NBER WORKING PAPER SERIES

BALANCED GROWTH DESPITE UZAWA

Gene M. Grossman
Elhanan Helpman
Ezra Oberfield
Thomas Sampson

Working Paper 21861
http://www.nber.org/papers/w21861

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2016

We are grateful to Chad Jones, Gianluca Violante, and Jonathan Vogel for discussions and suggestions. Grossman and Helpman thank the Einaudi Institute for Economics and Finance, Grossman thanks Sciences Po and the Kyoto Institute for Economic Research, Oberfield thanks New York University, and Sampson thanks the Center for Economic Studies for their hospitality. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Balanced Growth Despite Uzawa
Gene M. Grossman, Elhanan Helpman, Ezra Oberfield, and Thomas Sampson
NBER Working Paper No. 21861
January 2016
JEL No. E25,J24,O40

ABSTRACT

The evidence for the United States points to balanced growth despite falling investment-good prices and an elasticity of substitution between capital and labor less than one. This is inconsistent with the Uzawa Growth Theorem. We extend Uzawa's theorem to show that the introduction of human capital accumulation in the standard way does not resolve the puzzle. However, balanced growth is possible if schooling is endogenous and capital is more complementary with schooling than with raw labor. We describe balanced growth paths for a variety of neoclassical growth models with capital-augmenting technological progress and endogenous schooling. The balanced growth path in an overlapping-generations model in which individuals choose the duration of their education matches key features of the U.S. economic record.

Gene M. Grossman
International Economics Section
Department of Economics
Princeton University
Princeton, NJ 08544
and NBER
grossman@princeton.edu

Ezra Oberfield
Department of Economics
Princeton University
Fisher Hall
Princeton, NJ 08544
and NBER
edo@princeton.edu

Elhanan Helpman
Department of Economics
Harvard University
1875 Cambridge Street
Cambridge, MA 02138
and NBER
ehelpman@harvard.edu

Thomas Sampson
London School of Economics
t.a.sampson@lse.ac.uk

1 Introduction

Some key facts about economic growth have become common lore. Among those famously cited by Kaldor (1961) are the observation that output per worker and capital per worker have grown steadily, while the capital-output ratio, the real return on capital, and the shares of capital and labor in national income have remained fairly constant. Jones (2015) updates these facts using the latest available data. He reports that real per capita GDP in the United States has grown "at a remarkably steady average rate of around two percent per year" for a period of nearly 150 years, while the ratio of physical capital to output has remained nearly constant. The shares of capital and labor in total factor payments were very stable from 1945 through about 2000.[1]

These facts suggest to many the relevance of a "balanced growth path" and thus the need for models that predict sustained growth of output, consumption and capital at constant rates. Indeed, neoclassical growth theory was developed largely with this goal in mind. Apparently, it succeeded. As Jones and Romer (2010, p.225) conclude: "There is no longer any interesting debate about the features that a model must contain to explain [the Kaldor facts]. These features are embedded in one of the great successes of growth theory in the 1950s and 1960s, the neoclassical growth model."

Alas, "all is not well," as Hamlet might say. Jones (2015) highlights yet another fact that was noted earlier by Gordon (1990), Greenwood et al. (1997), Cummins and Violante (2002), and others: the relative price of capital equipment, adjusted for quality, has been falling steadily and dramatically since at least 1960. Figure 1 reproduces two series from FRED (Federal Reserve Economic Data, a database maintained by the Federal Reserve Bank of St. Louis).[2] In the period from 1947 to 2013, the relative price of investment goods has fallen at a compounded average rate of 2.0 percent per annum. The relative price of equipment has fallen at an even faster annual rate of 3.8 percent.

This observation of falling capital prices rests uncomfortably with the features of the economy that are thought to be needed to foster balanced growth. As Uzawa (1961) pointed out, and Schlicht (2006) and Jones and Scrimgeour (2008) later clarified, a balanced growth path in the two-factor neoclassical growth model with a constant and exogenous rate of population growth and a constant rate of labor-augmenting technological progress requires either an aggregate production function with a unitary elasticity of sub-

[1] As is well known from Piketty (2014) and many others before him and since, the share of capital in national income has been rising, and that of labor falling, since around 2000; see, for example, Elsby et al. (2013), Karabarbounis and Neiman (2014), and Lawrence (2015). It is not clear yet whether this is a temporary fluctuation around the longstanding division, part of a transition to a new steady-state division, or perhaps (as Piketty asserts) a permanent departure from stable factor shares.

[2] The FRED data for investment and equipment prices are based on updates of Gordon's (1990) numbers by Cummins and Violante (1990) and DiCecio (2009, Appendix A).

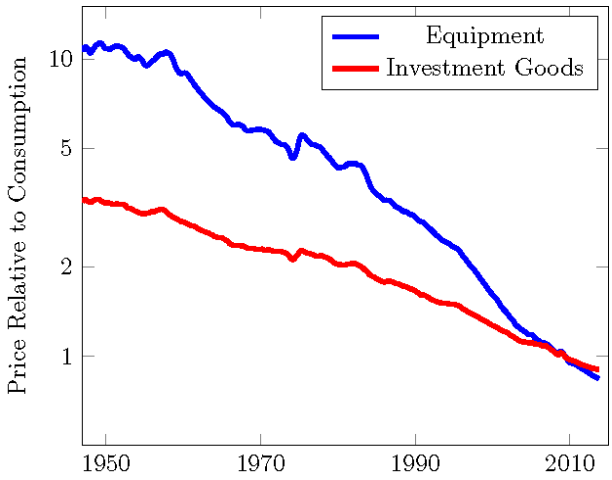

Figure 1: U.S. Relative Price of Equipment, 1947-2013
Source: Federal Reserve Bank Economic Data (FRED), Series PIRIC and PERIC.

stitution between capital and labor or else an absence of any capital-augmenting technological progress. The size of the elasticity of substitution between capital and labor is much debated and still controversial. Yet, a preponderance of the evidence suggests an elasticity well below one.[3] And the fact that the quality-adjusted prices of investment goods (and especially equipment) have been falling relative to the price of final output suggests that the rate of (embodied) capital-augmenting technological progress has not been nil.[4]

The Uzawa Growth Theorem rests on the impossibility of getting an endogenous rate of capital accumulation to line up with an exogenous growth rate of effective labor in the presence of capital-augmenting technological progress, unless the aggregate production function takes a Cobb-Douglas form. The "problem," it would seem, stems from the model's assumption of an inelastic supply of effective labor that does not adjust to capital deepening, even over time. If human capital could be accumulated endogenously, via investments in schooling, on-the-job training, or otherwise, then perhaps effective labor growth would fall into line with growth in effective capital, and a balanced growth path would be possible in a broader set of circumstances. Seen in this light, another fact about the U.S. growth experience appears to offer a way out. We reproduce—as did Jones (2015)—a figure from Goldin and Katz (2007).

[3] Chirinko (2008, p.671), for example, who surveyed and evaluated a large number of studies that attempted to measure this elasticity, concluded that "the weight of the evidence suggests a value of [the elasticity of substitution] in the range of 0.4 to 0.6." In research conducted since that survey, Karabarounis and Nieman (2014) estimate an elasticity of substitution greater than one, but Chirinko et al. (2011), Oberfield and Raval (2014), Chirinko and Mallick (2014), Herrendorf, et al. (2015), and Lawrence (2015) all estimate elasticities below one.

[4] Motivated by Uzawa's Growth Theorem, Acemoğlu (2003) and Jones (2005) provide theories of directed technical change in order to provide an explanation for the absence of capital-augmenting technical change. To be consistent with balanced growth, both look for restrictions that would lead endogenous technical change to be entirely labor-augmenting. Neither attempts to reconcile capital-augmenting technical change with balanced growth.

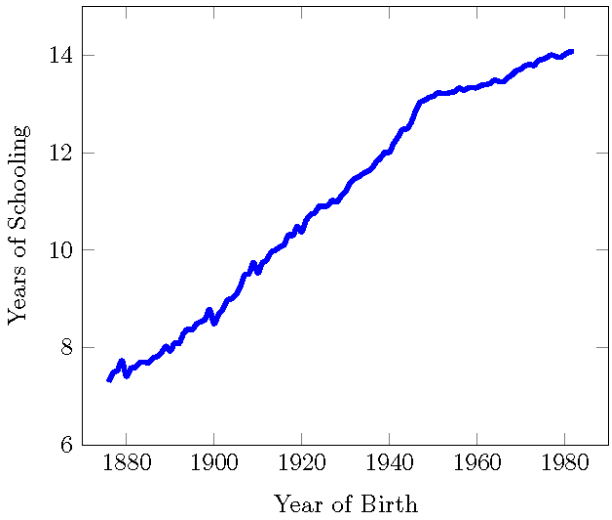

Figure 2: U.S. Education by Birth Cohort, 1876-1982
Source: Goldin and Katz (2007) and additional data from Lawrence Katz.

Figure 2 shows the average years of schooling measured at age thirty for all cohorts of native American workers born between 1876 and 1982.[5] Clearly, educational attainment has been rising steadily for more than a century. Put differently, there has been ongoing investment in "human capital." Indeed, Uzawa (1965), Lucas (1988), and others have established the existence of a balanced growth path in a neoclassical growth model that incorporates a standard treatment of human capital accumulation, albeit in settings that lack embodied or disembodied capital-augmenting technological progress.[6]

Unfortunately, the usual formulation of human capital does not do the trick. In the next section, we prove an extended version of the Uzawa Growth theorem that allows for accumulation of human capital. We specify an aggregate production function that has effective capital (the product of physical capital and a productivity-augmenting technology term) and human capital as arguments. Human capital is represented as an arbitrary function of technology-augmented "raw labor" and a variable that measures private investments in upgrading the labor input. In this setting, we show again that balanced growth requires either a unitary elasticity of substitution between physical capital and human capital, or else an absence of capital-augmenting technological progress. The intuition is similar to that provided by Jones and Scrimgeour for the original Uzawa theorem. Along a balanced growth path, physical capital that is produced from final goods inherits the trend in output growth.[7] But the growth rate of final

[5] We are grateful to Larry Katz for providing the unpublished data that allowed us to extend his earlier figure.

[6] Uzawa (1965) studies a model with endogenous accumulation of human capital in which education augments "effective labor supply" so as to generate convergence to a steady state. Lucas (1988) incorporates an externality in his measure of human capital, a possibility that we do not consider here. Acemoğlu (2009, pp. 371-374) characterizes a balanced growth path in a setting with overlapping generations.

[7] If the price of investment goods relative to consumption can change—something Jones and Scrimgeour did not consider—

output is a weighted average of the growth rates of effective capital and effective labor, with factor shares as weights. If these shares are to remain constant along a balanced growth path with an aggregate production function that is not Cobb-Douglas, then effective capital and effective labor must grow at common rates. It follows that the growth rate of output also mirrors the growth rate of effective capital. With the growth rate of final output equal to both the growth rate of (the value of) physical capital and the growth rate of effective capital, there is no room for capital productivity to improve or for the cost of investment to fall. And all of this is true whether effective labor grows partly due to endogenous investment in human capital or not.

But our findings in Section 2 also point to a way out of the bind. Ongoing increases in educational attainment such as those depicted in Figure 2 can potentially reconcile the existence of a balanced growth path with a sustained rise in capital or investment productivity and an elasticity of substitution between capital and labor less than unity, provided that schooling enters the aggregate production function differently than raw labor. Then investments in schooling can offset the change in the capital share that results from capital deepening (growth in effective capital relative to technology-augmented raw labor). It is possible—with just the right steady gains in education—for balanced growth to occur, with output and the value of capital growing at the same rates, effective capital growing at a faster rate than technology-augmented labor, and an index of schooling rising over time to keep the factor shares constant.

To be more precise, suppose that $F(K, L, s; t)$ is the output that can be produced with the technology available at time t by L units of "raw labor" working with K units of physical capital, when the economy has an education level summarized by the scalar measure s. Suppose that $F(\cdot)$ has constant returns to scale in K and L and that $\sigma_{KL} < 1$, where $\sigma_{KL} \equiv F_L F_K / F F_{LK}$ is the elasticity of substitution between capital and labor, holding schooling constant. We will show that a balanced growth path with constant factor shares, a growing index of education level, and positive capital-augmenting technological progress (embodied or disembodied) can emerge, but only if the ratio of the marginal product of schooling to the marginal product of labor rises as capital accumulates; i.e., $\partial (F_s/F_L)/\partial K > 0$. Clearly, this precludes a production function of the form $F(K, H; t)$, where $H = G(L, s; t)$ is a standard measure of human capital at time t, because then F_s/F_L is independent of K. A necessary condition for balanced growth in the presence of capital-augmenting technological progress and a non-unitary elasticity of substitution is a sufficient degree of complementarity between capital and education. Of course, many researchers have noted the empirical relevance of "capital-skill complementarity" (see, most prominently, Krusell, et al.,

the analogous requirement is that the *value* of the capital stock inherits the growth rate of output.

2000 and Autor, et al., 1998), albeit with varying interpretations of the word "skill" and of the word "complementarity." Our analysis makes clear that the appropriate sense of complementarity is a relative one: growth in the capital stock must raise the marginal productivity of schooling relatively more than it does the marginal productivity of raw labor. Moreover, if $\sigma_{KL} < 1$, then balanced growth requires that the technology $F(K, L, s; t)$ be characterized by *strict log supermodularity* in K and s, which is a stronger sense of complementarity than $F_{Ks} > 0$.

The fact that schooling gains *can* offset the effects of capital-augmenting technological progress on the capital share does not of course mean that they *will* do so in a reasonable model of schooling decisions. So we proceed in the subsequent sections to introduce optimizing behavior. In Section 3, we keep things simple at the cost of realism. We first solve a social planner's resource-allocation problem that incorporates a reduced-form specification of the trade-off between an index of an economy's schooling level and its available labor supply. The key simplifying assumptions in this section are that an economy's schooling can be represented by a scalar measure and that this choice variable can jump from one moment to the next. Under these assumptions, when the aggregate production function belongs to a specified class, the optimal growth trajectory converges to a balanced-growth path with constant rates of growth of output, consumption and capital, and a constant capital share in national income. Following the presentation of the planner's problem, we present two distinct models in which the market equilibrium shares the dynamic properties of the efficient solution. In both models, the economy is populated by a continuum of similar dynasties, each comprising a sequence of family members who survive for only infinitessimal lifespans. In the "time-in-school" model of Section 3.2, each individual decides what fraction of her brief existence to devote to schooling, thereby determining her productivity in her remaining time as a worker. Firms allocate capital to their various employees as a function of their productivity levels and therefore their schooling. In the "manager-worker" model of Section 3.3, individuals instead make a discrete educational choice. Those who devote a fixed fraction of their life to schooling are trained to work as managers with their remaining time. Those who do not opt for management training have their full life to serve as production workers. In this case, our measure of the economy's education is its ratio of manager hours to worker hours, and we assume that productivity of a production unit (workers combined with equipment) rises with this ratio due to improved monitoring. In both models the economy converges to a balanced-growth path for a specified class of production functions, all of whose members are characterized by stronger complementarity between capital and schooling than between capital and technology-augmented labor.

Section 4 adds features to the time-in-school model that make it more realistic. There, we allow the dynasties to comprise overlapping generations of finitely-lived family members. Each individual devotes the first part of her life to school and chooses a stopping date to enter the workforce so as to maximize the dynasty's utility. Once an individual begins working, productivity initially rises and ultimately falls with experience. Death happens stochastically according to a Poisson process. If the individual survives a sufficiently long career, eventually her productivity falls to zero and she "retires." In this setting, different birth cohorts make different education decisions, and so "schooling" does not have a scalar representation. Both an individual's education attainment and the distribution of education levels in the workforce are state variables that adjust gradually over time.

For a range of parameter values, the overlapping-generations model—like its counterpart with non-overlapping generations—admits a balanced-growth path for a class of production functions that has $\sigma_{KL} < 1$, even with ongoing capital-augmenting technological progress. On the balanced-growth path, the value of capital grows at the same rate as the value of output, the productivity-augmented capital stock grows faster than technology-augmented labor, educational attainment by birth cohort rises linearly with time, labor-force participation trends downward, and both aggregate factor shares and the real interest rate are constant. The growth rate of per capita output is increasing in the rate of labor-augmenting technological progress and the rate of capital-augmenting technological progress. Although we have no analytical result for the long-run effects of an acceleration or deceleration of technical change on income distribution, plausible parameter values selected to approximate those in the U.S. economy suggest that a slowdown in either form of technological progress will raise the capital share in national income.

Section 5 contains some concluding remarks.

2 The Extended Uzawa Growth Theorem and a Possible Way Out

In this section, we state and prove a version of the Uzawa Growth Theorem, following Schlicht (2006) and Jones and Scrimgeour (2008), and extend it to allow for falling investment-good prices and the possible accumulation of human capital. We also show how investments in schooling can loosen the straitjacket of the theorem, but only if capital accumulation boosts the marginal product of schooling proportionally more than it does the marginal product of raw labor.

