho}. \tag{84}$$

Recall that, under assumption (A2), $\varrho \geq 0$. According to the properties of functions $\bar{h}^X$ and $\bar{l}$, if $\varrho > 0$, then w_h^*/w_l^* is increasing in $\mathfrak{h}_l^E$, but decreasing in τ_r and τ_g. ■

References

[1] Acemoglu, Daron (1998). Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality, *Quarterly Journal of Economics* 113, 1055-1089.

[2] Acemoglu, Daron (2002). Directed Technical Change, *Review of Economic Studies* 69, 781-810.

[3] Acemoglu, Daron and David Autor (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings, in: Ashenfelter, Orley and David E. Card (eds.), *Handbook of Labor Economics*, Vol. 4, Amsterdam: Elsevier.

[4] Acemoglu, Daron and David Autor (2012). What Does Human Capital Do? A Review of Goldin and Katz's 'The Race between Education and Technology', *Journal of Economic Literature* 50, 426–463.

[5] Aghion, Philippe and Patrick Bolton (1997). A Theory of Trickle-Down Growth and Development, *Review of Economic Studies* 64, 151-172.

[6] Aghion, Philippe and Peter Howitt (2005). Growth with Quality-Improving Innovations: An Integrated Framework, in: P. Aghion and S. Durlauf (eds.), *Handbook of Economic Growth*, North-Holland, Amsterdam.

[7] Audretsch, David B. and Maryann P. Feldman (1996). R&D Spillovers and the Geography of Innovation and Production, *American Economic Review* 86, 630-640.

[8] Black, Sandra E., Paul Devereux and Kjell G. Salvanes (2005). Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital, *American Economic Review* 95, 437-449.

[9] Bloom, Dan (2010), Programs and Policies to Assist High School Dropouts in the Transition to Adulthood, *The Future of Children* 20 (1), pp. 89-108.

[10] Bricker, Jesse, Arthur B. Kennickell, Kevin B. Moore and John Sabelhaus (2012). Changes in U.S. Family Finances from 2007 to 2010: Evidence from the Survey of Consumer Finances, *Federal Reserve Bulletin* 98 (2), 1-80.

[11] Bureau of Economic Analysis (2015). National Income and Product Accounts - Government Current Receipts and Expenditures, retrieved from www.bea.gov on May 12, 2015.

[12] Bureau of Labor Statistics (2015). News Release, January 21, 2015, www.bls.gov/news.release/pdf/wkyeng.pdf.

[13] Che, Yi and Lei Zhang (2014). Human Capital and Firm Performance: Evidence from China's Higher Education Expansion in the Late 1990s, Shanghai Jiao Tong University (mimeo).

[14] Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez (2014). Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States, *Quarterly Journal of Economics* 129, 1553-1623.

[15] Corak, Miles (2013). Income Inequality, Equality of Opportunity, and Intergenerational Mobility, *Journal of Economic Perspectives* 27, 79-102.

[16] Crosley, Adair and Brandon Roberts (2007). Strengthening State Policies to Increase the Education and skills of Low-Wage Workers, The Working Poor Families Project, Policy Brief, Spring, pp. 1-8.

[17] Cunha, Flavio, James Heckman, Lance Lochner and Dimitriy Masterov (2006). Interpreting the Evidence on Life Cycle Skill Formation, in: Hanushek, Eric and Finis Welch (eds.), *Handbook of the Economics of Education* Vol. 1. Amsterdam: North-Holland, pp. 697-812.

[18] Deaton, Angus (2013). The Great Escape: Health, Wealth, and the Origins of Inequality, Princeton University Press.

[19] DeNavas-Walt, Carmen, Bernadette D. Proctor, Jessica C. Smith (2013). Income, Poverty, and Health Insurance Coverage in the United States: 2012, U.S. Census Bureau, Current Population Reports, U.S. Government Printing Office, Washington, DC.

[20] Dustmann, Christian, Johannes Ludsteck and Uta Schönberg (2009). Revisiting the German Wage Structure, *Quarterly Journal of Economics* 124, 843-881.

[21] Fernandez, Raquel and Richard Rogerson (1995). On the Political Economy of Education Subsidies, *Review of Economic Studies* 62, 249–262.

[22] Galor, Oded and Omar Moav (2000). Ability-Biased Technological Transition, Wage Inequality, and Economic Growth, *Quarterly Journal of Economics*,115, 469–497.

[23] Galor, Oded and Joseph Zeira (1993). Income Distribution and Macroeconomics, *Review of Economic Studies* 60, 35-52.

[24] Goldin, Claudia and Lawrence F. Katz (2007). Long-Run Changes in the Wage Structure: Narrowing, Widening, Polarizing, *Brookings Papers on Economic Activity* 2.

[25] Goldin, Claudia and Lawrence F. Katz (2008). The Race between Education and Technology, Harvard University Press.

[26] Gordon, Nora (2013). High School Graduation in the Context of Changing Elementary and Secondary Education Policy and Income Inequality: The Last Half Century, Georgetown University (mimeo).

[27] Gould, Elise and Hilary Wething (2012). U.S. Poverty Rates Higher, Safety Net Weaker than in Peer Countries, Economic Policy Institute, Issue Brief #339.

[28] Grossmann, Volker, Thomas M. Steger and Timo Trimborn (2013). Dynamically Optimal R&D Subsidization, *Journal of Economic Dynamics and Control* 37, 516–534.

[29] Grossmann, Volker, Thomas M. Steger and Timo Trimborn (2015). Quantifying Optimal Growth Policy, *Journal of Public Economic Theory*, forthcoming.

[30] Heckmann, James J. (1976). A Life-Cycle Model of Earnings, Learning, and Consumption. *Journal of Political Economy* 84, S11–44.

[31] Heston, Alan, Robert Summers and Bettina Aten (2012), Penn World Table Version 7.1, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, Nov 2012.

[32] Hirsch, Barry T. (1980). Poverty and Economic Growth: Has Trickle Down Petered Out?, *Economic Inquiry* 18, 151-58.

[33] Johnson, George E. (1997). Changes in Earnings Inequality: The Role of Demand Shifts, *Journal of Economic Perspectives* 11, 41-54.

[34] Jones, Charles I. (1995). R&D-based Models of Economic Growth, *Journal of Political Economy* 103, 759–784.

[35] Jones, Charles I. (2005). Growth and Ideas, in: P. Aghion and S. Durlauf (eds.), *Handbook of Economic Growth*, North-Holland, Amsterdam.

[36] Jones, Charles I. and John C. Williams (2000), Too Much of a Good Thing? The Economics of Investment in R&D, *Journal of Economic Growth* 5, 65-85.

[37] Hirsch, Barry T. (1980). Poverty and Economic Growth: Has Trickle Down Petered Out?, *Economic Inquiry* 18, 151–158.

[38] Katz, Lawrence F. and Kevin M. Murphy (1992). Changes in Relative Wages, 1963-1987: Supply and Demand factors, *Quarterly Journal of Economics* 107, 35-78.

[39] Kautz, Tim, James J. Heckman, Ron Diris, Bas ter Weel and Lex Borghans (2014). Fostering and Measuring Skills: Improving Cognitive and Non-Cognitive Skills to Promote Lifetime Success, IZA Discussion Paper No. 8696.

[40] Kuznets, Simon (1955). Economic Growth and Income Inequality, *American Economic Review* 45, 1-28.

[41] Lee, Sangheon, Deidre McCann and Jon C. Messenger (2007). Working Time Around the World – Trends in Working Hours, Laws and Policies in a Global Comparative Perspective, International Labour Organization, Routledge.

[42] Lochner, Lance J. and Alexander Monge-Naranjo (2011). The Nature of Credit Constraints and Human Capital, *American Economic Review* 101, 2487-2529.

[43] Lucas, R. E. (1988). On the Mechanics of Economic Development *Journal of Monetary Economics* 22, 3–42.

[44] Mankiw, N. Gregory (2013). Defending the One Percent, *Journal of Economic Perspectives* 27, 21-34.

[45] Matsuyama, Kiminori (2000). Endogenous Inequality, *Review of Economic Studies* 67, 743-759.

[46] Mehra, R. and Prescott, E. C. (1985). The Equity Premium: A Puzzle, *Journal of Monetary Economics* 15 (2), 145–161.

[47] Murray Z. Frank, Rajdeep Singh and Tracy Yue Wang (2012). The Effect of Dividend Taxation on Corporate Investment: Old Debate, New Evidence, University of Minnesota (mimeo).

[48] Norrbin, S. C. (1993). The Relation between Price and Marginal Cost in U.S. Industry: A Contradiction. *Journal of Political Economy* 101 (6), 1149–64.

[49] OECD (2009). OECD Science, Technology and Industry Scoreboard 2009, Paris.

[50] OECD (2014). Education at a Glance 2014, Paris.

[51] OECD (2015). Public Sector, Taxation and Market Regulation, OECD.Stat, Paris retrieved from http://stats.oecd.org/index.aspx?DataSetCode=TABLE_I4, retrived on May 11, 2015.

[52] Osikominu, Aderonke (2013). Quick Job Entry or Long-Term Human Capital Development? The Dynamic Effects of Alternative Training Schemes, *Review of Economic Studies* 80, 313-342.

[53] Piketty, Thomas (1997). The Dynamics of the Wealth Distribution and the Interest Rate with Credit Rationing, *Review of Economic Studies* 64, 173-189.

[54] Piketty, Thomas (2014). Capital in the Twenty-First Century, Harvard University Press.

[55] Plug, Erik and Wim Vijverberg (2003). Schooling, Family Background, and Adoption: Is It Nature or Is It Nurture?, *Journal of Political Economy* 111, 611-641.

[56] Schochet, Peter Z., John Burghardt, and Sheena McConnell (2008). Does Job Corps Work? Impact Findings from the National Job Corps Study, *American Economic Review* 98, 1864-1886.

[57] Stiglitz, Joseph E. (2012). The Price of Inequality: How Today's Divided Society Endangers Our Future, W. W. Norton and Company.

[58] Thornton, James R., Richard J. Agnello and Charles R. Link (1978). Poverty and Economic Growth: Trickle Down Peters Out, *Economic Inquiry* 16, 385–394.

[59] Trimborn, Timo, Karl-Josef Koch and Thomas M. Steger (2008). Multi-Dimensional Transitional Dynamics: A Simple Numerical Procedure, *Macroeconomic Dynamics* 12, 301–319.

[60] Von Weizsäcker, Carl-Christian (1966). Tentative Notes on a Two Sector Model with Induced Technical Progress, *Review of Economic Studies* 33, 245-251.

Online-Appendix

In this online-appendix, we first summarize the dynamical system and the balanced growth equilibrium (Appendix I). We also show that the long run values of individual asset holdings and consumption levels are indeterminate, i.e. depend on the initial value of relative asset holdings, $a_{h,0}/a_{l,0}$. We then derive expressions and relationships used for calibrating the model (Appendix II). Next, we provide algebraic details for the policy reforms analyzed in section 6 of the main paper (Appendix III). Finally (Appendix IV), we display transitional dynamics of all variables in response to policy reforms, when adjusting the top labor income tax rate τ_h (Fig. A.2-A.4). We also consider the policy reform implications on consumption of type$-l$ individuals when we adjust the fourth spending category and hold tax rates constant (Fig. A.5).