Let $Y_t = F(A_t K_t, B_t L_t, s_t)$ be a standard neoclassical production function with constant returns to scale in its first two arguments, where, as usual, Y_t is output, K_t is capital, L_t is labor, and where A_t

and B_t characterize the state of (disembodied) technology at time t, augmenting respectively the physical capital stock and the "raw" labor force.[8] We take s_t to be a scalar variable representing the education level in the economy.

At time t, the economy can convert one unit of output into q_t units of capital. Growth in q_t represents what Greenwood et al. (1997) have called "investment-specific technological change." This is a form of *embodied* technical change—familiar from the earlier work of Johansen (1959), Solow (1960) and others—inasmuch as new capital goods require less foregone consumption than did prior vintages of capital. The economy's resource constraint can be written as

$$Y_t = C_t + I_t/q_t \ ,$$

where C_t is consumption and I_t is the number of newly-installed units of capital. Investment in new capital augments the capital stock after the replacement of depreciation, which occurs at a fixed rate δ; i.e.,

$$\dot{K}_t = I_t - \delta K_t.$$

We begin with a lemma that extends slightly the one proved by Jones and Scrimgeour (2008) by incorporating ongoing investment-specific technological progress. Define a balanced-growth path (BGP) as a trajectory along which the economy experiences constant proportional rates of growth of Y_t, C_t, and K_t after some time T. Let $g_X = \dot{X}_t/X_t$ denote the growth rate of the variable X along the BGP. We have

Lemma 1 *Suppose g_q is constant. Then in any BGP with $C_t < Y_t$, $g_Y = g_C = g_K - g_q$.*

The proof, which closely follows Jones and Scrimgeour, is relegated to the appendix. The lemma states that the growth rates of consumption and the capital stock mirror that of total output. However, with the possibility of investment-specific technological progress, it is the value of the capital stock measured in units of the final good (and the resources used in investment) that grows at the same rate as output.[9]

Now define $\gamma_K \equiv g_A + g_q$. This can be viewed as the *total* rate of capital-augmenting technological change, combining the rate of disembodied progress (g_A) and the rate of embodied progress (g_q). Also,

[8] For ease of exposition and for comparability with the literature, we treat technology as a combination of components that augment physical capital and raw labor. However, as we show in the appendix, our Proposition 1 can readily be extended to any constant-returns to scale production function with the form $F(K_t, L_t, s_t; t)$.

[9] When capital goods are valued, their price p_t in terms of final goods must equal the cost of new investment, i.e., $p_t = 1/q_t$.

define, as we did before, $\sigma_{KL} \equiv (F_L F_K) / (F_{LK} F)$ to be the elasticity of substitution between capital and labor holding fixed the level of schooling. In the appendix we prove

Proposition 1 *Suppose that investment-specific technological progress occurs at constant rate g_q. If there exists a BGP along which the income shares of capital and labor are constant and strictly positive when factors are paid their marginal products, then*

$$(1 - \sigma_{KL}) \gamma_K = \sigma_{KL} \frac{F_L}{F_K} \frac{\partial (F_s / F_L)}{\partial K} \dot{s} \ . \tag{1}$$

The proposition stipulates a relationship between the combined rate of capital-augmenting technological progress and the change in schooling per worker that is needed to keep factor shares constant as the value of the capital stock and output grow at common rates.

We can now revisit the two cases that are familiar from the literature. First, suppose that there are no opportunities for investment in schooling, so that s remains constant. This is the setting considered by Uzawa (1961). Setting $\dot{s} = 0$ in (1) yields

Corollary 1 (Uzawa) *Suppose that s is constant. Then a BGP with constant and strictly positive factor shares can exist only if $\sigma_{KL} = 1$ or $\gamma_K = 0$.*

As is well known, balanced growth in a neoclassical economy without education requires either a Cobb-Douglas production function or an absence of capital-augmenting technological progress.[10]

Second, suppose that (effective) labor and schooling can be aggregated into an index of "human capital," $H(BL, s)$, such that net output can be written as a function of effective physical capital and human capital, as in Uzawa (1965), Lucas (1988), or Acemoğlu (2009). Denote this production function by $\tilde{F}[AK, H(BL, s)] \equiv F(AK, BL, s)$. Then $F_s / F_L = H_s / H_L$, which is independent of K. Setting $\partial (F_s / F_L) / \partial K = 0$ in (1) yields

Corollary 2 (Human Capital) *Suppose that there exists a measure of human capital, $H(BL, s)$, such that $F(AK, BL, s) \equiv \tilde{F}[AK, H(BL, s)]$. Then a BGP with constant and strictly positive factor shares can exist only if $\sigma_{KL} = 1$ or $\gamma_K = 0$.*

[10] Our Proposition 1 is predicated on constant and interior factor shares. But, in the Uzawa case, log differentiation of the production function with to respect to time, holding s constant, implies

$$g_Y = \theta_K (g_A + g_K) + (1 - \theta_K)(g_B + n)$$

where $\theta_K = AKF_K / Y$ is the capital share in national income. In a steady state in which Y and K grow at constant rates in response to constant rates of growth of A, B, L and q, θ_K must be constant as well. Note that Jones and Scrimgeour do not assume constant factor shares in their statement and proof of the Uzawa Growth Theorem.

In this case, ongoing accumulation of human capital cannot perpetually neutralize the effects of capital deepening on the factor shares.

However, Proposition 1 suggests that balanced growth with constant factor shares might be possible despite a non-unitary elasticity of substitution between capital and labor and the presence of capital-augmenting technological progress, so long as $\dot{s} \neq 0$ and $\partial \left(F_s / F_L \right) / \partial K \neq 0$. Suppose, for example, that $\sigma_{KL} < 1$, as seems most consistent with the empirical literature. Suppose further that educational attainment grows over time, again in line with observation. Then the existence of a BGP with constant factor shares requires $\partial \left(F_s / F_L \right) / \partial K > 0$; i.e., an increase in the capital stock must raise the marginal product of schooling by proportionally more than it does the marginal product of raw labor. In looser parlance, the technology must be characterized by "capital-skill complementarity," or by a "skill bias" in the capital-augmenting technological change.

The results in this section use only resource constraints (i.e., accounting) and the assumption that factors are paid their marginal products. We have, as yet, provided no model of savings, of investment, or of schooling decisions. Moreover, we have shown that a BGP with constant factor shares *might exist*, but not that one *does exist* under some reasonable set of assumptions about individual behavior and a reasonable specification of the aggregate production function. These are our next tasks, which we will perform in two stages. First, we study a simple environment in which the economy's level of education can be summarized by a scalar variable that can jump discretely from one moment to the next. Then, we extend our analysis to a more realistic setting in which individuals' education accumulates slowly over time and the distribution of educational levels in the economy evolves gradually.

3 Balanced Growth with Short Lifespans

We begin this section by posing a social planner's problem that incorporates a reduced-form treatment of schooling choice. In Section 3.1, the planner designs a time path for a scalar variable that summarizes the level of education in the workforce. The planner faces a trade-off between the level of schooling and the labor supply available for producing output. The economy experiences both labor-augmenting and capital-augmenting technological progress, and the elasticity of substitution between capital and labor in aggregate production is less than one. Here we show that the planner's allocation converges to a unique BGP for a specified class of production functions and under certain parameter restrictions. Moreover, if the efficient allocation can be characterized by balanced growth after some moment in time, then the

technology *must* have a representation with a production function in the specified class. We derive the steady-state growth rate of output for the planner's solution and the associated (and constant) capital share in income.

In the succeeding subsections, we develop a pair of models of individual behavior and aggregate production that generate the reduced-form education function of Section 3.1. Both models feature a continuum of dynasties and a sequence of family members that survive only for fleetingly brief lives. Generations are replaced continuously by new ones that begin afresh, without prior schooling. In Section 3.2, the representative family member decides the fraction of her life to devote to school, thereby determining as well her availability for gainful employment. Workers produce with the capital allocated to them by competitive firms and their productivity on the job depends on their educational attainment. In Section 3.3, by contrast, individuals face a discrete choice between pursuing an education that leaves them "skilled" or having more time for work. Those who attend school ultimately are employed by firms as "managers," while those who remain unskilled serve as "production workers." The productivity of a production unit varies with the ratio of managers to workers, i.e., the inverse of the managers' span of control. In Section 3.4, we describe how the model can be extended to include directed technical change, in the manner introduced by Acemoglu (2003). We conclude the section with a brief discussion that relates our findings to the recent literature on investment-specific technological progress.

3.1 A Planner's Problem with a Reduced-Form Education Function

The economy comprises a continuum of identical family dynasties of measure one. Each family has a continuum N_t of members alive at time t, where N_t grows at the exogenous rate n. Dynastic utility at some time t_0 is given by

$$u(t_0) = \int_{t_0}^{\infty} N_t e^{-\rho(t-t_0)} \frac{c_t^{1-\eta} - 1}{1-\eta} dt \ , \tag{2}$$

where c_t is consumption per family member at time t and ρ is the subjective discount rate.

Consider the problem facing a social planner who seeks to maximize utility for the representative dynasty subject to a resource constraint, an evolving technology, and an ongoing trade-off between some aggregate measure of educational attainment and contemporaneous labor supply. Write this trade-off in reduced form as $L_t = D(s_t) N_t$, with $D'(s_t) < 0$ for all s_t, where s_t is a scalar index of schooling and L_t is labor supply. The production function takes the form $Y_t = F(A_t K_t, B_t L_t, s_t)$, where A_t again converts physical capital to "effective capital" in view of the disembodied technology available at time t, and

similarly B_t converts raw labor to effective labor. Assuming, as we do, that $F(\cdot)$ has constant returns to scale in its first two arguments, we can express this function in intensive form as $f(k_t, s_t) \equiv F(k_t, 1, s_t)$, where $f(\cdot)$ is output per effective worker and $k_t = A_t K_t / B_t L_t$ is the ratio of effective capital to effective labor. The economy can convert one unit of the final good into q_t units of capital at time t. Capital depreciates at the constant rate δ.

We assume that the technology can be represented by a member of a class of production functions that take the following form.

Assumption 1 *The intensive production function can be written as* $f(k,s) = D(s)^{-\mu\beta} h[kD(s)^{\mu}]$, *with* $\mu > 0$ *and* $\beta \in (0,1)$, *where*

(i) $h(z)$ *is strictly increasing, twice differentiable, and strictly concave for all* $z \equiv kD(s)^{\mu} \geq 0$; *and*

(ii) $f(k,s)$ *is strictly log supermodular in* k *and* s.

Assumption 1 immediately implies that $\sigma_{KL} < 1$ and that $\partial(F_s/F_L)/\partial K > 0$.[11] Therefore, the technology satisfies the pre-requisites for the existence of a BGP, per Proposition 1, provided that the planner's optimal choice of schooling is rising over time.

We also impose some parameter restrictions. Let $\mathcal{E}_h(z) \equiv zh'(z)/h(z)$ be the elasticity of the $h(\cdot)$ function. Note that $\mathcal{E}_h(z)$ is strictly decreasing under Assumption 1.[12] Now define $d_{\max} \equiv \lim_{z\to 0} \mathcal{E}_h(z)$ and $d_{\min} \equiv \lim_{z\to\infty} \mathcal{E}_h(z)$. We adopt

Assumption 2 *(i)* $\beta \geq d_{\max}$; *(ii)* $\frac{\mu\beta - 1}{\mu - 1} \in (d_{\min}, d_{\max})$.

Part (i) of Assumption 2 ensures that the marginal product of schooling is non-negative for all levels of k and s.[13] Part (ii) guarantees that $\mu\beta > 1$ and that the optimal schooling choice is positive, as we will see below. To provide an example of a technology that satisfies Assumption 1, we can choose $h(z) = (1 + z^{-\alpha})^{-\beta/\alpha}$, with $\alpha > 0$, which results from a production function of the form $F(AK, BL, s) = (BL)^{1-\beta}\left\{(AK)^{-\alpha} + \left[D(s)^{-\mu} BL\right]^{-\alpha}\right\}^{-\beta/\alpha}$. In this case, $\mathcal{E}_h(z) = \beta/(1 + z^{\alpha})$. Clearly, $\mathcal{E}_h(z)$ is declining in z, and we have $d_{\min} = 0$ and $d_{\max} = \beta$.

[11] See the proof in the appendix.

[12] To see this, note that $d\mathcal{E}_h(z)/dz \propto -[\mathcal{E}_h(z) - \mathcal{E}_{h'}(z) - 1]$, where $\mathcal{E}_{h'}(z) \equiv zh''(z)/h'(z)$ is the elasticity of $h'(z)$. Using $f(k,s) = D(s)^{-\mu\beta} h[kD(s)^{\mu}]$, $D'(s) < 0$, and the fact that $f(k,s)$ is strictly log supermodular if and only if $f_{ks} f > f_k f_s$, it follows readily that $d\mathcal{E}_h(z)/dz < 0$.

[13] Assumption 1 implies $f_s(k,s) = \mu h(z) D(s)^{-\mu\beta - 1}[\beta - \mathcal{E}_h(z)]$, which is non-negative for all k and s if and only if $d_{\max} \leq \beta$.

We can write the planner's problem as

$$\max_{\{c_t,s_t\}} \int_{t_0}^{\infty} N_t e^{-\rho(t-t_0)} \frac{c_t^{1-\eta}-1}{1-\eta} dt$$

subject to

$$\begin{aligned} Y_t &\leq B_t L_t D(s_t)^{-\mu\beta} h\left[\frac{A_t K_t}{B_t L_t} D(s_t)^{\mu}\right] \ ; \\ L_t &= D(s_t) N_t \ ; \\ \dot{K}_t &= q_t (Y_t - N_t c_t) - \delta K_t \ . \end{aligned}$$

where the first constraint describes the technology at time t in view of Assumption 1, the second captures the trade-off between education and labor supply, and the last reflects the resource constraint that governs capital accumulation. The planner takes the initial capital stock, K_{t_0}, as given.

Substituting for $L_t = D(s_t) N_t$, we can re-write the first constraint as

$$Y_t \leq B_t N_t D(s_t)^{-(\mu\beta-1)} h\left[\frac{A_t K_t}{B_t N_t} D(s_t)^{\mu-1}\right] \ .$$

Now, since the schooling variable does not appear in the maximand or in the capital-accumulation equation, it is clear that the planner should choose s_t at every t to maximize contemporaneous output. The first-order condition $\partial Y_t / \partial s_t = 0$ implies

$$-(\beta\mu - 1) h\left[\frac{A_t K_t}{B_t N_t} D(s_t)^{\mu-1}\right] + (\mu - 1) h'\left[\frac{A_t K_t}{B_t N_t} D(s_t)^{\mu-1}\right] \frac{A_t K_t}{B_t N_t} D(s_t)^{\mu-1} = 0 \ ,$$

or[14]

$$\mathcal{E}_h\left[k_t D(s_t)^{\mu}\right] = \frac{\mu\beta - 1}{\mu - 1} \text{ for all } t \ . \tag{3}$$

In other words, the planner chooses education at every moment in time so that $z_t \equiv k_t D(s_t)^{\mu}$ remains constant. In this sense, the planner offsets (effective) capital deepening by increases in schooling.

[14] Note that

$$\begin{aligned} \frac{A_t K_t}{B_t N_t} D(s_t)^{\mu-1} &= \frac{A_t K_t}{B_t L_t} D(s_t)^{\mu} \\ &= k_t D(s_t)^{\mu} \ . \end{aligned}$$

Let z^* denote the optimal (and time invariant) value of z_t. Part (ii) of Assumption 2 ensures that there exists a solution for z^* and the fact that $\mathcal{E}_h(z)$ is strictly decreasing implies that the solution is unique.[15]

Once we have $z_t = z^*$, we can use Assumption 1 to solve for aggregate output as a function of the capital stock, the population size, and the state of technology. We find

$$Y_t = (B_t N_t)^{\frac{\mu(1-\beta)}{\mu-1}} (A_t K_t)^{\frac{\mu\beta-1}{\mu-1}} z^{*\frac{1-\mu\beta}{\mu-1}} h(z^*) \ . \tag{4}$$

Notice that (4) is a Cobb-Douglas function of effective capital and technology-augmented population, with exponents $\theta \equiv (\mu\beta - 1)/(\mu - 1)$ and $1 - \theta$, respectively. Now substituting for Y_t in the planner's constraints yields a standard and familiar dynamic optimization problem. As usual, we need the discount rate to be sufficiently large so that the integral in the maximand is bounded. In particular, we invoke

Assumption 3 $\rho > n + (1-\eta)\left[\gamma_L + \frac{\mu\beta-1}{(1-\beta)\mu}\gamma_K\right]$.

Assumption 3 ensures that the transversality condition for the dynamic optimization will be satisfied.

We will not rehearse the details of the transition path; these are familiar from neoclassical growth theory. In the appendix, we show that the planner chooses the initial per capita consumption level, c_{t_0}, so as to put the economy on the unique saddle path that converges to a steady state. On the BGP, consumption and output grow at constant rate g_Y and the capital stock grows at constant rate g_K.

We can readily calculate the growth rates of output and consumption along the BGP. From (3), we have

$$(\mu - 1) g_D + g_A + g_K - \gamma_L - n = 0$$

for all $t \geq t_0$. Noting that $Y_t = B_t N_t D(s_t)^{-(\mu\beta-1)} h(z^*)$, we also have

$$g_Y = \gamma_L + n - (\mu\beta - 1) g_D$$

along the optimal path. Finally, combining these two equations and using Lemma 1—which requires that $g_Y = g_K - g_q$ along any BGP—we find that $g_D = -\gamma_K/\mu(1-\beta)$ and $g_Y = n + \gamma_L + \gamma_K(\mu\beta - 1)/\mu(1-\beta)$ in the steady state, where $\gamma_K \equiv g_A + g_q$, as before.

[15] In the appendix, we show that the second-order condition is satisfied at $z_t = z^*$ under Assumption 1. Moreover, we show that the second-order condition would be violated if $f(k, s)$ were not log supermodular or, equivalently in this setting, if the elasticity of substitution between effective capital and effective labor exceeds one.

The growth of per capita income is increasing in the rate of labor-augmenting technological progress, just as in the neoclassical growth model without endogenous schooling. But now a BGP exists even when there is ongoing capital-augmenting technological progress or when the price of investment-goods is falling at a constant rate. Assumption 2 guarantees that $\mu\beta > 1$.[16] Therefore, the growth rate of per capita income also is increasing in γ_K, the combined rate of embodied and disembodied capital-augmenting progress.

We have not as yet introduced any market decentralization, which we will do only for the specific models described in Sections 3.2 and 3.3 below. However, in anticipation that capital will be paid its marginal product in a competitive equilibrium, we can define the capital share in national income at time t as $\theta_{Kt} = (\partial Y_t/\partial K_t)\, K_t/Y_t$. Using (4), we see that $\theta_{Kt} = (\mu\beta - 1)/(\mu - 1) \equiv \theta$ for all $t \geq t_0$. That is, the planner chooses the trajectories for the capital stock and schooling such that the capital share remains constant, both along the transition path and in the steady state. Notice that the growth rate and the capital share both are increasing in μ and β; in this sense, fast growth and a high capital share go hand in hand.

For future reference, we summarize our findings in the following proposition.

Proposition 2 *Suppose there is a trade-off between labor supply and a summary measure of schooling given by $L_t = D(s_t)\, N_t$. Let Assumptions 1, 2, and 3 hold. Then along the optimal trajectory from any initial capital stock, K_{t_0}, the economy converges to a BGP. On the BGP,*

(i) aggregate output and aggregate consumption grow at the common rate $g_Y = n + \gamma_L + \frac{\mu\beta - 1}{(1-\beta)\mu}\gamma_K$;

(ii) schooling evolves such that $g_D = -\frac{\gamma_K}{\mu(1-\beta)}$;

(iii) the capital share is constant and equal to $\theta_K = \frac{\mu\beta - 1}{\mu - 1}$.

We offer now some remarks about the role of Assumption 1. This assumption restricts the form of the intensive production function. But we could as well have made an assumption directly about the gross output function, $F(AK, BL, s)$. Then we would have stipulated that this function takes the form $\tilde{F}\left[AKD(s)^a, BLD(s)^{-b}\right]$ for some quasi-concave function $\tilde{F}(\cdot)$ with constant-returns to scale in the two arguments and some $a > 0$ and $b > 0$. Written in this way, $h[kD(s)^\mu]$ is equivalent to $\tilde{F}[kD(s)^\mu, 1]$,

[16] Assumption 1(i) implies $d_{\min} \geq 0$. So, Assumption 2 implies $(\mu\beta - 1)/(\mu - 1) > 0$. Thus, if $\mu > 1$, $\mu\beta > 1$. Suppose $\mu < 1$ and $\mu\beta < 1$. Then Assumption 2(i) and Assumption 2(ii) imply $(\mu - 1)\beta < (\mu\beta - 1)$, which in turn implies $\beta > 1$. This contradicts Assumption 1.

so we would need assumptions about $\tilde{F}(\cdot)$ that are equivalent to Assumption 2(i) and (ii). Clearly, we would have $a = \mu(1-\beta)$ and $b = \mu\beta$.

Evidently, schooling enters the gross output function in a way that augments the productivity of labor while diminishing the productivity of capital.[17] Of course, with the analog to Assumption 2(i), the combined effect of schooling on gross output is positive. The decline in the productivity of capital is just what is needed, along the BGP, to keep the schooling-plus-technology augmented capital stock growing in line with output. To see that this is so, notice that $D(s)^a A_t q_t$ is constant along the BGP. The effect of the optimal schooling is as if to neutralize the effect of the capital-augmenting progress and the declining investment-good prices on the growth of the effective capital stock.

One might wonder whether we are able to dispense with the functional-form restriction of Assumption 1. The answer to this question is no. In the appendix, we prove that if $L_t = D(s_t) N_t$ and if the solution to the social planner's problem exhibits balanced growth after some time T with increasing schooling and a constant capital share $\theta_K \in (0,1)$, then either there is no capital-augmenting technological progress ($\gamma_K = 0$) or else the technology can be represented along the equilibrium trajectory by a production function with the form $\tilde{F}\left[A_t K D(s)^a, B_t L D(s)^{-b}\right]$, with $a > 0$ when $\gamma_K > 0$ and $b = 1 + a\theta_K/(1-\theta_K)$. In other words, Assumption 1 is not only sufficient for the existence of a BGP with $\gamma_K > 0$ and $\sigma_{KL} < 1$, but it is essentially necessary as well. As with any model that generates balanced growth, knife-edge restrictions are required to maintain the balance; our model is no exception to this rule.

3.2 Balanced Growth in a "Time-in-School" Model

We provide now a first example of a market economy that generates the reduced form described in Section 3.1. The competitive equilibrium of this economy mimics the planner's optimal allocation, and so the market economy converges to a BGP with the properties summarized in Proposition 2.

As above, the representative family has a continuum N_t of members at time t. Each life is fleetingly brief; an individual attends school for the first fraction of her momentary existence and then joins the workforce for the remainder of her life. The variable s_t now represents the fraction of life that the representative member of the generation alive at time t devotes to education; she spends the remaining fraction $1 - s_t$ working. In this case, $D(s) = 1 - s$, so that the family's labor supply is $L_t = N_t(1 - s_t)$. Given the brevity of life, there is no discounting of an individual's wages relative to her time in school.

[17] This observation should not be misinterpreted. Under Assumption 1(ii) that $f(k,s)$ is strictly log supermodular, it remains true that capital is more complementary to schooling than it is to labor, in the sense that F_s/F_L rises with capital accumulation.

But dynasties do discount the earnings (and well being) of future generations relative to those currently alive. Every new cohort starts from scratch with no schooling.

Each individual chooses her consumption, savings, and schooling to maximize total dynastic utility, which at time t_0 is given by (2). Each individual supposes that other family members in her own and subsequent generations will behave similarly. Savings are used to purchase units of physical capital, which are passed on within the family from one generation to the next. The N_t members of the representative dynasty collectively inherit K_t units of capital at time t, considering that the aggregate capital stock is fully owned by the population and there is a unit continuum of dynasties in the economy.

Firms produce output using capital, labor, and the technology available at the time. A firm that employs K_t units of physical capital and that hires L_t time units from workers with schooling s_t at time t produces $F\left(A_tK_t, B_tL_t, s_t\right) = \tilde{F}\left[A_tK_t\left(1-s_t\right)^{\mu(1-\beta)}, B_tL_t\left(1-s_t\right)^{-\mu\beta}\right]$ units of output. Then the intensive production function takes the form $f\left(k,s\right) = \left(1-s\right)^{-\mu\beta} h\left[k\left(1-s\right)^{\mu}\right]$. The functions $h\left(\cdot\right)$ and $f\left(\cdot\right)$ have the properties described in Assumption 1. The parameter restrictions in Assumptions 2 and 3 also apply. Aggregate output is simply the sum of the outputs produced by all firms.

The competitive firms take the rental rate per unit of capital, R_t, and the wage schedule per unit of time, $W_t\left(s_t\right)$, as given, where the latter conveys the competitive wage rate for a worker with schooling s_t. A firm that hires workers with this level of education chooses L_t and k_t to maximize $B_tL_t\left[f\left(k_t,s_t\right) - r_tk_t - w_t\left(s_t\right)\right]$, where $r_t \equiv R_t/A_t$ is the rental rate per effective unit of capital and $w_t\left(s_t\right) \equiv W_t\left(s_t\right)/B_t$ is the wage per effective unit of labor. Profit maximization implies, as usual, that

$$f_k\left(k_t,s_t\right) = r_t \tag{5}$$

and[18]

$$f\left(k_t,s_t\right) - r_tk_t = w_t\left(s_t\right). \tag{6}$$

We define the functions $\kappa\left(s,r\right)$ and $\omega\left(s,r\right)$ such that $f_k\left[\kappa\left(s,r\right),s\right] \equiv r$ and $\omega\left(s,r\right) \equiv f\left[\kappa\left(s,r\right),s\right] - r\kappa\left(s,r\right)$. Then, in equilibrium, $k_t = \kappa\left(s_t,r_t\right)$ and $w_t\left(s_t\right) = \omega\left(s_t,r_t\right)$.

Schooling choices have no persistence for the family. Therefore, an individual alive at time t who seeks to maximize dynastic utility should choose s to maximize her own wage income, $B_t\left(1-s\right)\omega\left(s,r_t\right)$, taking the rental rate per unit of effective capital as given. The rental rate will determine, via (5), how

[18] Equation (6) is the zero-profit condition, which is implied by the optimal choice of L_t in an equilibrium with positive output.

much capital the individual will be allocated by her employer as a reflection of her schooling choice. The individual's education decision is separable from her choice of consumption, much as the planner's choice of s_t in Section 3.1 was separable from the choice of c_t and $\dot{K}_t$.

The first-order condition for income maximization at time t requires

$$(1 - s_t)\,\omega_s\,(s_t, r_t) = \omega\,(s_t, r_t)\,.$$

But using $\omega\,(s, r_t) \equiv f\,[\kappa\,(s, r_t)\,, s] - r_t\kappa\,(s, r_t)$ and noting (5), we have $\omega_s\,(s_t, r_t) = f_s\,[\kappa\,(s_t, r_t)\,, s_t]$. In other words, the marginal effect of schooling on the wage reflects only the direct effect of schooling on per capita output; the extra output that comes from a greater capital allocation to more highly educated workers, $f_k\kappa_s$, just offsets the extra part of revenue that the firm must pay for that capital, $r\kappa_s$. Consequently, we can rewrite the first-order condition as

$$(1 - s_t)\,f_s\,[\kappa\,(s_t, r_t)\,, s_t] = f\,[\kappa\,(s_t, r_t)\,, s] - f_k\,[\kappa\,(s_t, r_t)\,, s_t]\,\kappa\,(s_t, r_t)\ . \tag{7}$$

Now replace $f\,(k, s)$ by $(1 - s)^{-\mu\beta}\,h\,[k\,(1 - s)^{\mu}]$ and use this representation to calculate $f_s\,(\cdot)$ and $f_k\,(\cdot)$ as well. After rearranging terms, this yields

$$(\mu\beta - 1)\,h\,[\kappa\,(s_t, r_t)\,(1 - s_t)^{\mu}] = (\mu - 1)\,h'\,[\kappa\,(s_t, r_t)\,(1 - s_t)^{\mu}]\,\kappa\,(s_t, r_t)\,(1 - s_t)^{\mu}$$

or

$$\mathcal{E}_h\,[\kappa\,(s_t, r_t)\,(1 - s_t)^{\mu}] = \frac{\mu\beta - 1}{\mu - 1}\ .$$

Evidently, the individual's choice of schooling to maximize wage income matches the planner's choice of s_t in (3), once we recognize that $D\,(s_t) = 1 - s_t$ in the "time-in-school" model. Part (ii) of Proposition 2 then implies

$$\dot{s}_t = (1 - s_t)\,\frac{\gamma_K}{\mu\,(1 - \beta)}\ .$$

On a BGP, schooling rises over time, but at a declining rate.

A dynasty's intertemporal optimization also yields the same consumption and savings decisions as in the planner's problem. The family members adjust consumption in response to the real interest rate, ι_t, according to

$$\frac{\dot{c}_t}{c_t} = \frac{1}{\eta}\,(\iota_t - \rho)\ .$$

When combined with the intertemporal budget constraint and the no-arbitrage condition[19], $\iota_t = R_t/p_t + g_p - \delta$, where $p_t = 1/q_t$ is the equilibrium price of a unit of capital, this Euler equation generates the same time path for aggregate capital as in the planner's allocation; see the appendix for details.

Of course, it is no surprise that the market equilibrium with perfect competition and complete markets mimics the planner's solution. The point we wish to emphasize is that the time-in-school model converges to a BGP and that the wage schedule $\omega(s, r_t)$ gives the family members the appropriate incentives to extend their time in school from one generation to the next. Here, the faster accumulation of effective capital relative to effective labor sets in motion a sequence of events that preserves balance. An increase in effective capital lowers the rental rate. This causes the returns to education to rise, due to the complementarity between schooling and capital. With $\sigma_{KL} < 1$, the direct effect of the capital deepening is a fall in the capital share. But, since—as we have noted—Assumption 1 implies that we can write $Y_t = \tilde{F}\left[A_t K_t (1-s_t)^a, B_t L_t (1-s_t)^{-b}\right]$ with $a = \mu(1-\beta)$ and $b = \mu\beta$, schooling effectively augments the productivity of labor while diminishing that of capital. This in turn raises the capital share. While it is fairly natural that the accumulation of effective capital and the gains in education should have opposing effects on the capital share, the functional-form restrictions of Assumption 1 ensure that the scale is perfectly balanced and the capital share is constant and equal to $(\mu\beta - 1)/(\mu - 1)$.

3.3 Balanced Growth in a "Manager-Worker" Model

In Section 3.2, we described an environment in which individuals choose their time in school and education improves productivity. In that model, firms allocate capital equipment to individual workers and output is the sum of all that is produced by the various individuals. The model yields the same trade-off between education and labor supply that was captured in reduced form in the planner's problem of Section 3.1.

In this section, we present an entirely different model that yields a similar reduced form. Now we imagine teams that combine "managers" and "production workers." Firms allocate capital equipment to teams according to their productivity. Only production workers are directly responsible for operating equipment and thus for generating output. But the productivity of a team depends on its ratio of managers to workers, as in the hierarchical models of management proposed by Beckmann (1977), Rosen (1982) and others.

The family structure, demographics, and preferences are the same as before. Lifespans are short. Each individual decides whether to devote a fixed fraction m of her potential working life to school. If

[19] The no-arbitrage condition states that the real interest rate on a short-term bond equals the dividend rate on a unit of physical capital plus the rate of capital gain on capital equipment (positive or negative), minus depreciation.

she opts to do so, she will acquire the skills needed to serve as a manager and she will have $1-m$ units of time left to perform this function. Those who do not go for management training are employed as production workers. They will use all of their available time to earn unskilled wages.

Let L_t be the time units supplied by production workers at time t and let M_t be the time units supplied by managers. Since production workers devote all of their time to their jobs, L_t is also the number of production workers. Managers are in school a fraction m of their time, so the number of managers is $M_t/(1-m)$. The population divides between workers and managers, so

$$L_t + \frac{M_t}{1-m} = N_t \ . \tag{8}$$

We take $s_t = M_t/L_t$ to be our index of schooling. This is the ratio of managerial hours to hours of production workers and the inverse of the typical manager's "span of control." It measures, for example, the time that a manager can spend monitoring a typical one of her underlings. With this definition, (8) implies $L_t + L_t s_t/(1-m) = N_t$, so $D(s) = [1 + s/(1-m)]^{-1}$ is the share of production workers in the total population.

Monitoring makes the workers and their equipment more productive. In particular, we suppose that the production function at time t can be written as $\tilde{F}\left[D(s)^{\mu(1-\beta)} A_t K, D(s)^{-\mu\beta} B_t L\right]$, with $\tilde{F}(\cdot)$ homogeneous of degree one in its two arguments. With $s = M/L$, this implies that output is a constant-returns to scale function of the three inputs, $A_t K, B_t L$ and $B_t M$. It also implies that the intensive production function $f(k,s)$ has the form $D(s)^{-\mu\beta} h[kD(s)^{\mu}]$. An example of a production function with this form is

$$Y = (B_t L)^{1-\beta} \left[(A_t K)^{-\alpha} + D(s)^{\mu\alpha} (B_t L)^{-\alpha}\right]^{-\frac{\beta}{\alpha}} \ .$$

In this model, the education decision for the representative individual born at time t is simple: pursue schooling if $(1-m) W_{Mt} > W_{Lt}$ and not if the inequality runs in the opposite direction, where W_{Mt} and W_{Lt} are the market wages of managers and production workers at time t, respectively. In an equilibrium with a positive number of managers, every individual must be indifferent between the two occupations, so that

$$(1-m) W_{Mt} = W_{Lt} \ . \tag{9}$$

Over time, the accumulation of effective capital exerts upward pressure on the skill premium, because the functional form of Assumption 1 ensures that capital is more complementary with managers than it is

with production workers. This provides the incentive for a greater fraction of the new generation to gain skills and then the expanding relative supply of managers to workers restores the indifference condition, (9).