Appendix I. Dynamical System and Balanced Growth Equilibrium

Differential equations:

$$\dot{l} = \xi(h_l^E)^\gamma l^\eta - \delta_H l, \tag{85}$$

$$\dot{h} = \xi(h_h^E)^\beta h^\eta - \delta_H h, \tag{86}$$

$$\frac{\dot{\tilde{c}}_h}{\tilde{c}_h} = \frac{(1-\tau_r)r - \rho}{\sigma} - g, \tag{87}$$

$$\frac{\dot{m}}{m} = \delta_H + \rho - n + (\sigma - 1)g - \left(\eta \mathfrak{h}_h^E + \beta \frac{1-\tau_h}{1-\vartheta}\right) \xi(\mathfrak{h}_h^E)^{\beta-1} h^{\eta+\beta-1}, \tag{88}$$

$$\frac{\dot{\tilde{a}}_h}{\tilde{a}_h} = (1-\tau_r)r - n + (1-\tau_h)\frac{\tilde{w}_h h}{\tilde{a}_h} - (1-\vartheta)\frac{\tilde{w}_h h_h^E}{\tilde{a}_h} - \frac{\tilde{c}_h}{\tilde{a}_h} - g, \tag{89}$$

$$\frac{\dot{\tilde{c}}_l}{\tilde{c}_l} = \frac{(1-\tau_r)r - \rho}{\sigma} - g, \tag{90}$$

$$\frac{\dot{\tilde{a}}_l}{\tilde{a}_l} = (1-\tau_r)r - n + (1-\tau_l)\frac{\tilde{w}_l l}{\tilde{a}_l} - \frac{\tilde{c}_l}{\tilde{a}_l} + \frac{\tilde{T}}{\tilde{a}_l} - g, \tag{91}$$

$$\frac{\dot{\tilde{A}}_H}{\tilde{A}_H} = \bar{\nu}(\tilde{A}_H)^{\phi-1}(h_H^A)^{1-\theta} - g, \tag{92}$$

$$\frac{\dot{\tilde{A}}_L}{\tilde{A}_L} = \bar{\nu}(\tilde{A}_L)^{\phi-1}(h_L^A)^{1-\theta} - g, \tag{93}$$

$$\dot{p}_H^A = \frac{1-\tau_r}{1-\tau_g}\left(rp_H^A - \frac{(\kappa-1)\left(\frac{\alpha}{\kappa}P_H^X\right)^{\frac{1}{1-\alpha}} h^X}{(r+\delta_K)^{\frac{\alpha}{1-\alpha}}}\right) - np_H^A, \tag{94}$$

$$\dot{p}_L^A = \frac{1-\tau_r}{1-\tau_g}\left(rp_L^A - \frac{(\kappa-1)\left(\frac{\alpha}{\kappa}P_L^X\right)^{\frac{1}{1-\alpha}} l^X}{(r+\delta_K)^{\frac{\alpha}{1-\alpha}}}\right) - np_L^A. \tag{95}$$

Algebraic equations:

$$h_h^E = \mathfrak{h}_h^E h, \tag{96}$$

$$\begin{aligned}\tilde{a}_h + \chi\tilde{a}_l = {} & \tilde{A}_H\left(\frac{\alpha P_H^X}{\kappa(r+\delta_K)}\right)^{\frac{1}{1-\alpha}} h^X + \\ & \tilde{A}_L\left(\frac{\alpha P_L^X}{\kappa(r+\delta_K)}\right)^{\frac{1}{1-\alpha}} l^X + p_H^A\tilde{A}_H + p_L^A\tilde{A}_L,\end{aligned} \tag{97}$$

$$P_H^X = \left[1 + \left(\frac{\tilde{A}_H h^X}{\tilde{A}_L l^X}\right)^{-\frac{\psi-1}{\psi}}\right]^{\frac{1}{\varepsilon-1}}, \tag{98}$$

$$P_L^X = \left[1 + \left(\frac{\tilde{A}_H h^X}{\tilde{A}_L l^X}\right)^{\frac{\psi-1}{\psi}}\right]^{\frac{1}{\varepsilon-1}}, \tag{99}$$

$$m\xi\beta(\mathfrak{h}_h^E)^{\beta-1}h^{\eta+\beta-1} = (\tilde{c}_h)^{-\sigma}(1-\vartheta)\tilde{w}_h, \tag{100}$$

$$\tilde{w}_h = (1-\alpha)\left(\frac{\alpha}{\kappa(r+\delta_K)}\right)^{\frac{\alpha}{1-\alpha}} \tilde{A}_H\left(P_H^X\right)^{\frac{1}{1-\alpha}}, \tag{101}$$

$$\tilde{w}_l = (1-\alpha)\left(\frac{\alpha}{\kappa(r+\delta_K)}\right)^{\frac{\alpha}{1-\alpha}} \tilde{A}_L\left(P_L^X\right)^{\frac{1}{1-\alpha}}, \tag{102}$$

$$p_H^A \bar{\nu} \left(\tilde{A}_H\right)^{\phi-1} (h_H^A)^{-\theta} = \frac{\tilde{w}_h}{\tilde{A}_H}, \tag{103}$$

$$p_L^A \bar{\nu} \left(\tilde{A}_L\right)^{\phi-1} (h_L^A)^{-\theta} = \frac{\tilde{w}_h}{\tilde{A}_L}, \tag{104}$$

$$h^X + h_H^A + h_L^A + h_h^E = (1 - \chi \mathfrak{h}_l^E) h, \tag{105}$$

$$l^X = \chi l. \tag{106}$$

Steady state values:

$$r^* = \frac{\rho + \sigma g}{1 - \tau_r}. \tag{107}$$

$$l^* = \left(\frac{\xi(\mathfrak{h}_l^E h^*)^\gamma}{\delta_H}\right)^{\frac{1}{1-\eta}}; \tag{108}$$

$$h^* = \left[\frac{\xi(\mathfrak{h}_h^{E*})^\beta}{\delta_H}\right]^{\frac{1}{1-\beta-\eta}}, \tag{109}$$

where

$$\mathfrak{h}_h^{E*} = \frac{1-\tau_h}{1-\vartheta} \frac{\beta \delta_H}{\rho - n + (\sigma - 1)g + (1-\eta)\delta_H}. \tag{110}$$

Under assumption (A2), $\varrho \geq 0$, h^{X*} is implicitly defined by

$$(1+\Gamma)h^X = (1 - \chi \mathfrak{h}_l^E) h^* - \mathfrak{h}_h^{E*} h^* - \Gamma \left(h^X\right)^\varrho (\chi l^*)^{1-\varrho} \tag{111}$$

as a unique value, where

$$\Gamma = \frac{1 - \frac{1}{\kappa}}{\frac{1}{\alpha} - 1} \frac{(1-\tau_r)g}{\rho + \sigma g - (1-\tau_g)n}, \tag{112}$$

$$\varrho = \frac{2 - \phi - \psi + \theta(\psi - 1)}{1 - \theta - \psi(\phi - \theta)}. \tag{113}$$

If $\varrho < 0$ (i.e. (A2) is violated), as discussed in the Remark in Appendix B, (111) defines either two solutions or no interior solution for h^{X*}. Using our baseline calibration in Tab. 1 (where $\varrho = 0$), except changing the derived elasticity of substitution from $\psi = 1.5$ to $\psi = 1.9$ (thus, $\varrho = -8$), there are two steady state values h^{X*}, only the higher one being stable; see Fig. A.1.

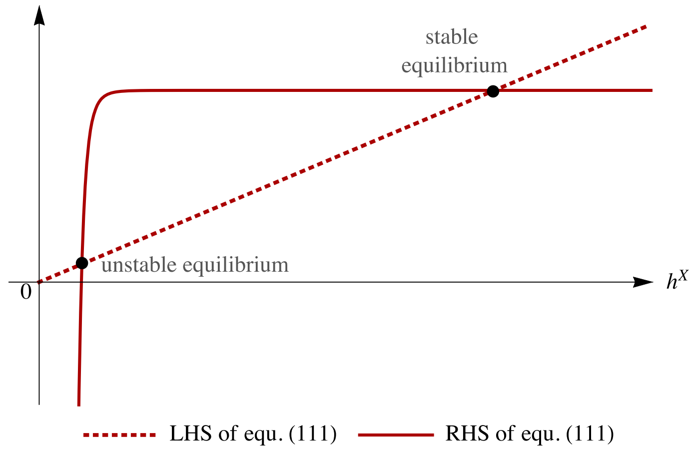

Figure A.1: Multiple steady states when the second inequality of (A3) is violated ($\varrho < 0$). Parameters like in Tab. 1 except $\psi = 1.9$.

Using h^{X*} we find long run values of h_H^A and h_L^A:

$$h_H^{A*} = \Gamma h^{X*}, \tag{114}$$

$$h_L^{A*} = \Gamma \left(h^{X*}\right)^{\varrho} (\chi l^*)^{1-\varrho}. \tag{115}$$

Setting $\dot{\tilde{A}}_H = \dot{\tilde{A}}_L = 0$ in (92) and (93) yields, by using h_H^{A*}, h_L^{A*}, the steady state values of $\tilde{A}_H$ and $\tilde{A}_L$:

$$\tilde{A}_H^* = \left(\frac{\bar{\nu} \left(h_H^{A*}\right)^{1-\theta}}{g} \right)^{\frac{1}{1-\phi}}, \tag{116}$$

$$\tilde{A}_L^* = \left(\frac{\bar{\nu} \left(h_L^{A*}\right)^{1-\theta}}{g} \right)^{\frac{1}{1-\phi}}. \tag{117}$$

Using $\tilde{A}_H^*$, $\tilde{A}_L^*$, h^{X*} and $l^X = \chi l^*$ in (98) and (99) gives us long run values P_H^{X*} and P_L^{X*}, respectively.

Using $\tilde{A}_H^*$, $\tilde{A}_L^*$, P_H^{X*}, P_L^{X*} and r^* in (101) and (102) give us long run values $\tilde{w}_h^*$ and $\tilde{w}_l^*$, respectively.

Using $\tilde{A}_H^*$, $\tilde{A}_L^*$, h_H^{A*}, h_L^{A*} and $\tilde{w}_h^*$ in (103) and (104) give us long run values p_H^{A*} and p_L^{A*}, respectively.

Finally, setting $\dot{\tilde{a}}_h = \dot{\tilde{a}}_l = 0$ in (89) and (91) implies

$$0 = (1-\tau_r)r - n + (1-\tau_h)\frac{\tilde{w}_h h}{\tilde{a}_h} - (1-\vartheta)\frac{\tilde{w}_h \mathfrak{h}_h^E h}{\tilde{a}_h} - \frac{\tilde{c}_h}{\tilde{a}_h} - g, \tag{118}$$

$$0 = (1-\tau_r)r - n + (1-\tau_l)\frac{\tilde{w}_l l}{\tilde{a}_l} - \frac{\tilde{c}_l}{\tilde{a}_l} + \frac{\tilde{T}}{\tilde{a}_l} - g. \tag{119}$$

Using long run values $\tilde{w}_h^*$, $\tilde{w}_l^*$, $\mathfrak{h}_h^{E*}$, h^*, r^* and $\tilde{T}^*$, there are the four equations (97), (100), (118), (119) left for the five remaining unknown long run values of $\tilde{c}_h$, $\tilde{c}_l$, $\tilde{a}_h$, $\tilde{a}_l$ and m. Unlike the long run values of other variables, $\tilde{c}_h^*$, $\tilde{c}_l^*$, $\tilde{a}_h^*$, $\tilde{a}_l^*$ and m^* depend on initial conditions. The initial values of assets holdings $a_{h,0}$, $a_{l,0}$, are related to the initial values of the number of machines, $A_{H,0}$, $A_{L,0}$, according to (97):

$$\begin{aligned} a_{h,0} + \chi a_{l,0} &= A_{H,0}\left(\frac{\alpha P_{H,0}^X}{\kappa(r_0+\delta_K)}\right)^{\frac{1}{1-\alpha}} h_0^X + \\ & A_{L,0}\left(\frac{\alpha P_{L,0}^X}{\kappa(r_0+\delta_K)}\right)^{\frac{1}{1-\alpha}} l_0^X + p_{H,0}^A A_{H,0} + p_{L,0}^A A_{L,0}. \end{aligned} \tag{120}$$