In the appendix, we use $W_{Mt} = \tilde{F}_M \left\{ [1 + s_t/(1-m)]^{-\mu(1-\beta)} A_t K_t, [1 + s_t/(1-m)]^{\mu\beta} B_t L_t \right\}$ and $W_{Lt} = \tilde{F}_L \left\{ [1 + s_t/(1-m)]^{-\mu(1-\beta)} A_t K_t, [1 + s_t/(1-m)]^{\mu\beta} B_t L_t \right\}$ to show that (9) implies

$$\mathcal{E}_h \left[k_t \left(1 + \frac{s_t}{1-m} \right)^{-\mu} \right] = \frac{\mu\beta - 1}{\mu - 1}.$$

This gives the same index of schooling as in the planner's solution (3). It follows that the economy converges to a BGP, with a constant rate of output growth and a constant capital share given by parts (i) and (iii) of Proposition 2, respectively, and with an ever increasing ratio of manager hours to worker hours.

3.4 Balanced Growth with Directed Technical Change

In this subsection, we describe briefly how the short-lifespans model can be extended to incorporate endogenous innovation and directed technical change. We follow closely the approach developed by Acemoglu (2003). In his model, final goods are assembled from two intermediate goods, each of which is produced from sets of upstream, differentiated varieties. One set of varieties is produced by capital alone, the other set by labor alone. Innovation takes the form of the invention of new varieties of one type of upstream intermediate or the other. Expansion in a set of upstream intermediates augments the productivity for assembling the associated downstream intermediate. A fixed stock of scientists conducts R&D. If they choose to invent new varieties of capital intermediates, then the resulting innovation generates capital-augmenting technology gains in the aggregate production function. If they invent new varieties of labor intermedates, then the innovation appears as labor-augmenting in the aggregate production function.

To extend our reduced-form model of Section 3.1 to incorporate directed technical change in the manner of Acemoglu (2003), let one type of intermediate be produced by L_t and the other by K_t, where $L_t = D(s_t) N_t$, as before. Each downstream intermediate good is a CES aggregate of the quantities used of the upstream differentiated intermediates. Then A_t reflects the measure of capital varieties available at time t and B_t reflects the measure of labor varieties; each is a power function of the measure of varieties, where the exponent is a function of the elasticity of substitution between the upstream goods. The

aggregate production function, expressed as a function of B_tL_t and A_tK_t, obeys Assumption 1 as before.

Now recall the "indirect production function" that we derived in (4). This indirect production function results when the planner chooses s_t to maximize aggregate output, Y_t. After substituting for L_t in the aggregate production function and using the optimal $D(s_t)$ prescribed by (3), it is a function of the technology-augmented population size, B_tN_t, and the technology-augmented capital stock, A_tK_t. Moreover, it is a *Cobb-Douglas* function of these two quantities. But, as Acemoglu has already shown, when the aggregate production function is Cobb-Douglas, there is a unique long-run equilibrium with balanced growth in which both capital-augmenting and labor-augmenting technological progress take place at well-determined rates. Moreover the growth rates of both types of intermediates will be positive if there is enough curvature in the trade-off between the two types of innovation. It follows that, in an extended version of the short-lifespans model with directed technical change, long-run growth is balanced, and the equilibrium is characterized by positive rates of both capital-augmenting and labor-augmenting technological progress.[20]

3.5 Relationship to the Literature on Investment-Specific Technological Change

Before leaving this section, it may be useful to relate our results to the large literature that has studied the long-run implications of investment-specific technological change. In his seminal paper on embodied technical progress, Solow (1960) did not close his model to solve for a steady state, but he indicated how this could be done. However, Solow employed a Cobb-Douglas production function throughout this paper, and his discussion about closing the model relies on this assumption. Sheshinski (1967) demonstrated convergence to a BGP in an extended version of the Johansen (1959) model with both embodied and disembodied technological progress. Although he does not restrict attention to any particular production function, he does insist that both forms of progress are Harrod-neutral, i.e., they augment the productivity of labor. So, the technology gains in Sheshinski's paper, while embodied in vintages of capital, are nonetheless assumed to be labor-augmenting. These findings are echoed in Greenwood et al. (1997), who resurrected the literature on technological improvements that are embodied in new equipment. They

[20] In (4), the exponents on A_tK_t and B_tN_t are $\theta = (\mu\beta - 1)/(\mu - 1)$ and $1 - \theta$, respectively. Borrowing the specfication of innovation from Acemoglu (2003), we obtain $\dot{A}_t/A_t = b_A\phi(S_{At})$ and $\dot{B}_t/B_t = b_B\phi(S_{Bt})$, where S_{At} and S_{Bt} are the number of scientists engaged in inventing capital-intensive varieties and labor-intensive varieties at time t, respectively. The allocation satisfies the resource constraint, $S_{At} + S_{Bt} = \bar{S}$, where $\bar{S}$ is the total supply of scientists. We assume that $\phi(0) = 0$ and that $\phi(\cdot)$ is concave. Since the planner can change the allocation of scientists at any moment in time without regard to preferences, the path of population growth, or the path of capital accumulation, it is optimal to allocate the stock of scientists so as to maximize the instantaneous rate of increase in total factor productivity, $A_t^{\theta}B_t^{1-\theta}$. The growth rate of TFP is $\theta\dot{A}t/A_t + (1-\theta)\dot{B}_t/B_t = \theta b_A\phi(S_A) + (1-\theta)b_b\phi(s_B)$. With sufficient concavity of $\phi(\cdot)$, the solution converges to a BGP with $S_A > 0$ and $S_B > 0$.

studied an economy that has no opportunities for schooling in which two types of capital ("equipment" and "structures") and labor are combined to produce consumption goods. Unlike Sheshinski, they do not assume that embodied progress is Harrod-neutral and, consequently, they are led to conclude that a Cobb-Douglas production function is necessary to generate balanced growth, in keeping with the dictates of the Uzawa Growth Theorem.

Krusell et al. (2000) posit a technology with capital-skill complementarity according to which output is produced from equipment, structures and two types of labor ("skilled" and "unskilled"). Leaving aside their distinction between equipment and structures, their model is one with capital and two types of labor, much like our manager-worker model in Section 3.3 above. Although their production function incorporates capital-skill complementarity, it does not satisfy the dictates of our Assumption 1. Nor do they endogenously determine the supplies of skilled and unskilled workers. They, and much of the substantial literature that has adopted their production function, do not address the prospects for balanced growth with ongoing declines in investment-good prices and endogenous schooling, but instead focus on the transition dynamics that result from a specified sequence of relative price changes and of factor supplies. Two recent papers do try to generate balanced growth in models of investment-specific technological progress that is not Harrod-neutral. He and Liu (2008) introduce endogenous schooling into the Krusell et al. model, so that the relative supplies of skilled and unskilled labor are determined in the general equilibrium. They define a BGP to be an equilibrium trajectory along which equipment, structures and output all grow at constant rates and the fraction of skilled workers converges to a constant. With this definition, they conclude (see their Proposition 1) that balanced growth is consistent with ongoing investment-specific technological change only when the aggregate production function takes a Cobb-Douglas form. Maliar and Maliar (2011) study a similar environment, but assume instead that the stocks of skilled and unskilled labor grow at constant and exogenous rates. They show that, with a falling relative price of equipment, balanced growth requires technological regress in the component of technical change that reflects the productivity of capital, such that (in our notation) $\gamma_K = 0$. In contrast to these papers, we have shown that balanced growth is in fact compatible with a falling relative price of capital, non-negative growth in capital productivity, and $\sigma_{KL} \neq 1$, provided that capital and schooling are sufficiently complementary. Our result requires that the aggregate production function falls into the class defined by Assumption 1 and that an appropriate index of the economy's educational outcome is rising over time.

4 Balanced Growth with Overlapping Generations

In Section 3, we illustrated how balanced growth could emerge in an economy with endogenous education. But we did so in a model of short lifespans in which each cohort lives for an instant and is replaced by the next without any overlap. In such a setting, it was possible to summarize the economy's education in a scalar variable and to allow that variable to jump from one moment to the next. This approach was pedagogically convenient, because it laid bare the mechanism at work. But our treatment of schooling was surely unrealistic, inasmuch as educational attainment typically varies by birth cohort and the distribution of educational outcomes adjusts slowly over time.

In this section, we introduce overlapping generations. We enrich the "time-in-school" model of Section 3.2 by assuming that individuals live for a finite (but stochastic) time, the first part of which they spend in school. A representative member of a cohort chooses at birth the duration of her tenure in the classroom and joins the labor force once her schooling is complete. We allow productivity to rise and then fall with experience, thereby capturing the employment life cycle that ultimately leads to retirement. Our goal once again is to uncover conditions that allow for a BGP with ongoing capital-augmenting technological progress and falling investment-goods prices, and to study the properties of such a growth path. We will find, for example, that in the overlapping-generations (OLG) model, the capital share in national income varies with the form and speed of technological change, unlike what we found to be true for an economy with short lifespans, where the capital share is independent of γ_K and γ_L.

A potential obstacle to our constructing a balanced growth path in an OLG model is that, if younger cohorts obtain more schooling and enter the workforce later in life than their more senior counterparts, the age distribution of the employment pool will not be stationary over time. As we show below, it turns out that the particular restriction on the production function that maintains balance between capital, labor, and schooling—the analogue to Assumption 1—leads naturally to an evolving age distribution in the workforce that retains a simple structure, thereby preserving balance across cohorts and facilitating aggregation.

As before, the economy is populated by a unit mass of identical dynasties.[21] A representative dynasty comprises a continuum N_t of individuals at time t. Each individual gives birth to a new member of her dynasty with an instantaneous probability λ and faces an instantaneous probability υ of death. These

[21] For continuity with the previous section and comparability with the literature, we continue to assume that families maximize dynastic utility, including the discounted well-being of unborn generations. We could obtain similar results, including the existence of a BGP, in a Yaari (1965) economy with (negative) life insurance and no bequests by following the path laid out by Blanchard and Fischer (1989, ch.3).

hazard rates remain constant over time. Therefore, the size of a dynasty is given by

$$N_t = e^{(\lambda - \upsilon)(t-t_0)} N_{t_0},$$

and the size at time t of the surviving cohort born at b is $\lambda N_b e^{-\upsilon(t-b)}$. The population growth rate is $n = \lambda - \upsilon$.

Conditional on survival, there are three phases of life: schooling, work, and retirement. An individual obtains s years of schooling, has a working life of $\bar{u}$ years after leaving school, and then retires. Let u be a worker's labor-market experience. We assume that a firm that employs K units of capital and L workers with schooling s and experience u produces output $F(AK, BL, s, u)$, where $F(AK, BL, s, u) = 0$ for $u \geq \bar{u}$. Thus, workers with experience beyond $\bar{u}$ cease to be productive and exit the labor market. The wage rate of an individual with schooling s and experience u at time t is $W_t(s, u)$. There is disembodied capital-augmenting technical change at rate g_A, labor-augmenting technical change at rate γ_L, and investment-specific technical change at rate g_q. The goods-market clearing condition $C_t + I_t/q_t = Y_t$ and the capital accumulation equation $\dot{K}_t = I_t - \delta K_t$ remain as before.

We assume that individuals must obtain their education at the beginning of their lives.[22] Each individual designs a "stopping rule," i.e., a duration s that she intends to remain in school conditional on survival. These choices are made to maximize the expected present discounted value of lifetime earnings, because that is optimal for the dynasty as a whole.[23] For an individual born at time b, expected discounted wage earnings at birth are given by

$$\int_{b+s}^{\infty} e^{-\int_b^t \iota_z dz} e^{-\upsilon(t-b)} W_t(s, t-b-s)\, dt = \int_0^{\bar{u}} e^{-\int_b^{b+s+u} \iota_z dz} e^{-\upsilon(s+u)} W_{b+s+u}(s, u)\, du\ , \tag{10}$$

where we have used the fact that an individual born at b who obtains s years of schooling has labor market experience $u = t - s - b$ at time t. Let s_b be the optimal schooling duration chosen by an individual born at b. Then a person born at b starts to work at time $b + s_b$ and retires at time $b + s_b + \bar{u}$. On the balanced growth path, educational attainment rises over time, so that the entry date $b + s_b$ is strictly increasing in b. We denote by $\Upsilon(\tau) = \tau - s_{\Upsilon(\tau)}$ the birth date of an individual who enters the workforce at time τ.

At time t_0 a representative dynasty chooses a future path of consumption $\{c_t \geq 0\}_{t_0}^{\infty}$ to maximize

[22] Blinder and Weiss (1976) have shown that models of life-cycle human-capital investments typically admit cycling with stretches in and out of school, unless the discount rate is sufficiently high or sufficiently low. Of course, the data show that most individuals concentrate their formal education at the beginning of life. To avoid the complications of (unrealistic) cycling, we assume that the education technology requires an uninterrupted period of schooling.

[23] Considering the continuum of family members, a dynasty faces no uncertainty. So its members can self-insure and behave as if risk-neutral with respect to investment decisions.

dynastic utility,

$$\int_{t_0}^{\infty} e^{-\rho(t-t_0)} N_t \frac{c_t^{1-\eta}-1}{1-\eta} dt \ ,$$

subject to the budget constraint

$$\int_{t_0}^{\infty} e^{-\int_{t_0}^{t} \iota_z dz} N_t c_t dt = p_{t_0} K_{t_0} + \int_{t_0}^{\infty} e^{-\int_{t_0}^{t} \iota_z dz} \int_{\Upsilon(t-\bar{u})}^{\Upsilon(t)} \lambda N_b e^{-\upsilon(t-b)} W_t(s_b, t-b-s_b) db dt.$$

On the right-hand side of the budget constraint, we have the value of the dynasty's capital at time t_0 plus, for all future periods t, the discounted (to time t_0) present value of wage income of all surviving dynasty members who remain employed at time t.

The solution to the dynasty's intertemporal maximization problem yields the Euler equation, $\dot{c}_t/c_t = (\iota_t - \rho)/\eta$, as usual. Moreover, by differentiating the budget constraint, we again obtain the no-arbitrage equation $\iota_t = R_t/p_t + g_p - \delta$, from which it follows that

$$\eta \frac{\dot{c}_t}{c_t} = \frac{R_t}{p_t} + g_p - \delta - \rho, \tag{11}$$

much as is true in the model with short lifespans.

Let us revisit the problem facing firms, before returning to the individuals' schooling choices. Much is the same as before. The main difference is that a firm may hire workers from different cohorts who vary in their schooling and experience. Firms must decide how much capital to allocate to each of their workers. However, with constant returns to scale and competitive firms that earn zero profits, it is as if each worker type indexed by s and u is hired by a separate firm, or by a separate unit of the firm. At time t, a firm that employs workers with schooling s and experience $u \leq \bar{u}$ maximizes profits by choosing the number L of such workers and the capital K with which to equip them so as to maximize $F(A_t K, B_t L, s, u) - R_t K - W_t(s,u) L$, where $W_t(s,u)$ is the competitive wage earned at time t by a worker with schooling s and experience u. The first-order conditions imply, as before,

$$r_t = f_k[\kappa(s,u,r_t), s, u],$$

and

$$\omega(s,u,r_t) = f[\kappa(s,u,r_t), s, u] - r_t \kappa(s,u,r_t)$$

for all workers $\{s,u\}$ that are present in the workforce at time t, where $f(k,s,u) \equiv F(A_tK/B_tL,1,s,u)$ is the intensive production function, $r_t = R_t/A_t$ is the rental rate per effective unit of capital, $\kappa(s,u,r_t)$ is the effective capital to effective labor ratio that the firm applies to workers of type $\{s,u\}$ when the rental rate per effective unit of capital is r_t, and $\omega(s,u,r_t)$ is the wage per effective unit of labor for workers of this type. In equilibrium, the wage schedule at time t satisfies $W_t(s,u)/B_t \equiv w_t(s,u) = \omega(s,u,r_t)$ and the sum total of the equipment allocated to all workers exhausts the available supply of capital, or

$$K_t = \frac{B_t}{A_t}\int_{\Upsilon(t-\bar{u})}^{\Upsilon(t)} \lambda N_b e^{-\upsilon(t-b)}\kappa(s_b, t-s_b-b, r_t)db.$$

Despite being two dimensional, the wage schedule for effective labor $w_t(s,u)$ changes over time only due to changes in r_t, which implicitly determines how much effective capital is allocated to a worker with schooling s and experience u.