Thus, to find long run values $\tilde{c}_h^*$, $\tilde{c}_l^*$, $\tilde{a}_h^*$, $\tilde{a}_l^*$ and m^*, we can fix a long run value $\mathfrak{s} \equiv a_h^*/a_l^*$ belonging to some configuration of initial conditions $a_{h,0}$, $a_{l,0}$, $A_{H,0}$, $A_{L,0}$, which fulfills (120). In this sense, $\tilde{c}_h^*$, $\tilde{c}_l^*$, $\tilde{a}_h^*$, $\tilde{a}_l^*$ and m^* are *indeterminate*. Evaluating (97) at long run values and using $\tilde{a}_h^* = \mathfrak{s}\tilde{a}_l^*$, implies

$$\begin{aligned} \tilde{a}_l^* &= \frac{1}{\mathfrak{s}+\chi}\left(\tilde{A}_H^*\left(\frac{\alpha P_H^{X*}}{\kappa(r^*+\delta_K)}\right)^{\frac{1}{1-\alpha}} h^{X*} + \right. \\ & \left. \tilde{A}_L^*\left(\frac{\alpha P_L^{X*}}{\kappa(r^*+\delta_K)}\right)^{\frac{1}{1-\alpha}} l^{X*} + p_H^{A*}\tilde{A}_H^* + p_L^{A*}\tilde{A}_L^*\right). \end{aligned} \tag{121}$$

Evaluating (119) at $\tilde{a}_l^*$ and the other long run values gives us $\tilde{c}_l^*$. Similarly, evaluating

(118) at $\tilde{a}_h^* = \mathfrak{s}\tilde{a}_l^*$ and the other long run values gives us $\tilde{c}_h^*$. Finally, evaluating (100) at $\tilde{c}_h^*$ and the other long run values gives us m^*.

Appendix II. Calibration

Public Sector: The total tax revenue ($\mathcal{TTR}$) is the sum of the revenue from taxation of labor income and returns to asset holding,

$$\mathcal{TTR} = \tau_h N_h w_h h + \tau_l N_l w_l l + \tau_r r K + \tau_g (\dot{P}_H^A A_H + \dot{P}_L^A A_L) + \tau_r (\pi_H A_H + \pi_L A_L). \quad (122)$$

Note that $\dot{P}_H^A / N_h = \dot{p}_H^A + n p_H^A$ and $\dot{P}_L^A / N_h = \dot{p}_L^A + n p_L^A$, as given by the right-hand side of (60) and (61), respectively. Moreover, combining (4) with (25) and (26), we can write

$$K = A_H \left(\frac{\alpha P_H^X}{\kappa(r + \delta_K)} \right)^{\frac{1}{1-\alpha}} H^X + A_L \left(\frac{\alpha P_L^X}{\kappa(r + \delta_K)} \right)^{\frac{1}{1-\alpha}} L^X. \quad (123)$$

Inserting (29), (30) and (123) in (122) and using (32) and $\chi = N_l / N_h$ we obtain

$$\begin{aligned} \Xi \equiv \; & \frac{\mathcal{TTR}}{N_h e^{gt}} = \tau_h \tilde{w}_h h + \tau_l \tilde{w}_l \chi l + \frac{(1-\tau_r)\tau_g}{1-\tau_g} r (p_H^A \tilde{A}_H + p_L^A \tilde{A}_L) + \\ & \frac{\tau_r - \tau_g}{1-\tau_g} \frac{(\kappa - 1) \left(\frac{\alpha}{\kappa}\right)^{\frac{1}{1-\alpha}}}{(r + \delta_K)^{\frac{\alpha}{1-\alpha}}} \left(\left(P_H^X\right)^{\frac{1}{1-\alpha}} \tilde{A}_H h^X + \left(P_L^X\right)^{\frac{1}{1-\alpha}} \tilde{A}_L \chi l \right) + \\ & \tau_r r \left(\tilde{A}_H \left(\frac{\alpha P_H^X}{\kappa(r + \delta_K)} \right)^{\frac{1}{1-\alpha}} h^X + \tilde{A}_L \left(\frac{\alpha P_L^X}{\kappa(r + \delta_K)} \right)^{\frac{1}{1-\alpha}} \chi l \right). \end{aligned} \quad (124)$$

Hence, if tax rates $(\tau_h, \tau_l, \tau_g, \tau_r)$ are time-invariant, Ξ is stationary in the long run and can be obtained using h^*, $\tilde{w}_h^*$, $\tilde{w}_l^*$, $\tilde{A}_H^*$, $\tilde{A}_L^*$, P_H^{X*}, P_L^{X*}, p_H^{A*}, p_L^{A*}, h^{X*} in (124).

Aggregate transfer payments to type$-l$ individuals read as $N_l T$. Thus, the fraction of transfer payments in total tax revenue is $s^T \equiv N_l T / \mathcal{TTR}$. Using $\chi = N_l / N_h$ and $\tilde{T}_t \equiv T_t e^{-gt}$, we thus obtain

$$s^T = \frac{\chi \tilde{T}}{\Xi}. \quad (125)$$

Hence, a time-invariant s^T goes along with a stationary growth-adjusted transfer, $\tilde{T}$, in the long run.

Public expenditure for employing human capital to educate type$-l$ individuals reads as $E_l \equiv w_h N_l h_l^E$. Thus, recalling $\mathfrak{h}_l^E = h_l^E/h$, the fraction of tax revenue devoted to subsidize education of low-skilled workers, $s_l^E \equiv E_l/\mathcal{TTR}$, is time-invariant in the long run and given by

$$s_l^E = \frac{\chi \tilde{w}_h \mathfrak{h}_l^E h}{\Xi}. \tag{126}$$

Finally, public expenditure for subsidizing education costs of type$-h$ individuals (at rate ϑ) is given by $E_h \equiv \vartheta w_h N_h h_h^E$. Thus, the fraction of tax revenue devoted to subsidize education of high-skilled workers, $s_h^E \equiv E_h/\mathcal{TTR}$, is time-invariant in the long run and given by

$$s_h^E = \frac{\vartheta \tilde{w}_h h_h^E}{\Xi}. \tag{127}$$

Assuming that the economy is in BGE and for given (s^T, s_l^E, s_h^E), policy parameters $(\tilde{T}, \mathfrak{h}_l^E, \vartheta)$ are implicitly defined by the equation system (85)-(106), (124)-(127).