To generate a BGP, we need a functional-form assumption and parameter restrictions that are analogous to Assumptions 1, 2, and 3 for the model with short lifespans. Now we adopt

Assumption 4 *The intensive production function can be written as* $f(k,s,u) = e^{\mu\beta s}h(ke^{-\mu s},u)$, *with* $\mu > 0$ *and* $\beta \in (0,1)$, *where*

(i) $h(z,u)$ *is strictly increasing, twice differentiable, and strictly concave in* $z \equiv ke^{-\mu s}$ *for all* $z > 0$ *and* $0 \leq u < \bar{u}$ *and* $h(z,u) = 0$ *for all* $u \geq \bar{u}$; *and*

(ii) $f(k,s,u)$ *is log supermodular in* k *and* s *for all* $u \in [0,\bar{u})$.

An example of a function that satisfies Assumption 4(i) and (ii) is $h(z,u) = \tilde{h}(u)(1+z^{-\alpha})^{-\beta/\alpha}$, where $\alpha > 0$ and $\tilde{h}(u)$ is first increasing and subsequently decreasing in u for $0 \leq u < \bar{u}$ and zero for $u \geq \bar{u}$. This generates an aggregate production function of the form $F(AK,BL,s,u) = \tilde{h}(u)(BL)^{1-\beta} \times \left[(AK)^{-\alpha} + (e^{\mu s}BL)^{-\alpha}\right]^{-\beta/\alpha}$.

Recall that Assumption 1 implies $\sigma_{KL} < 1$ in the model with short lifespans. By the same token (and by an analogous argument), Assumption 4 implies $\sigma_{KL} < 1$ when $u < \bar{u}$ in the OLG model. Now define $\mathcal{E}_{h,z}(z,u)$ to be the elasticity of $h(z,u)$ with respect to $z \equiv ke^{-\mu s}$. Assumption 4 implies that $\mathcal{E}_{h,z}(z,u)$ is strictly decreasing in z when $u < \bar{u}$, in analogy to what came before. Moreover, the elasticity $\mathcal{E}_{h,z}(z,u)$ equals the capital share in revenue at a firm (or unit) that employs workers with schooling s and experience u. To ensure that output is non-decreasing in schooling, we must have $\mathcal{E}_{h,z}(z,u) \leq \beta$.

Now define $d_{\max}(u) = \lim_{z\to 0} \mathcal{E}_{h,z}(z,u)$ and $d_{\min}(u) = \lim_{z\to\infty} \mathcal{E}_{h,z}(z,u)$. Let $d_{\max} = \inf_{0\le u<\bar{u}} d_{\max}(u)$ and $d_{\min} = \sup_{0\le u<\bar{u}} d_{\min}(u)$.[24] We impose the following parameter restrictions.

Assumption 5 *(i)* $\beta \ge \sup_{0\le u<\bar{u}} d_{\max}(u)$*; (ii)* $\frac{d_{\min}}{1-d_{\min}} < \Omega < \frac{d_{\max}}{1-d_{\max}}$ *(iii)* $(1-\beta)\,\mu > \gamma_K$*; and (iv)* $\frac{\mu\beta-\lambda}{(1-\beta)\mu} > \Omega$*, where* $\Omega \equiv \frac{1}{(1-\beta)\mu-\gamma_K}\left\{n-\rho+(1-\eta)\gamma_L+(\mu\beta-\lambda)\left[1-\frac{\eta\gamma_K}{(1-\beta)\mu}\right]\right\}$.

Part (i) guarantees that output is non-decreasing in schooling. Part (ii) provides for optimal schooling that is positive and finite. Part (iii) will be required for educational attainment to rise over time. Finally, part (iv) is analogous to Assumption 3 in the model with short lifespans inasmuch as it ensures that the integrals in the dynasty's budget constraint are finite.[25] Note that since $d_{\min} < d_{\max} \le \beta$, parts (ii) and (iv) of Assumption 5 together imply that $\mu\beta > \lambda$.

We return to the choice of schooling. Let us conjecture the existence of a BGP along which output, aggregate consumption, and the capital stock grow at constant rates. On a BGP, the goods-market clearing condition implies—as in Lemma 1—that aggregate consumption (as well as the value of the capital stock) must grow at the same rate as output. Then the dynasty's intertemporal optimization requires a constant real interest rate,

$$\iota = \eta\,(g_Y - n) + \rho$$

and the no-arbitrage condition $\iota_t = R_t/p_t + g_p - \delta$ then implies that r_t declines at constant rate $g_A - g_p = g_A + g_q = \gamma_K$. Using this observation, we show in the appendix that, along a BGP, choosing s_b to maximize the expected present discounted value of wages by the cohort born at b is equivalent to a maximization problem involving the choice of $x_b \equiv r_b e^{[(1-\beta)\mu-\gamma_K]s}$. Moreover, the latter maximization problem is independent of the birthdate b. We prove that the problem has a unique solution, x^*, provided (as we ultimately must assume) that the second-order condition is satisfied. It follows that s_b and r_b are tied together along any BGP by

$$x^* = r_b e^{[(1-\beta)\mu-\gamma_K]s_b} \quad \text{for all } b. \tag{12}$$

Now differentiate the relationship between r_b and s_b, and use the fact that, on a BGP, r_b falls at rate

[24] Whenever $h(z,u)$ is log separable in z and u, $\mathcal{E}_{h,z}(\cdot)$ is independent of u. Then $d_{\min}(u)$ and $d_{\max}(u)$ are constants. Moreover, when $h\,(z,u) = h_u\,(u)\left(1+z^{-\alpha}\right)^{-\beta/\alpha}$, $d_{\min} = 0$ and $d_{\max} = \beta$.

[25] Part (iv) can alternatively be written as

$$\rho > n + (1-\eta)\left[\gamma_L + \frac{\mu\beta-\lambda}{(1-\beta)\,\mu}\gamma_K\right],$$

which is closer in form to what appears in Assumption 3.

γ_K. Then schooling by birth cohort must evolve according to

$$\dot{s}_b = \frac{\gamma_K}{(1-\beta)\,\mu - \gamma_K}\ ; \tag{13}$$

that is, educational attainment rises *linearly* over time.[26] This prediction of the model seems roughly in accord with the U.S. experiences (as depicted in Figure 2) for the birth cohorts from 1876 until approximately 1955, and then again for the later cohorts, albeit with schooling then growing more slowly than before.

We take a momentary detour to comment on the role played by retirement in our model. Recall that productivity falls to zero after experience reaches $\bar{u}$, at which point a surviving individual leaves the workforce. We will see shortly that $\bar{u}$ has no effect on the steady-state growth rate. We introduced the assumption that productivity falls to zero in order to counteract an implication of the (common but clearly unrealistic) assumption that death occurs with a constant hazard rate. Given the evolution of educational attainment dictated by (13), there must have been some birth cohorts in the distant past for whom the non-negativity constraint that $s \geq 0$ was binding. With a constant probability of death, some members of these ancient cohorts must still be alive at time t. Indeed, without retirement, there would be a mass of workers at every moment with schooling $s = 0$. The presence of such individuals in the labor force would complicate aggregation in the model. It seems best to assume that individuals must eventually leave the workforce given the assumption (made for convenience) that individuals might live unreasonably long lives.

Our next task is to calculate aggregate output, Y_t. Define the function $\Phi(z,u)$ as the inverse of $h_z(z,u)$, so that $z \equiv \Phi[h_z(z,u),u]$. Then $\kappa(s,u,r) = e^{\mu s}\Phi\left[re^{(1-\beta)\mu s},u\right]$. At time t, a worker with schooling s and experience u uses $B_t\kappa(s,u,r_t) = B_t e^{\mu s}\Phi\left[r_t e^{(1-\beta)\mu s},u\right]$ units of effective capital and produces $B_t e^{\mu\beta s}h\left\{\Phi\left[r_t e^{(1-\beta)\mu s},u\right],u\right\}$ units of output. Only individuals born at times b between $\Upsilon(t-\bar{u})$ and $\Upsilon(t)$ are employed at time t. Since r_t declines at rate γ_K and schooling evolves according to (13), it follows that at time t an individual born at $b \in [\Upsilon(t-\bar{u}),\Upsilon(t)]$ with experience $u = t-b-s_b$, produces a flow $B_t e^{\mu\beta s_b}h\left\{\Phi\left[e^{-\gamma_K u}x^*,u\right],u\right\}$ of output.

An individual with experience u at time t was born at $\Upsilon(t-u)$ and has $t-u-\Upsilon(t-u)$ years of schooling. Therefore, using (13) to relate the schooling of individuals born at $\Upsilon(t-u)$ to the schooling

[26] Assumption 5(iii) ensures that $\dot{s}_b > 0$. To see the parallel with the short lifespan model, (12) could be written instead as $x^* = r_t e^{(1-\beta)\mu s_{\Upsilon(t)}}$, thereby relating a cohort's schooling to the cost of effective capital upon entry into the workforce. Similarly, we could rewrite (13) as $d\left[s_{\Upsilon(t)}\right]/dt = \gamma_K/(1-\beta)\,\mu$, measuring the rate of increase in the schooling among those just entering the workforce.

of those born at $\Upsilon(t_0)$, we have

$$s_{\Upsilon(t-u)} = t - u - \Upsilon(t-u) = t_0 - \Upsilon(t_0) + \frac{\gamma_K}{(1-\beta)\mu}(t - t_0 - u) \ . \tag{14}$$

Since the size at time t of the cohort born at b is $\lambda N_b e^{-\upsilon(t-b)} = \lambda e^{(\lambda-\upsilon)(b-t_0)} N_{t_0} e^{-\upsilon(t-b)}$, the number of workers with experience u at time t is $L_t(u) = \lambda N_{t_0} e^{-\lambda t_0} e^{-\upsilon(t-t_0)} e^{\lambda\Upsilon(t-u)} = \lambda N_t e^{\lambda[\Upsilon(t-u)-t]}$ and using (14) gives

$$L_t(u) = \lambda N_t e^{\lambda[\Upsilon(t_0)-t_0]} e^{-\lambda\frac{\gamma_K}{(1-\beta)\mu}(t-t_0)} e^{-\lambda\left[1-\frac{\gamma_K}{(1-\beta)\mu}\right]u} \ . \tag{15}$$

Combining these observations, aggregate output at time t is given by

$$Y_t = B_t \int_0^{\bar{u}} L_t(u)\, e^{\mu\beta s_{\Upsilon(t-u)}} h\left\{\Phi\left[e^{-\gamma_K u} x^*, u\right], u\right\} du \ .$$

Since $x^* = r_{t-u} e^{(1-\beta)\mu s_{\Upsilon(t-u)}} = r_t e^{\gamma_K u} e^{(1-\beta)\mu s_{\Upsilon(t-u)}}$, aggregate output at t is

$$Y_t = B_t \int_0^{\bar{u}} L_t(u)\, r_t^{-\frac{\beta}{1-\beta}} \left[e^{-\gamma_K u} x^*\right]^{\frac{\beta}{1-\beta}} h\left\{\Phi\left[e^{-\gamma_K u} x^*, u\right], u\right\} du \ . \tag{16}$$

Using (16), we can readily calculate the growth rate of output on a BGP, which is[27]

$$g_Y = n + \gamma_L + \frac{\mu\beta - \lambda}{(1-\beta)\mu}\gamma_K \ . \tag{17}$$

Note the similarity between (17) and g_Y in part (ii) of Proposition 2, except that the progeneration rate λ enters the former but does not exist as a separate parameter in the model with short lifespans. Note too that Assumption 5(ii) and (iv) imply $\mu\beta > \lambda$, as we have observed previously, so the growth rate again is increasing in both the rate of labor-augmenting technological progress and the total rate of capital-augmenting technological progress.[28]

How do factor shares evolve along the BGP that we have just described? Recall that an individual

[27] In performing this calculation, we use $L_t(u) = e^{\left[n-\frac{\gamma_K}{(1-\beta)\mu}\lambda\right](t-t_0)} L_{t_0}(u)$, $B_t = e^{\gamma_L(t-t_0)} B_{t_0}$, and $r_t = e^{-\gamma_K(t-t_0)} r_{t_0}$.

[28] The transversality condition on the BGP requires $\iota > g_Y$, which in turn requires

$$\frac{\mu\beta - \lambda}{(1-\beta)\mu} > \Omega \ ,$$

where we recall that

$$\Omega \equiv \frac{1}{(1-\beta)\mu - \gamma_K}\left\{n - \rho + (1-\eta)\gamma_L + (\mu\beta - \lambda)\left[1 - \frac{\eta\gamma_K}{(1-\beta)\mu}\right]\right\} \ .$$

This condition has been assumed to hold in part (iv) of Assumption 5.

with schooling s and experience u works at time t with $B_t e^{\mu s}\Phi\left[r_t e^{(1-\beta)\mu s}, u\right]$ effective units of capital. This equals $B_t e^{\mu s \Upsilon(t-u)}\Phi\left[e^{-\gamma_K u}x^*, u\right]$ for individuals who are still in the workforce at time t, given that schooling evolves according to (13) and r_t declines at rate γ_K. Then, since the schooling level of a worker with experience u at time t is given by (14) and there are $L_t(u)$ such individuals in the labor force, it follows from the capital-market clearing condition that

$$K_t = \frac{B_t}{A_t}\int_0^{\bar{u}} L_t\left(u\right) e^{\mu s \Upsilon(t-u)}\Phi\left[e^{-\gamma_K u}x^*, u\right] du \ .$$

Now, using $r_t = \frac{R_t}{A_t}$ and $x^* = r_t e^{\gamma_K u} e^{(1-\beta)\mu s \Upsilon(t-u)}$, capital income amounts to

$$R_t K_t = r_t B_t \int_0^{\bar{u}} L_t\left(u\right) r_t^{-\frac{1}{1-\beta}}\left[e^{-\gamma_K u}x^*\right]^{\frac{1}{1-\beta}}\Phi\left[e^{-\gamma_K u}x^*, u\right] du \ . \tag{18}$$

Thus, on the balanced growth path, aggregate capital income grows at the same rate g_Y as aggregate output, which implies that the capital (and labor) share is constant. Combining (15), (16), and (18) yields

$$\theta_K = \frac{\int_0^{\bar{u}} e^{-\left[\lambda+\frac{\mu\beta-\lambda}{(1-\beta)\mu}\gamma_K\right]u} e^{-\gamma_K u}x^*\Phi\left[e^{-\gamma_K u}x^*, u\right] du}{\int_0^{\bar{u}} e^{-\left[\lambda+\frac{\mu\beta-\lambda}{(1-\beta)\mu}\gamma_K\right]u} h\left\{\Phi\left[e^{-\gamma_K u}x^*, u\right], u\right\} du}. \tag{19}$$

We are ready to summarize our main findings for the model with overlapping generations. We have

Proposition 3 *Suppose that Assumptions 4 and 5 hold in the model with overlapping generations. Then the OLG economy has a unique balanced growth path. On the BGP,*

(i) aggregate output, aggregate consumption, and aggregate wages grow at rate

$$g_Y = n + \gamma_L + \frac{\mu\beta-\lambda}{(1-\beta)\mu}\gamma_K \ ;$$

(ii) the educational attainment of new cohorts rises according to

$$\dot{s}_b = \frac{\gamma_K}{(1-\beta)\mu - \gamma_K} \ ;$$

(iii) the aggregate capital share is constant.

Before leaving this section, we offer several further observations about the BGP in the OLG model. First, since at time t there are $L_t\left(u\right)$ workers with experience u, (15) implies that the time t labor force is

$$L_t = \lambda N_t e^{\lambda[\Upsilon(t_0)-t_0]} e^{-\lambda\frac{\gamma_K}{(1-\beta)\mu}(t-t_0)} \int_0^{\bar{u}} e^{-\lambda\left[1-\frac{\gamma_K}{(1-\beta)\mu}\right]u} du \ ,$$

so that the labor force participation rate is

$$\frac{L_t}{N_t} = \lambda e^{\lambda[\Upsilon(t_0)-t_0]} e^{-\lambda\frac{\gamma_K}{(1-\beta)\mu}(t-t_0)} \int_0^{\bar{u}} e^{-\lambda\left[1-\frac{\gamma_K}{(1-\beta)\mu}\right]u} du \ .$$

It follows that, on the BGP, the participation rate declines at a constant rate $\lambda\gamma_K/(1-\beta)\mu$. That is, as time devoted to school rises over time, a smaller fraction of the population works. The FRED data (Series LNS11300001) show that labor force participation among men has been declining steadily in the United States since the start of the series in 1948.

Second, the fraction of workers with less than u years of experience at time t is

$$\frac{\int_0^u L_t(z)\,dz}{L_t} = \frac{\int_0^u e^{-\lambda\left[1-\frac{\gamma_K}{(1-\beta)\mu}\right]z} dz}{\int_0^{\bar{u}} e^{-\lambda\left[1-\frac{\gamma_K}{(1-\beta)\mu}\right]z} dz} \ .$$

This fraction is constant over time. In other words, the distribution of experience among those in the labor force does not vary along the BGP. There are, however, shifts in the distribution of schooling in the labor force. At time t the fraction of workers with experience below u also equals the fraction of workers with at least $t_0 - \Upsilon(t_0) + \frac{\gamma_K}{(1-\beta)\mu}(t-t_0-u)$ years of schooling. Therefore, the schooling of workers at all levels of experience increases by $\gamma_K/(1-\beta)\mu$ per year. Consequently, the entire density of schooling shifts to the right at this constant rate.

We can also calculate the returns to schooling and the returns to experience along the BGP. We find[29]

$$\frac{\partial \log W_t(s,u)}{\partial s} = \mu\beta - \mu(1-\beta)\frac{\mathcal{E}_{h,z}\left\{\Phi\left[r_t e^{(1-\beta)\mu s}, u\right], u\right\}}{1-\mathcal{E}_{h,z}\left\{\Phi\left[r_t e^{(1-\beta)\mu s}, u\right], u\right\}}$$

and

$$\frac{\partial \log W_t(s,u)}{\partial u} = \frac{h_u\left(\Phi\left[r_t e^{(1-\beta)\mu s}, u\right], u\right)}{h\left(\Phi\left[r_t e^{(1-\beta)\mu s}, u\right], u\right)} \frac{1}{1-\mathcal{E}_{h,z}\left\{\Phi\left[r_t e^{(1-\beta)\mu s}, u\right], u\right\}} \ .$$

In the cross-section, log wages increase with educational attainment (holding experience constant), albeit at a declining rate. This is reminiscent of a Mincer wage equation (Mincer, 1974), except that Mincer posited a linear relationship between log wages and years of schooling. Over time, the returns to schooling

[29] We use $\omega(s,u,r) = e^{\mu\beta s}\left\{h\left(\Phi\left[re^{(1-\beta)\mu s}, u\right], u\right) - re^{(1-\beta)\mu s}\Phi\left[re^{(1-\beta)\mu s}, u\right]\right\}$.

rise as r_t declines, for workers with a given s and u. Finally, if $h(z,u)$ is log separable in z and u, then the returns to experience holding u constant are increasing in s, while the returns to experience conditional on s and u fall over time.

Finally, we turn to the determinants of the long-run capital share. Unfortunately, the expression in (19) does not provide a simple and transparent relationship between θ_K and the rates of technical progress, in large part because the economy is populated by individuals with different levels of schooling and varied experience who therefore work with different amounts of capital. To illustrate how changes in technological progress impact the capital share and the wage profile, we resort to numerical simulation of a parameterized version of the model.

For the simulation exercise, we use the production function

$$F(A_tK, B_tL, s, u) = \tilde{h}(u)\,(B_tL)^{1-\beta}\left[A_tK^{-\alpha} + (e^{\mu s}B_tL)^{-\alpha}\right]^{-\beta/\alpha} \text{ for } u < \bar{u},$$

which, as we discussed above, corresponds to $h(z,u) = \tilde{h}(u)\left[1 + z^{-\alpha}\right]^{-\beta/\alpha}$. We adopt a simple, quadratic experience profile, $\tilde{h}(u) = 1 + 0.2\left[1 - (2u/\bar{u} - 1)^2\right]$ and specify a working life of $\bar{u} = 40$ years. We set the birth and death rates equal to $\lambda = \upsilon = 0.01$. For each calibration, we choose the production function parameters α, β, and μ, so that in the baseline case ($\gamma_K = 0.02$ and $\gamma_L = 0.01$) the capital share is 0.35, the average local elasticity of substitution between capital and labor is 0.6, and educational attainment increases by one year each decade.

We will find that the sensitivity of the capital share to changes in technological progress is governed by the real interest rate. For this reason, we must choose the intertemporal elasticity of substitution and the discount rate with care. What interest rate should we target? On the one hand, the low riskless rate of return in the U.S. economy over many decades suggest that we ought to choose parameters to match a low value of ι. On the other hand, our model features equality on the margin between the internal rate of return on schooling and the discount rate. But rates of return on schooling have been high in the United States and elsewhere, which suggests choosing parameters that yield a higher value for ι. It is impossible to choose parameters that simultaneously match the low riskless rate and the high rate of return on schooling.[30] Instead of taking a strong stand on the appropriate interest rate for our model, we present comparative statics under both low-interest-rate and high-interest-rate scenarios.

Table 1 shows two sets of simulation results. In the top part of the table, which presents a scenario

[30] The gap may be explained by factors outside the model such as financing constraints, risk compensation, or a utility cost of schooling.

Low Interest Rate: $\rho = .01,\ \eta = 1$					
γ_K	γ_L	Growth in per capita Income	Annual Increase in Schooling	Capital Share	Interest Rate
0.03	0.01	0.028	0.158	0.348	0.038
0.02	0.01	0.022	0.1	0.35	0.032
0.01	0.01	0.016	0.048	0.352	0.026
0.02	0.02	0.032	0.1	0.35	0.042
0.02	0.01	0.022	0.1	0.35	0.032
0.02	0	0.012	0.1	0.35	0.022

High Interest Rate: $\rho = .01,\ \eta = 3$					
γ_K	γ_L	Growth in per capita Income	Annual Increase in Schooling	Capital Share	Interest Rate
0.03	0.01	0.038	0.158	0.288	0.123
0.02	0.01	0.028	0.1	0.35	0.095
0.01	0.01	0.019	0.048	0.402	0.068
0.02	0.02	0.038	0.1	0.303	0.125
0.02	0.01	0.028	0.1	0.35	0.095
0.02	0	0.018	0.1	0.394	0.065

Table 1: Response of Long-Run Growth Rate, Schooling, Capital Share, and Interest Rate to Changes in Rates of Technological Progress

with a baseline interest rate of 3.2% per year, a decrease in the rate of capital-augmenting or investment-specific technical change of one percentage point per year reduces the output growth rate by a little more than half a percentage point per year and reduces the rate of increase in educational attainment by a half-year of schooling per decade. However, in this case, the capital share moves hardly at all. In the bottom part of the table, which presents a scenario with a higher baseline rate of 9.5 percent per year, a decline in γ_K has a similar impact on output growth and on the rate of increase in educational attainment, but the impact on the capital share is much more substantial.

What accounts for this difference? The impact of a change in γ_K on factor shares reflects the responsiveness of schooling decisions to changes in the rate of technological progress. When the interest rate is low, an individual's choice of schooling reflects the allocation of effective capital she anticipates throughout her lifetime. As γ_K falls, so does the allocation of effective capital later in life, and this reduces the optimal time in school. The relatively elastic response of schooling cushions the impact of capital deepening on the factor shares, much as in the model with short lifespans, for which we found that θ_K is independent of γ_K and γ_L. Indeed, we show in the appendix that, in the OLG model, as the interest rate approaches the growth rate of the economy from above, the capital share approaches $(\mu\beta - \lambda) / (\mu - \lambda)$,

which also is independent of γ_K and γ_L. In contrast, if individuals discount future wages heavily, then their educational decisions will largely reflect their capital allocation and the state of technology shortly after their time in school. The response of s to a change in γ_K or γ_L will be muted by the relative disregard for capital and technology later in life. With a dampened response of schooling, the impact on the capital share is greater. Interestingly, a slowdown in the rate of labor-augmenting technical change has similar effects on output growth and the capital share as a fall in the rate of capital-augmenting technological progress in both cases.

5 Conclusion

Over at least the last half century, the United States has experienced balanced growth; nearly constant growth rates of output per worker, capital, and consumption, and roughly constant factor shares until quite recently. Uzawa's Growth Theorem established that, in a conventional neoclassical growth model, balanced growth can be realized only if technical change is purely labor-augmenting or the elasticity of substitution between capital and labor is unity. But the price of capital equipment has been falling precipitously over time and the elasticity of substitution appears to be significantly less than one. We have shown that if labor quantity and labor quality do not enter the aggregate production function symmetrically, capital-augmenting technological change can be reconciled with balanced growth, provided that schooling increases over time and that capital and schooling are sufficiently complementary. Our model matches trends for the U.S. economy that suggest balance as well as others that may appear to reflect unbalanced growth, such as a linear increase in educational attainment and a falling labor-force participation rate. We achieve this while also matching conventional estimates of the capital-labor elasticity of substitution and life-cycle earnings profiles.

The basic mechanism in our model is straightforward: over time, growing stocks of effective capital raise the returns to schooling, which induces individuals to spend more time in school. Inasmuch as capital and labor are complements, capital accumulation tends to lower capital's share in national income, but this is offset by the subsequent rise in schooling, because capital and schooling are also complements. When capital and schooling are more complementary than capital and labor, the second effect can neutralize the first. Although the presence of these offsetting forces is natural enough, restrictions on how schooling enters the production function are needed to maintain exact balance along an equilibrium trajectory. The restrictions are in a sense analogous to those usually imposed on preferences in a dynamic model

in order to generate balanced growth. Specifically, while it may be natural to assume that income and substitution effects offset one another as wages rise, the intratemporal utility function must be specified in a particular away so as to maintain perfect balance along an equilibrium trajectory. Just as balanced-growth preferences are consistent with a range of intertemporal elasticities of substitution and labor-supply elasticities, so too are the restrictions we impose on the production function consistent with a range of elasticities of substitution between capital and labor and between capital and schooling.

References

[1] Acemoğlu, Daron, 2003. "Labor and Capital-Augmenting Technical Change," *Journal of the European Economic Association 1*(1), 1-37.

[2] Acemoğlu, Daron, 2009. *Introduction to Modern Economic Growth,* Princeton, NJ: Princeton University Press.

[3] David Autor, Katz, Lawrence F. and Krueger, Alan B., 1998. "Computing Inequality: Have Computers Changed the Labor Market?" *Quarterly Journal of Economics 113 (4)*, 1169-1214.

[4] Beckmann, Martin J, 1977. "Management Production Functions and the Theory of the Firm," *Journal of Economic Theory 14(1)*, 1-18.

[5] Blanchard, Olivier J. and Fischer, Stanley, 1989. *Lectures on Macroeconomics*, Cambridge, MA: The MIT Press.

[6] Blinder, Alan S. and Weiss, Yoram, 1975. "Human Capital and Labor Supply: A Synthesis," *Journal of Political Economy 83(3)*, 449-72.

[7] Chirinko, Robert S., 2008. "σ: The Long and the Short of It," *Journal of Macroeconomics 30(2)*, 671-86.

[8] Chirinko. Robert S. and Mallick, Debdulal, 2014. "The Substitution Elasticity, Factor Shares, Long-Run Growth, and the Low-Frequency Panel Model," CESifo Working Paper Series No. 4895, CESifo Group.

[9] Chirinko, Robert S., Fazzari, Steven M., and Meyer, Andrew P., 2011. "A New Approach to Estimating Production Function Parameters: The Elusive Capital-Labor Substitution Elasticity," *Journal of Business & Economic Statistics 29 (4)*, 587-594.

[10] Cummins, Jason and Violante, Giovanni L., 2002. "Investment-Specific Technical Change in the US (1947-2000): Measurement and Macroeconomic Consequences," *Review of Economic Dynamics 5(2)*, 243-84.

[11] DiCecio, Riccardo, 2009. "Sticky Wages and Sectoral Labor Co-Movement," *Journal of Economic Dynamics and Control 33(3)*, 538-53.

[12] Elsby, Michael W. L., Hobijn, Bart, and Sahin, Aysegul, 2013. "The Decline of the U.S. Labor Share," *Brookings Papers on Economic Activity 47(2)*, 1-63.

[13] Goldin, Claudia and Katz, Lawrence F., 2007. "Long-Run Changes in the Wage Structure: Narrowing, Widening, and Polarization," *Brookings Papers on Economic Activity 38(2)*, 135-68.

[14] Gordon, Robert J., 1990. *The Measurement of Durable Goods Prices*, Chicago, IL: University of Chicago Press.

[15] Greenwood, Jeremy, Hercovitz, Zvi and Krusell, Per, 1997. "Long-Run Implications of Investment-Specific Technological Change," *American Economic Review 87(3)*, 342-62.

[16] He, Hui and Liu, Zheng, 2008. "Investment-Specific Technological Change, Skill Accumulation, and Wage Inequality," *Review of Economic Dynamics, 11(2)*, 314-34.

[17] Herrendorf, Berthold, Herrington, Christopher and Valentinyi, Akos, 2013. "Sectoral Technology and Structural Transformation," CEPR Discussion Papers No. 9386, London UK.

[18] International Monetary Fund, 2014. *Economic Outlook*, IMF: Washington DC.

[19] Johansen, Leif, 1959. "Substitution versus Fixed Production Coefficients in the Theory of Economic Growth: A Synthesis," *Econometrica 27(2)*, 157-176.

[20] Jones, Charles I., 2005. "The Shape of Production Functions and the Direction of Technical Change," *Quarterly Journal of Economics 120(2)*, 517-549.

[21] Jones, Charles I., 2015. "The Facts of Economic Growth," in preparation for the *Handbook of Macroeconomics*.

[22] Jones, Charles I. and Romer, Paul M., 2010. "The New Kaldor Facts: Ideas, Institutions, Population, and Human Capital," *American Economic Journal: Macroeconomics 2(1)*, 224-45.

[23] Jones, Charles I. and Scrimgeour, Dean, 2008. "A New Proof of Uzawa's Steady-State Growth Theorem," *Review of Economics and Statistics 90(1)*, 180-2.

[24] Kaldor, Nicholas, 1961. "Capital Accumulation and Economic Growth" in F.A. Lutz and D.C. Hague, eds., *The Theory of Capital*, New York: St. Martins Press.

[25] Karabarbounis, Loukas and Neiman, Brent, 2014. "The Global Decline of the Labor Share," *Quarterly Journal of Economics 129(1)*, 61-103.

[26] Krusell, Per, Ohanian, Lee E., Rios-Rull, José-Victor, and Violante, Giovanni L., 2000. "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica 68(5)*, 1029-63.

[27] Lawrence, Robert Z., "Recent Declines in Labor's Share in US Income: A Preliminary Neoclassical Account," NBER Working Paper No. 21296, National Bureau of Economic Research.

[28] Lucas, Robert E., Jr., 1988. "On the Mechanics of Economic Development," *Journal of Monetary Economics 22(1)*, 3-47.

[29] Mincer, Jacob A., 1974. *Schooling, Experience, and Earnings*, New York: The National Bureau of Economic Research.

[30] Maliar, Lilia and Maliar, Serguei, 2011. "Capital-Skill Complementarity and Balanced Growth," *Economica 78(310)*, 240-59.

[31] Oberfield, Ezra and Raval, Devesh, 2014. "Micro Data and Macro Technology," NBER Working Paper No. 20452, National Bureau of Economic Research.

[32] Picketty, Thomas, 2014. *Capital in the Twenty-First Century*, Cambridge, MA: Harvard University Press.

[33] Rosen, Sherwin, 1983. "Authority, Control, and the Distribution of Earnings," *The Bell Journal of Economics 13(2)*, 311-23.

[34] Schlicht, Ekkehart, 2006. "A Variant of Uzawa's Theorem," *Economics Bulletin 5(6)*, 1–5.

[35] Sheshinski, Eytan, 1967. "Balanced Growth and Stability in the Johansen Vintage Model," *Review of Economic Studies 34(2)*, 239-48.

[36] Solow, Robert M., 1960. "Investment and Technical Progress," in K.J. Arrow, S. Karlin and P. Suppes, eds., *Mathematical Methods in the Social Sciences*, 1959, Stanford, CA: Stanford University Press.

[37] Uzawa, Hirofumi, 1961. "Neutral Inventions and the Stability of Growth Equilibrium," *Review of Economic Studies 28(2)*, 117-24.

[38] Uzawa, Hirofumi, 1965. "Optimal Technical Change in an Aggregate Model of Economic Growth," *International Economic Review 6(1)*, 18-31.

[39] Yaari, Menahem E., 1965. "Uncertain Lifetime, Life Insurance, and the Theory of the Consumer," *Review of Economic Studies 32(2)*, 137–50.

Appendix for "Balanced Growth Despite Uzawa"

by

Gene M. Grossman, Elhanan Helpman, Ezra Oberfield and Thomas Sampson

Proofs from Section 2

Proof of Lemma 1

By assumption $C_t < Y_t$. Therefore, the resource constraint $Y_t = C_t + I_t/q_t$ ensures $I_t > 0$. The capital accumulation equation is $\dot{K}_t = I_t - \delta K_t$ implying

$$g_K = \frac{\dot{K}_t}{K_t} = \frac{I_t}{K_t} - \delta.$$

On a BGP g_K is constant meaning that since $I_t > 0$ the growth rates of I and K must be the same. Thus, $g_I = g_K$.