Skill premium: The steady state skill premium to which we calibrate our model is given by

$$\Omega^* \equiv \frac{w_h^*}{w_l^*} \frac{h^*}{l^*}, \tag{128}$$

where h^* and l^* are defined in part (v) and (vi) of Proposition 1, respectively, and w_h^*/w_l^* is given by (84).

R&D intensity: Substituting (25) and (26) into (1) we obtain

$$Y = \left(\frac{\alpha}{\kappa(r + \delta_K)} \right)^{\frac{\alpha}{1-\alpha}} \left[\left(A_L L^X \left(P_L^X \right)^{\frac{\alpha}{1-\alpha}} \right)^{\frac{\varepsilon-1}{\varepsilon}} + \left(A_H H^X \left(P_H^X \right)^{\frac{\alpha}{1-\alpha}} \right)^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}}. \tag{129}$$

Using expression (101) for $\tilde{w}_h$ and (129) in (18) we find that the long run R&D intensity is given by

$$R\&D^* \equiv \frac{(1-\alpha)\frac{\tilde{A}_H^*}{\tilde{A}_L^*} \left(P_H^{X*} \right)^{\frac{1}{1-\alpha}} \left(h_H^{A*} + h_L^{A*} \right)}{\left[\left(l^{X*} \left(P_L^{X*} \right)^{\frac{\alpha}{1-\alpha}} \right)^{\frac{\varepsilon-1}{\varepsilon}} + \left(\frac{\tilde{A}_H^*}{\tilde{A}_L^*} h^{X*} \left(P_H^{X*} \right)^{\frac{\alpha}{1-\alpha}} \right)^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}}}, \tag{130}$$

where $l^{X*} \equiv \chi l^*$, h^{X*} is given by (78), $h_H^{A*} \equiv \Gamma(\tau_r, \tau_g) h^{X*}$ (recall (70)) and $h_L^{A*} \equiv \Gamma(\tau_r, \tau_g) \left(h^{X*} \right)^{\varrho} (l^{X*})^{1-\varrho}$ (recall (76)). Moreover, using $h_H^A = h_H^{A*}$ in (66) and $h_L^A = h_L^{A*}$

in (67) gives us $\tilde{A}_H^*$ and $\tilde{A}_L^*$, respectively. Finally, evaluating the right-hand sides of (41) and (42) at long run values gives us P_H^{X*} and P_L^{X*}, respectively.

Returns to education: To derive R_h, first, we compute a first-order approximation of (7), $\dot{h} = \xi(h_h^E)^\beta h^\eta - \delta_H h \equiv G(h)$, around the steady state value

$$h^* = \left(\frac{\xi}{\delta_H}\right)^{\frac{1}{1-\eta}} (h_h^E)^{\frac{\beta}{1-\eta}}, \tag{131}$$

i.e. $\dot{h} \cong G(h^*) + G'(h^*)q$, where $q \equiv h - h^*$. Since $G(h^*) = 0$ by definition and $G'(h^*) = (\eta - 1)\delta_H$, we have $\dot{h} = \dot{q} \cong -(1-\eta)\delta_H q$. Thus, $q(t) = q(0)e^{-(1-\eta)\delta_H t}$, implying

$$h(t) \cong h^* + (h_0 - h^*)e^{-(1-\eta)\delta_H t}. \tag{132}$$

With a discount rate $R > 0$ we can thus approximate a present discounted value (PDV)

$$\int_0^\infty [h(t) - h_0]\, e^{-Rt}\mathrm{d}t \cong \frac{h^* - h_0}{R} + \frac{h_0 - h^*}{R + (1-\eta)\,\delta_H}. \tag{133}$$

Permanently raising teaching input h_h^E by one unit involves, at constant wage rate w_h, a PDV of costs equal to $w_h \int_0^\infty e^{-Rt}\mathrm{d}t = w_h/R$. The internal rate of return of type$-h$ individuals, R_h, equalizes their PDV of a gain in wage income, $w_h \int_0^\infty [h(t) - h_0]\, e^{-Rt}\mathrm{d}t$ (at constant wage rate w_h), with the PDV of costs. Using (133), R_h thus solves

$$w_h \left(\frac{h^* - h_0}{R_h} + \frac{h_0 - h^*}{R_h + (1-\eta)\,\delta_H}\right) = \frac{w_h}{R_h}, \text{ i.e.} \tag{134}$$

$$R_h = (1-\eta)\delta_H(h^* - h_0 - 1). \tag{135}$$

Since we look at a permanent increase in teaching input by one unit and start from an initial long run equilibrium, we can approximate $\partial h^*/\partial h_h^E \cong h^* - h_0$. Thus, the long run rate of return reads as

$$
\begin{aligned}
R_h^* &= (1-\eta)\delta_H \left(\left. \frac{\partial h^*}{\partial h_h^E} \right|_{h_h^E = \mathfrak{h}_h^{E*} h^*} - 1 \right) \\
&= (1-\eta)\delta_H \left(\frac{1-\vartheta}{1-\tau_h} \frac{\rho - n + (\sigma-1)g + (1-\eta)\delta_H}{(1-\eta)\delta_H} - 1 \right), \qquad (136)
\end{aligned}
$$

where for the latter equation we used (131) and substituted $\mathfrak{h}_h^{E*}$ and h^* from (14) and (15), respectively.[27]

Similarly, for type$-l$ individuals, according to (8), $\dot{l} = 0$ implies a long run human capital level

$$
l^* = \left(\frac{\xi}{\delta_H} \right)^{\frac{1}{1-\eta}} (h_l^E)^{\frac{\gamma}{1-\eta}}. \qquad (137)
$$

First-order approximating around l^* implies

$$
l(t) \cong l^* + (l_0 - l^*) e^{-(1-\eta)\delta_H t}. \qquad (138)
$$

The internal rate of return, R_l, is defined by

$$
w_l \int_0^\infty [l(t) - l_0] \, e^{-R_l t} \mathrm{d}t = w_h \int_0^\infty e^{-R_l t} \mathrm{d}t. \qquad (139)
$$

Solving (139) for R_l, approximating $\partial l^*/\partial h_l^E \cong l^* - l_0$ and using (137) suggests

$$
\begin{aligned}
R_l^* &= (1-\eta)\delta_H \left(\frac{w_l^*}{w_h^*} \times \left. \frac{\partial l^*}{\partial h_l^E} \right|_{h_l^E = \mathfrak{h}_l^{E*} h^*} - 1 \right) = \\
&\quad (1-\eta)\delta_H \left(\frac{w_l^*}{w_h^*} \frac{\gamma}{1-\eta} \left(\frac{\xi}{\delta_H} \right)^{\frac{1}{1-\eta}} (\mathfrak{h}_l^E h^*)^{-\frac{1-\gamma-\eta}{1-\eta}} - 1 \right), \qquad (140)
\end{aligned}
$$

where h^* is given by (15) and w_h^*/w_l^* is given by (84).