Differentiating the resource constraint and rearranging gives

$$(g_C - g_Y)\frac{C_t}{Y_t} + (g_I - g_q - g_Y)\frac{I_t/q_t}{Y_t} = 0.$$

Substituting for $\frac{I_t/q_t}{Y_t} = 1 - \frac{C_t}{Y_t}$ in this expression and using $g_I = g_K$ we have

$$(g_K - g_q - g_C)\frac{C_t}{Y_t} = g_K - g_q - g_Y.$$

If both sides of this expression equals zero we immediately obtain $g_Y = g_C = g_K - g_q$ as claimed in the lemma. Otherwise, since the growth rates are constant on a BGP it must be that C and Y grow at the same rate implying $g_Y = g_C$. But then the resource constraint implies $\frac{I_t/q_t}{Y_t} = 1 - \frac{C_t}{Y_t}$ is constant and, since $g_I = g_K$, this ensures $g_Y = g_K - g_q$. Therefore, the lemma holds.

Proof of Proposition 1

Since factors are paid their marginal products the capital share is $\theta_K = A_t K_t F_K\left(A_t K_t, B_t L_t, s_t\right)/Y_t$. Note also that because F has constant returns to scale in its first two arguments $F_K\left(A_t K_t, B_t L_t, s_t\right) = F_K(k_t, 1, s_t)$ where

$k_t = A_t K_t / B_t L_t$. Therefore, on a BGP where the capital share is positive and constant we have[31]

$$0 = \frac{\dot{\theta}_K}{\theta_K} = g_A + g_K - g_Y + \frac{d \log F_K\left(k_t, 1, s_t\right)}{dt} = \gamma_K + \frac{d \log F_K\left(k_t, 1, s_t\right)}{dt},$$

where the final equality uses Lemma 1 and $\gamma_K = g_A + g_q$.

Taking the derivative of F_K and using $kF_{KK} + F_{KL} = 0$ we have

$$\gamma_K = -\frac{F_{KK}\dot{k}_t + F_{Ks}\dot{s}_t}{F_K} = \frac{F_{LK}}{F_K}\frac{\dot{k}_t}{k_t} - \frac{F_{Ks}\dot{s}_t}{F_K} = \frac{1}{\sigma_{KL}}\frac{F_L}{F}\frac{\dot{k}_t}{k_t} - \frac{F_{Ks}\dot{s}_t}{F_K}.$$

Since $1 - \theta_K = F_L/F$ this can be rearranged to give

$$\sigma_{KL}\gamma_K = \left(1 - \theta_K\right)\frac{\dot{k}_t}{k_t} - \sigma_{KL}\frac{F_{Ks}\dot{s}_t}{F_K}. \tag{20}$$

To simplify (20) it will be useful to derive an expression for F_{Ks}/F_K. Note that

$$\frac{\partial\left[F_s/F_L\right]}{\partial K} = \frac{F_{Ks}}{F_L} - \frac{F_{LK}F_s}{F_L^2} = \frac{F_K}{F_L}\left(\frac{F_{Ks}}{F_K} - \frac{1}{\sigma_{KL}}\frac{F_s}{F}\right). \tag{21}$$

Rearranging, we have $\frac{F_{Ks}}{F_K} = \frac{F_L}{F_K}\frac{\partial[F_s/F_L]}{\partial K} + \frac{1}{\sigma_{KL}}\frac{F_s}{F}$. Plugging this into (20) gives

$$\sigma_{KL}\gamma_K = \left(1 - \theta_K\right)\frac{\dot{k}_t}{k_t} - \sigma_{KL}\frac{F_L}{F_K}\frac{\partial\left[F_s/F_L\right]}{\partial K}\dot{s}_t - \frac{F_s\dot{s}_t}{F}. \tag{22}$$

Finally, differentiating the production function $Y_t = F(A_t K_t, B_t L_t, s_t)$ yields

$$\begin{aligned} g_Y &= \theta_K\left(g_A + g_K\right) + \left(1 - \theta_K\right)\left(g_B + g_L\right) + \frac{F_s\dot{s}_t}{F}, \\ &= g_A + g_K - \left(1 - \theta_K\right)\frac{\dot{k}_t}{k_t} + \frac{F_s\dot{s}_t}{F}. \end{aligned}$$

Using Lemma 1 and $\gamma_K = g_A + g_q$ this implies

[31] Instead of assuming constant factor shares, this expression can also be obtained by assuming the rental price of capital R_t declines at rate g_q. To see this differentiate $R_t = A_t F_K\left(k_t, 1, s_t\right)$.

$$\gamma_K = (1-\theta_K)\frac{\dot{k}_t}{k_t} - \frac{F_s \dot{s}_t}{F}.$$

Substituting this expression into (22) gives equation (1). This completes the proof.

Generalization of Proposition 1

Proposition 1 assumes technical change is factor augmenting, but we can generalize the proposition by relaxing this restriction. Suppose the production function is $Y = \hat{F}(K, L, s; t)$ where technical change is captured by the dependence of $\hat{F}$ on t. We can decompose technical change into a Harrod-neutral component and a non-Harrod-neutral residual. Technical change is Harrod-neutral if, holding the capital-output ratio and schooling fixed, it does not affect the marginal product of capital (Uzawa 1961). Therefore, we can define the non-Harrod-neutral component of technical change as the change in the marginal product of capital for a given capital-output ratio and schooling.

Let φ be the capital-output ratio and define $\hat{\kappa}(\varphi, s; t)$ by

$$\varphi = \frac{\hat{\kappa}(\varphi, s; t)}{\hat{F}(\hat{\kappa}(\varphi, s; t), 1, s; t)}.$$

$\hat{\kappa}(\varphi, s; t)$ is the capital-labor ratio that ensures the capital-output ratio equals φ given s and t. Differentiating this expression with respect to t while holding s and φ constant and using $\theta_K = \hat{\kappa}\hat{F}_K/\hat{F}$ implies

$$\frac{\hat{\kappa}_t}{\hat{\kappa}} = \frac{1}{1-\theta_K}\frac{\hat{F}_t}{\hat{F}}. \tag{23}$$

When technical change is Harrod-neutral $\hat{\kappa}_t \frac{\partial}{\partial \hat{\kappa}} \log \hat{F}_K + \frac{\partial}{\partial t} \log \hat{F}_K = 0$. Thus, we define the non-Harrod-neutral component of technical change Ψ by

$$\Psi \equiv -\sigma_{KL}\left[\hat{\kappa}_t \frac{\partial}{\partial \hat{\kappa}} \log \hat{F}_K(\hat{\kappa}(\varphi, s; t), 1, s; t) + \frac{\partial}{\partial t} \log \hat{F}_K(\hat{\kappa}(\varphi, s; t), 1, s; t)\right].$$

From this definition we have

$$\begin{aligned}\Psi &= -\sigma_{KL}\left(\frac{\hat{F}_{KK}\hat{\kappa}_t}{\hat{F}_K}+\frac{\hat{F}_{Kt}}{\hat{F}_K}\right),\\ &= -\sigma_{KL}\left(\frac{\hat{F}_{KK}}{\hat{F}_K}\frac{\hat{\kappa}}{1-\theta_K}\frac{\hat{F}_t}{\hat{F}}+\frac{\hat{F}_{Kt}}{\hat{F}_K}\right),\\ &= \frac{\hat{F}_t}{\hat{F}}-\sigma_{KL}\frac{\hat{F}_{Kt}}{\hat{F}_K},\end{aligned} \tag{24}$$

where the second line follows from (23) and the third line uses $\hat{\kappa}\hat{F}_{KK}=-\hat{F}_{KL}$, the definition of σ_{KL} and $1-\theta_K=\hat{F}_L/\hat{F}$. Note that in the case where technical change is factor augmenting we have $\hat{F}(K,L,s;t)=F(A_tK,B_tL,s)$ which implies $\Psi=(1-\sigma_{KL})g_A$.

Using the expression for Ψ given in (24) we obtain the following generalization of Proposition 1.

Proposition 4 *Suppose the production function is $Y=\hat{F}(K,L,s;t)$ and that investment-specific technological progress occurs at constant rate g_q. If there exists a BGP along which the income shares of capital and labor are constant and strictly positive when factors are paid their marginal products, then*

$$(1-\sigma_{KL})\,g_q+\Psi=\sigma_{KL}\frac{\hat{F}_L}{\hat{F}_K}\frac{\partial\left[\hat{F}_s/\hat{F}_L\right]}{\partial K}\dot{s}.$$

To avoid repetition, we omit the proof of Proposition 4 since it follows the same series of steps used to prove Proposition 1. Suppose either s is constant as in Corollary 1 or the production function can be written in terms of a measure of human capital $H(L,s,t)$ implying $\frac{\partial\left[\hat{F}_s/\hat{F}_L\right]}{\partial K}=0$ as in Corollary 2. Then Proposition 4 implies that a BGP with constant and strictly positive factor shares can exist only if $\Psi=0$ and either $\sigma_{KL}=1$ or $g_q=0$. Thus, a BGP is possible only if technical change that affects the production function is Harrod-neutral and either the elasticity of substitution between capital and labor equals one or there is no investment-specific technological change.

Proofs from Section 3

Implications of Assumption 1

$f(k,s)$ is strictly log supermodular if and only if $f_{ks}f>f_kf_s$. Using Assumption 1 to compute these derivatives it follows that f is strictly log supermodular if and only if $\mathcal{E}_h(z)-\mathcal{E}_{h'}(z)-1>0$ where $\mathcal{E}_h(z)=zh'(z)/h(z)$ and $\mathcal{E}_{h'}(z)=zh''(z)/h'(z)$. Now the elasticity of substitution σ_{KL} between capital and labor is

$$\sigma_{KL} = \frac{-f_k}{kf_{kk}}\left(1 - \frac{kf_k}{f}\right) = \frac{\mathcal{E}_h\left[k(1-s)^\mu\right] - 1}{\mathcal{E}_{h'}\left[k(1-s)^\mu\right]}. \tag{25}$$

Concavity of h implies $\mathcal{E}_{h'} < 0$ meaning that $\sigma_{KL} < 1$ if and only if $\mathcal{E}_h - \mathcal{E}_{h'} - 1 > 0$ which, as observed above, is equivalent to f being strictly log supermodular.

From the definition of the intensive-form production function in Assumption 1 we have

$$\frac{F_{Ks}F}{F_K F_s} = \frac{f_{ks}f}{f_k f_s} = \frac{\beta - 1 - \mathcal{E}_{h'}\left[k(1-s)^\mu\right]}{\beta - \mathcal{E}_h\left[k(1-s)^\mu\right]}.$$

Substituting this expression and (25) into (21) gives

$$\frac{\partial\left(F_s/F_L\right)}{\partial K} = \frac{F_K F_s}{F_L F}\left(\frac{\beta - 1 - \mathcal{E}_{h'}\left[k(1-s)^\mu\right]}{\beta - \mathcal{E}_h\left[k(1-s)^\mu\right]} - \frac{\mathcal{E}_{h'}\left[k(1-s)^\mu\right]}{\mathcal{E}_h\left[k(1-s)^\mu\right] - 1}\right).$$

Therefore, $\partial\left(F_s/F_L\right)/\partial K > 0$ if and only if the marginal product of each input is positive and $(\beta-1)(\mathcal{E}_h - \mathcal{E}_{h'} - 1) < 0$. Since $\beta < 1$ this inequality holds if and only if f is strictly log supermodular.

Second Order Condition of the Planner's Problem

The planner chooses s_t to maximize Y_t. The first order condition is

$$D'(s_t)D(s_t)^{-\mu\beta}\left\{-\left(\beta\mu - 1\right)h\left[\frac{A_t K_t}{B_t N_t}D\left(s_t\right)^{\mu-1}\right] + \left(\mu - 1\right)h'\left[\frac{A_t K_t}{B_t N_t}D\left(s_t\right)^{\mu-1}\right]\frac{A_t K_t}{B_t N_t}D\left(s_t\right)^{\mu-1}\right\} = 0,$$

and the second order condition is

$$(\mu-1)D(s_t)^{-\mu(\beta-1)-2}\left[D'(s_t)\right]^2\frac{A_t K_t}{B_t N_t}\left\{-(\mu\beta-1)h'(z^*) + (\mu-1)z^*h''(z^*) + (\mu-1)h'(z^*)\right\} < 0.$$

Using $\mu\beta - 1 = (\mu-1)\mathcal{E}_h(z^*)$ the second order condition holds if and only if

$$\mathcal{E}_h(z^*) - \mathcal{E}_{h'}(z^*) - 1 > 0,$$

which holds for any possible z^* if and only if f is strictly log supermodular.

Transition Dynamics of the Planner's Problem

After solving for optimal schooling we can write the planner's problem as

$$\max_{\{c_t\}} \int_{t_0}^{\infty} N_t e^{-\rho(t-t_0)} \frac{c_t^{1-\eta}-1}{1-\eta} dt$$

subject to

$$\dot{K}_t = q_t \left[Y(K_t) - N_t c_t \right] - \delta K_t.$$

where $Y(K_t)$ is given by (4).

Solving this problem we find the planner chooses a consumption path that satisfies

$$\frac{\dot{c}_t}{c_t} = -\frac{\rho + \delta + g_q}{\eta} + \frac{1}{\eta} \frac{\mu\beta - 1}{\mu - 1} q_t \frac{Y(K_t)}{K_t}. \tag{26}$$

Now let $\tilde{Y}_t = e^{-g_Y(t-t_0)} Y(K_t)$, $\tilde{C}_t = e^{-g_Y(t-t_0)} N_t c_t$ and $\tilde{K}_t = e^{-g_K(t-t_0)} K_t$ where g_Y is defined in part (i) of Proposition 2 and $g_K = g_Y + g_q$. Using (26) and the capital accumulation equation together with the fact that q_t, A_t, B_t and N_t grow at constant rates g_q, g_A, γ_L and n, respectively, we have

$$\tilde{Y}_t = \tilde{Y}\left(\tilde{K}_t\right) = A_{t_0}^{\frac{\mu\beta-1}{\mu-1}} \left(B_{t_0} N_{t_0}\right)^{\frac{\mu(1-\beta)}{\mu-1}} {z^*}^{\frac{1-\mu\beta}{\mu-1}} h\left(z^*\right) \tilde{K}_t^{\frac{\mu\beta-1}{\mu-1}},$$

$$\dot{\tilde{C}}_t = \left[-g_Y + n - \frac{\rho + \delta + g_q}{\eta} + \frac{q_{t_0}}{\eta} \frac{\mu\beta - 1}{\mu - 1} \frac{\tilde{Y}(\tilde{K}_t)}{\tilde{K}_t} \right] \tilde{C}_t, \tag{27}$$

$$\dot{\tilde{K}}_t = -(g_Y + g_q + \delta)\tilde{K}_t + q_{t_0} \left[\tilde{Y}\left(\tilde{K}_t\right) - \tilde{C}_t \right]. \tag{28}$$

Since consumption and schooling can jump, K_t (or, equivalently $\tilde{K}_t$) is the economy's only state variable. The pair of differential equations (27) and (28) govern the evolution of the economy from any initial condition K_{t_0}.

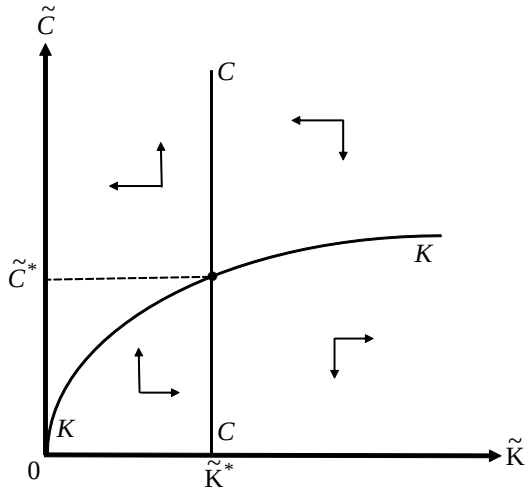

Figure 3: Transitional dynamics and stability of the balanced growth path

Figure 3 depicts a familiar phase diagram. The vertical line labeled CC has $\tilde{K} = \tilde{K}^*$ such that

$$\frac{\tilde{Y}(\tilde{K}^*)}{\tilde{K}^*} = \frac{1}{q_{t_0}} \frac{\mu - 1}{\mu\beta - 1} \left[\eta \left(g_Y - n\right) + \rho + \delta + g_q\right].$$

From (27), we see that $\dot{\tilde{C}}_t = 0$ along this line. The curve labeled KK has $\tilde{C} = \tilde{Y}(\tilde{K}) - (g_Y + g_q + \delta)\tilde{K}/q_{t_0}$. This curve, which from (28) depicts combinations of $\tilde{C}$ and $\tilde{K}$ such that $\dot{\tilde{K}}_t = 0$, can be upward sloping (as drawn) or hump-shaped. In either case, the two curves intersect on the upward sloping part of KK.[32] The intersection gives the unique steady-state values of $\tilde{K} = \tilde{K}^*$ and $\tilde{C} = \tilde{C}^*$, which in turn identify the unique BGP. As is clear from the figure, the BGP is reached by a unique equilibrium trajectory that is saddle-path stable.