Appendix III. Policy Experiments – Algebraic Details (Section 6)

Total public spending for other purposes than transfers and education finance reads

[27] The analysis by Grossmann et al. (2015) suggests that the long run human capital stock, h, is socially optimal when $\vartheta = \tau_h$. In this case, $R_h^* = \rho - n + (\sigma - 1)g > 0$, according to assumption (A1).

as $E_{other} \equiv (1 - s_h^E - s_l^E - s^T)\mathcal{TTR}$, i.e. other public spending per type$-h$ worker adjusted for steady state productivity growth, $\Upsilon \equiv E_{other}e^{-gt}/N_h$, is given by

$$\Upsilon = \Xi - \vartheta \tilde{w}_h \mathfrak{h}_h^E h - \chi \tilde{w}_h \mathfrak{h}_l^E h - \chi \tilde{T}, \tag{141}$$

according to (125), (126), (127) and relationship $\mathcal{TTR} = N_h e^{gt}\Xi$ (recall 124).

Let us denote Υ_0^* as the long run value of Υ for the baseline calibration (pre-reform steady state) given in Table 1. In section 6, we examine the following policy reforms and evaluate its impact on the economic situation of type$-l$ workers.

1. Education expansion on behalf of type$-h$ workers: Consider the dynamical system (85)-(106), (124), (127) and

   $$\vartheta \tilde{w}_h h_h^E + \chi \tilde{w}_h \mathfrak{h}_l^E h + \chi \tilde{T} + \Upsilon_0^* = \Xi. \tag{142}$$

   We consider a permanent change in s_h^E (by 0.5 percentage points) and let either τ_h and, possibly, τ_r and τ_g adjust accordingly such that (142) holds at all times and both $\tilde{T}$ and $\mathfrak{h}_l^E$ as well as the other tax rates are kept at their initially calibrated levels.

2. Education expansion on behalf of type$-l$ workers: Consider the dynamical system (85)-(106), (124), (126) and (142). We consider a permanent change in s_l^E (by 0.5 percentage points) and let either τ_h and, possibly, τ_r and τ_g adjust accordingly such that (142) holds at all times and both $\tilde{T}$ and ϑ as well as the other tax rates are kept at their initially calibrated levels.

3. Increasing transfers: Consider the dynamical system (85)-(106), (124), (125) and (142). We consider a permanent change in s^T (by 0.5 percentage points) and let τ_h and, possibly, τ_r and τ_g adjust accordingly such that (142) holds at all times and both $\mathfrak{h}_l^E$ and ϑ as well as the other tax rates are kept at their initially calibrated levels.

Appendix IV. Trajectories (Section 6)

- Fig. A.2 displays the trajectories in response to education expansion on behalf of type$-h$ workers (increase in ϑ).

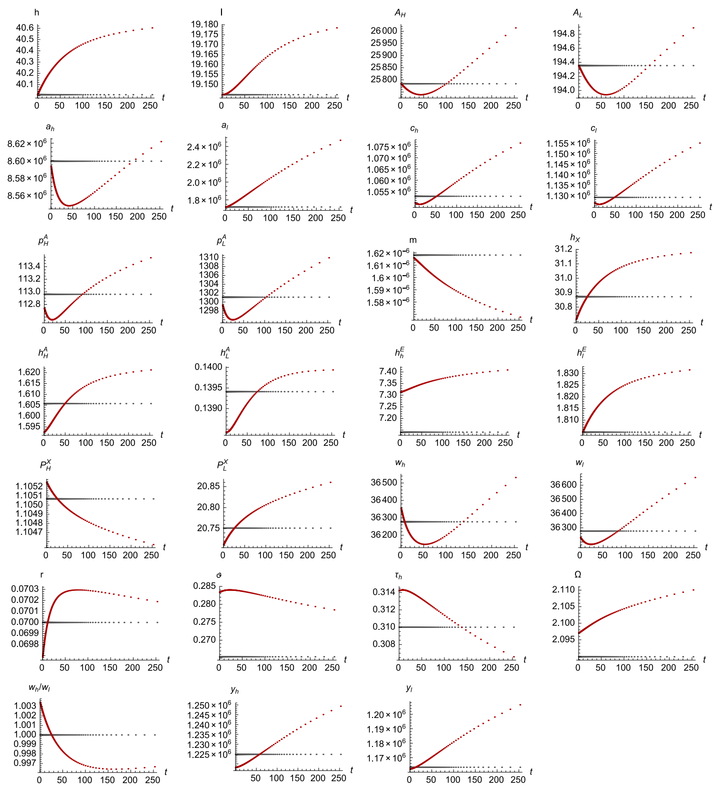

Figure A.2: Transitional dynamics in response to policy shock "expanding higher education" (increase in ϑ), related to Fig. 1-3. Set of parameters as in Table 1.

- Fig. A.3 displays the trajectories in response to education expansion on behalf of type$-l$ workers (increase in $\mathfrak{h}_l^E$).

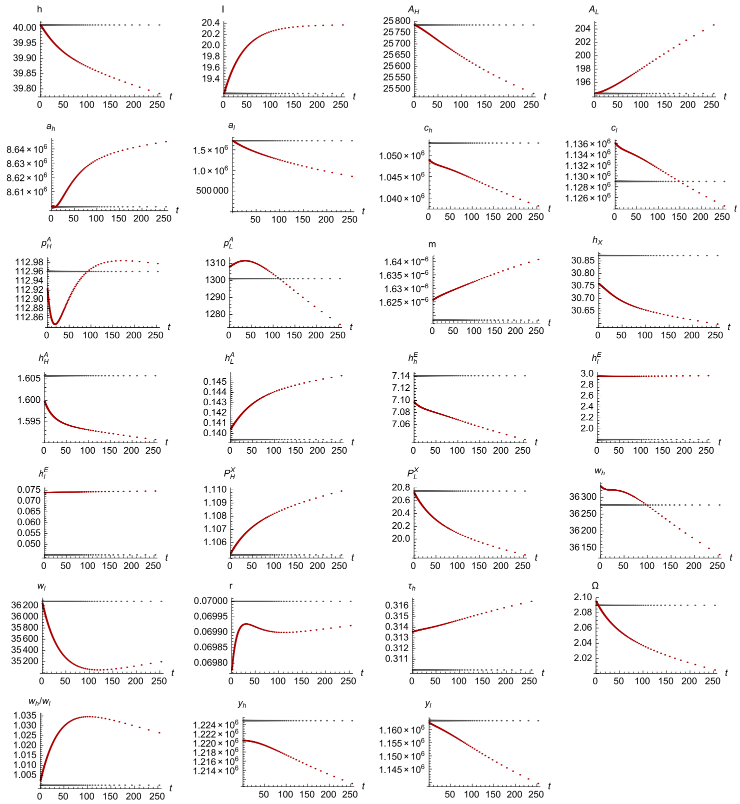

Figure A.3: Transitional dynamics in response to policy shock "expanding skills of low-ability workers" (increase in h_l^E), related to Fig. 1-3. Set of parameters as in Table 1.

- Fig. A.4 displays the trajectories in response to expansion of transfers (increase in $\tilde{T}$).

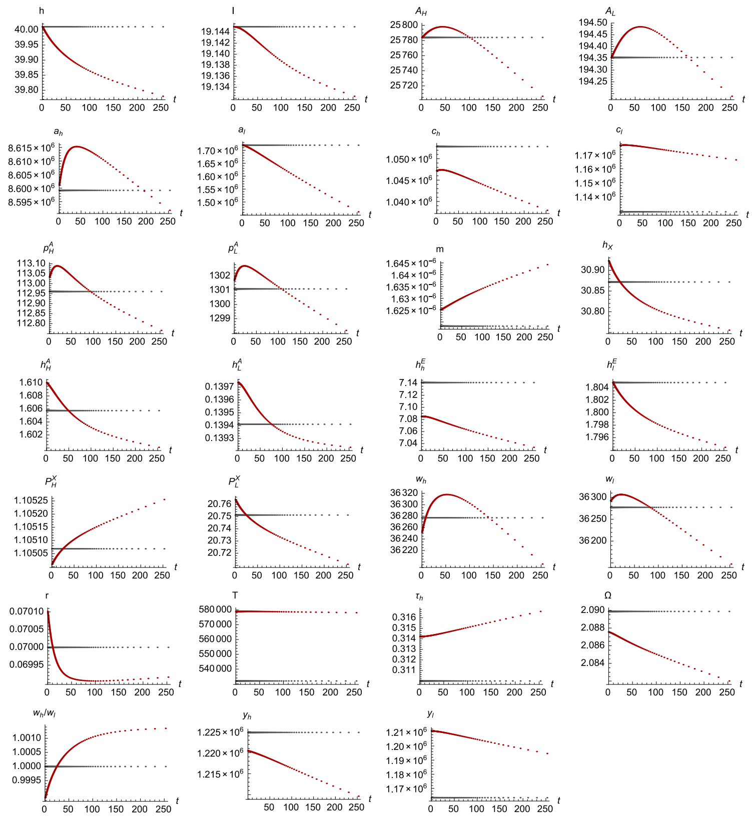

Figure A.4: Transitional dynamics in response to policy shock "expanding transfers" (increase in $\tilde{T}$), related to Fig. 1-3. Set of parameters as in Table 1.

- Finally, Fig. A.5 displays the implications of the three policy reforms for consumption of low-ability individuals, except that additional public spending is financed in a non-distortionary way; that is, all tax rates are kept constant. Comparison to Fig. 2 allows us to examine the role of tax distortions.

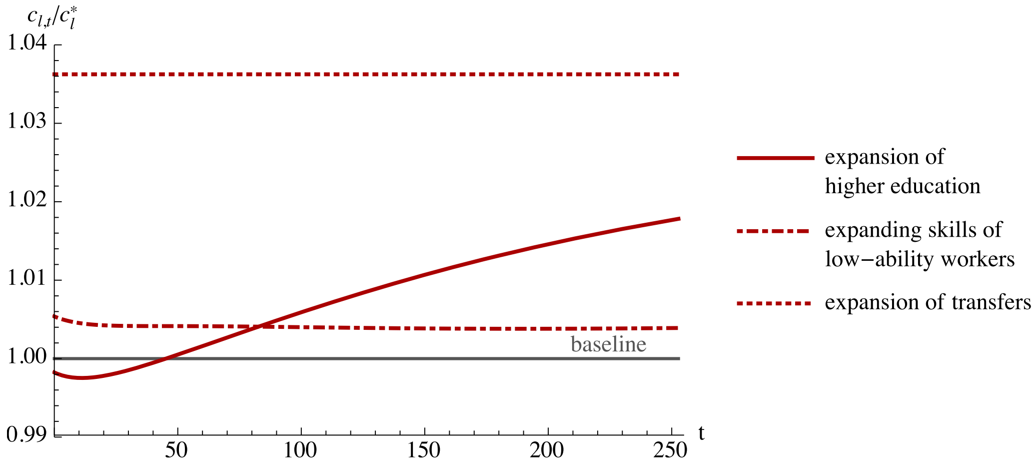

Figure A.5: Time paths of normalized consumption of type-l individuals, c_l/c_l^*, in response to three policy reforms under a non-distortionary financing scheme (i.e. constant tax rates). Set of parameters as in Table 1.