Necessity of Functional Form

Consider an economy that satisfies the assumptions required for Lemma 1 to hold and has production function $F\left(K, L, s; t\right)$ which is constant returns to scale in its first two arguments. Suppose factors are paid their marginal products and schooling is chosen to satisfy

$$s_t = \arg\max_s F(K_t, L_t, s; t) \text{ subject to } L_t = D\left(s\right) N_t.$$

We assume this optimization problem has a unique interior maximum.

[32] To see this, note that $\tilde{Y}'\left(\tilde{K}_t\right) = \frac{\mu\beta - 1}{\mu - 1} \frac{\tilde{Y}(\tilde{K}_t)}{\tilde{K}_t}$. Consequently, the slope of the KK curve is $\frac{\mu\beta - 1}{\mu - 1} \frac{\tilde{Y}(\tilde{K}_t)}{\tilde{K}_t} - \frac{g_Y + g_q + \delta}{q_{t_0}}$ which is positive when $\tilde{K} = \tilde{K}^*$ by (3).

Suppose the economy is on a BGP from time T onwards with constant and strictly positive factor shares. With a slight abuse of notation define $\tilde{F}$ by

$$\tilde{F}(K, L, s; t) = \tilde{F}\left[A_t K D(s)^a, B_t L D(s)^{-b}\right] \equiv F\left[A_t K D(s)^a, B_t L D(s)^{-b}, s_T; T\right],$$

where $b = 1 + a\theta_K/(1-\theta_K)$, while A_t and B_t are defined by

$$A_t \equiv e^{g_Y(t-T)} D(s_t)^{-a} \frac{K_T}{K_t},$$
$$B_t \equiv e^{g_Y(t-T)} D(s_t)^{b} \frac{L_T}{L_t}.$$

Since a and b jointly satisfy a single restriction, $\tilde{F}$ defines a one dimensional family of functions.

Differentiating the definitions of A_t and B_t together with the constraint $L_t = D(s_t)N_t$ and using Lemma 1 we obtain

$$\gamma_K \equiv \frac{\dot{A}_t}{A_t} + g_q = a(n - g_L),$$
$$\gamma_L \equiv \frac{\dot{B}_t}{B_t} = g_Y - n - \frac{\theta_K}{1-\theta_K}\gamma_K.$$

γ_K is the total rate of capital-augmenting technical change, while γ_L is the rate of labor-augmenting technical change. When both n and the labor force growth rate g_L are constant then γ_K and γ_L are also constant. Also, provided schooling is increasing over time $n > g_L$ implying that $a > 0$ if and only if γ_K is strictly positive.

We can now prove the following proposition. Part (i) shows that on the BGP F has a one dimensional family of representations of the form $\tilde{F}\left[A_t K D(s)^a, B_t L D(s)^{-b}\right]$. From the expressions for γ_K and γ_L above we see that each member of this family has a different combination of capital-augmenting and labor-augmenting technical change. When we say the production function can be represented by $\tilde{F}$ we mean that the equilibrium allocation on the BGP is the same under $\tilde{F}$ as under F. However, this does not imply that counterfactual experiments using $\tilde{F}$ will necessarily coincide with counterfactuals under F. The first order impact of some policy changes (e.g., schooling subsidies, capital taxation) depends on σ_{KL} and $\sigma_{Ks} \equiv (F_K F_s)/(F_{Ks} F)$. Therefore, in part (ii) of the proposition we show that if σ_{KL} is constant on the BGP then $\sigma_{KL} = \tilde{\sigma}_{KL} \equiv (\tilde{F}_K \tilde{F}_L)/(\tilde{F}_{KL}\tilde{F})$ and that

$\tilde{\sigma}_{Ks} \equiv (\tilde{F}_K \tilde{F}_s)/(\tilde{F}_{Ks} \tilde{F})$ can be written as a function of $\tilde{\sigma}_{KL}$, a and b. Consequently, if σ_{KL} and σ_{Ks} are constant on the BGP then there exist unique values of a and b such that $\tilde{\sigma}_{KL} = \sigma_{KL}$ and $\tilde{\sigma}_{Ks} = \sigma_{Ks}$. Thus, knowing σ_{KL} and σ_{Ks} is sufficient to separate the roles played by capital-augmenting and labor-augmenting technical change. Moreover, when a and b are chosen appropriately counterfactual analysis using $\tilde{F}$ instead of F will, to a first order, give accurate predictions.

Proposition 5 *Suppose for all $t \geq T$ the economy's equilibrium trajectory $\{Y_t, K_t, L_t, s_t\}$ is a BGP with constant and strictly positive factor shares. On the BGP,*

(i) The production function F can be represented by $\tilde{F}$ in the sense that for all $t \geq T$

$$\begin{aligned}
\tilde{F}(K_t, L_t, s_t; t) &= F(K_t, L_t, s_t; t), \\
\tilde{F}_K(K_t, L_t, s_t; t) &= F_K(K_t, L_t, s_t; t), \\
\tilde{F}_L(K_t, L_t, s_t; t) &= F_L(K_t, L_t, s_t; t), \\
\tilde{F}_s(K_t, L_t, s_t; t) &= F_s(K_t, L_t, s_t; t);
\end{aligned}$$

(ii) $\tilde{\sigma}_{KL}$ and $\tilde{\sigma}_{Ks}$ satisfy

$$\frac{1}{\tilde{\sigma}_{Ks}} - 1 = (a+b)\left(\frac{1}{\tilde{\sigma}_{KL}} - 1\right),$$

and if σ_{KL} is constant then $\tilde{\sigma}_{KL} = \sigma_{KL}$.

Proof. Without loss of generality let $T = 0$. Output at $t \geq 0$ is given by

$$\begin{aligned}
F(K_t, L_t, s_t; t) = Y_t = e^{g_Y t} Y_0 = e^{g_Y t} F(K_0, L_0, s_0; 0) &= F\left(e^{g_Y t} K_0, e^{g_Y t} L_0, s_0; 0\right), \\
&= F\left(A_t K_t D(s_t)^a, B_t L_t D(s_t)^{-b}, s_0; 0\right), \\
&= \tilde{F}(K_t, L_t, s_t; t).
\end{aligned}$$

To show the marginal products of capital are equal, we use the facts that the capital share is constant over time and capital is paid its marginal product. Therefore

$$\begin{aligned}\frac{K_t F_K\left(K_t, L_t, s_t; t\right)}{Y_t} = \theta_K = \frac{K_0 F_K\left(K_0, L_0, s_0; 0\right)}{Y_0} &= \frac{e^{g_Y t} K_0 F_K\left(e^{g_Y t} K_0, e^{g_Y t} L_0, s_0; 0\right)}{e^{g_Y t} Y_0}, \\ &= \frac{A_t K_t D\left(s_t\right)^a F_K\left(A_t K_t D\left(s_t\right)^a, B_t L_t D\left(s_t\right)^{-b}, s_0; 0\right)}{Y_t}, \\ &= \frac{K_t \tilde{F}_K\left(K_t, L_t, s_t; t\right)}{Y_t}.\end{aligned}$$

Dividing each side by K_t/Y_t gives $F_K\left(K_t, L_t, s_t; t\right) = \tilde{F}_K\left(K_t, L_t, s_t; t\right)$. Identical logic using the labor share gives $F_L\left(K_t, L_t, s_t; t\right) = \tilde{F}_L\left(K_t, L_t, s_t; t\right)$.

To complete the proof of part (i) we show equality of the marginal products of schooling. Optimal schooling choice implies

$$\frac{D'\left(s_t\right) L_t}{D\left(s_t\right)} = -\frac{F_s\left(K_t, L_t, s_t; t\right)}{F_L\left(K_t, L_t, s_t; t\right)}.$$

This means the ratio of the marginal product of schooling to output can be written as

$$\frac{F_s\left(K_t, L_t, s_t; t\right)}{Y_t} = -\left(1 - \theta_K\right) \frac{D'\left(s_t\right)}{D\left(s_t\right)}.$$

We now show that same equation holds for $\tilde{F}$. Differentiating $\tilde{F}$ with respect to s and dividing by output gives

$$\begin{aligned}\frac{\tilde{F}_s\left(K_t, L_t, s_t; t\right)}{Y_t} &= \frac{1}{Y_t} \frac{D'\left(s_t\right)}{D\left(s_t\right)} \left[a A_t K_t D\left(s_t\right)^a F_K\left(A_t K_t D\left(s_t\right)^a, B_t L_t D\left(s\right)^{-b}, s_0; 0\right)\right. \\ &\qquad \left. -b B_t L_t D\left(s_t\right)^{-b} F_L\left(A_t K_t D\left(s_t\right)^a, B_t L_t D\left(s_t\right)^{-b}, s_0; 0\right)\right], \\ &= \left[a\theta_K - b\left(1 - \theta_K\right)\right] \frac{D'\left(s_t\right)}{D\left(s_t\right)}, \\ &= -(1 - \theta_K) \frac{D'\left(s_t\right)}{D\left(s_t\right)}.\end{aligned}$$

To prove part (ii) we start by noting that when σ_{KL} is constant on the BGP, the homogeneity of F implies

$$
\begin{aligned}
\sigma_{KL} &= \frac{F_K\left(K_0, L_0, s_0; 0\right) F_L\left(K_0, L_0, s_0; 0\right)}{F_{KL}\left(K_0, L_0, s_0; 0\right) F\left(K_0, L_0, s_0; 0\right)}, \\
&= \frac{F_K\left(e^{g_Y t} K_0, e^{g_Y t} L_0, s_0; 0\right) F_L\left(e^{g_Y t} K_0, e^{g_Y t} L_0, s_0; 0\right)}{F_{KL}\left(e^{g_Y t} K_0, e^{g_Y t} L_0, s_0; 0\right) F\left(e^{g_Y t} K_0, e^{g_Y t} L_0, s_0; 0\right)}, \\
&= \frac{F_K\left(A_t K_t D\left(s_t\right)^a, B_t L_t D\left(s_t\right)^{-b}, s_0; 0\right) F_L\left(A_t K_t D\left(s_t\right)^a, B_t L_t D\left(s_t\right)^{-b}, s_0; 0\right)}{F_{KL}\left(A_t K_t D\left(s_t\right)^a, B_t L_t D\left(s_t\right)^{-b}, s_0; 0\right) F\left(A_t K_t D\left(s_t\right)^a, B_t L_t D\left(s_t\right)^{-b}, s_0; 0\right)}, \\
&= \frac{\tilde{F}_K\left(K_t, L_t, s_t; t\right) \tilde{F}_L\left(K_t, L_t, s_t; t\right)}{\tilde{F}_{KL}\left(K_t, L_t, s_t; t\right) \tilde{F}\left(K_t, L_t, s_t; t\right)}, \\
&= \tilde{\sigma}_{KL}.
\end{aligned}
$$

Next define $\hat{h}(z) \equiv F(z, 1, s_0; 0)$. Then we have

$$
\tilde{F}\left(K, L, s; t\right) = B_t L D(s)^{-b} \hat{h}\left[\frac{A_t K}{B_t L} D(s)^{a+b}\right].
$$

Noting the equivalence between this expression and the functional form assumed in Assumption 1 and using reasoning analogous to that employed above to derive the implications of Assumption 1 we have

$$
\begin{aligned}
\tilde{\sigma}_{KL} &= \frac{\mathcal{E}_{\hat{h}}\left[\frac{A_t K}{B_t L} D(s)^{a+b}\right] - 1}{\mathcal{E}_{\hat{h}'}\left[\frac{A_t K}{B_t L} D(s)^{a+b}\right]}, \\
\tilde{\sigma}_{Ks} &= \frac{\frac{b}{a+b} - \mathcal{E}_{\hat{h}}\left[\frac{A_t K}{B_t L} D(s)^{a+b}\right]}{\frac{b}{a+b} - 1 - \mathcal{E}_{\hat{h}'}\left[\frac{A_t K}{B_t L} D(s)^{a+b}\right]}.
\end{aligned}
$$

On the BGP we also have

$$
\theta_K = \frac{K_t \tilde{F}_K\left(K_t, L_t, s_t; t\right)}{Y_t} = \mathcal{E}_{\hat{h}}\left[\frac{A_t K_t}{B_t L_t} D(s_t)^{a+b}\right].
$$

Combining these expressions and using $b = 1 + a\theta_K / (1 - \theta_K)$ we have that on the BGP

$$
\frac{1}{\tilde{\sigma}_{Ks}} - 1 = (a + b)\left(\frac{1}{\tilde{\sigma}_{KL}} - 1\right).
$$

This completes the proof. ■

Transition Dynamics in the "Time-in-School" Model

Start by observing that combining $r_t = R_t/A_t$ with the no arbitrage condition $\iota_t = R_t/p_t + g_p - \delta$ and $p_t = 1/q_t$ gives

$$r_t = \frac{1}{q_t A_t}\left(\iota_t + g_q + \delta\right). \tag{29}$$

Individuals' optimal schooling choices imply $\kappa(s_t, r_t)(1-s_t)^{\mu} = z^*$ for all $t \geq t_0$ where z^* takes the same value as in the planner's problem. Therefore, aggregate output is given by (4) as in the planner's problem.

Using the functional form assumption imposed on f, the first order condition for profit maximization (5) yields

$$r_t = (1-s_t)^{\mu(1-\beta)} h'(z^*).$$

Substituting this expression into the capital market clearing condition $k_t = \kappa(s_t, r_t)$ and using (29) shows the real interest rate satisfies

$$\iota_t = -g_q - \delta + q_t A_t^{\frac{\mu\beta-1}{\mu-1}} \left(\frac{B_t N_t}{K_t} z^*\right)^{\frac{\mu(1-\beta)}{\mu-1}} h'(z^*).$$

Combining this equation with the representative dynasty's Euler equation $\dot{c}_t/c_t = (\iota_t - \rho)/\eta$ and using $\mathcal{E}_h(z^*) = (\mu\beta - 1)/(\mu - 1)$ and (4) gives

$$\frac{\dot{c}_t}{c_t} = -\frac{\rho + \delta + g_q}{\eta} + \frac{1}{\eta}\frac{\mu\beta - 1}{\mu - 1} q_t \frac{Y(K_t)}{K_t}.$$

Noting that this equation is identical to equation (26) we see that consumption per capita satisfies the same differential equation as in the planner's problem. Since the capital accumulation equation is also the same in both cases we conclude that consumption and the aggregate capital stock follow the same equilibrium trajectory as in the planner's problem.

Schooling Choice in the "Manager-Worker" Model

Recall that the production function can be written as $\tilde{F}\left[A_t K D(s)^{\mu(1-\beta)}, B_t L D(s)^{-\mu\beta}\right] = B_t L D(s)^{-\mu\beta} h\left[k D(s)^{\mu}\right]$ where $s = M/L$ and $D(s) = [1 + s/(1-m)]^{-1}$. Since $W_{Mt} = \tilde{F}_M$ and $W_{Lt} = \tilde{F}_L$, differentiating yields

$$W_{Mt} = \mu B_t D(s_t)^{-\mu\beta} \frac{D'(s_t)}{D(s_t)} h\left[k_t D(s_t)^{\mu}\right] \left\{-\beta + \mathcal{E}_h\left[k_t D(s_t)^{\mu}\right]\right\},$$
$$W_{Lt} = B_t D(s_t)^{-\mu\beta} h\left[k_t D(s_t)^{\mu}\right] \left(1 - \mathcal{E}_h\left[k_t D(s_t)^{\mu}\right] + \mu \frac{s_t D'(s_t)}{D(s_t)} \left\{\beta - \mathcal{E}_h\left[k_t D(s_t)^{\mu}\right]\right\}\right).$$

Substituting these expressions into (9) and using $D'(s) = -D(s)^2/(1-m)$ implies that, in equilibrium,

$$\mathcal{E}_h\left[k_t \left(1 + \frac{s_t}{1-m}\right)^{-\mu}\right] = \frac{\mu\beta - 1}{\mu - 1}.$$

The fact that $\mathcal{E}_h(z)$ is declining in z ensures stability of the equilibrium.

Proofs from Section 4

Derivation of Equation (12)

The function $\Phi(z,u)$ is defined as the inverse of $h_z(z,u)$ so that $z = \Phi\left[h_z(z,u), u\right]$. Using Assumption 4 and the definition of Φ, the first-order conditions from profit maximization imply

$$\kappa(s,u,r) = e^{\mu s} \Phi\left[r e^{(1-\beta)\mu s}, u\right],$$
$$\omega(s,u,r) = e^{\mu\beta s} \left\{h\left(\Phi\left[r e^{(1-\beta)\mu s}, u\right], u\right) - r e^{(1-\beta)\mu s} \Phi\left[r e^{(1-\beta)\mu s}, u\right]\right\}.$$

Therefore, the wage at time $b+s+u$ of an individual born at b who has s years of schooling and u years of experience is

$$W_{b+s+u}(s,u) = B_{b+s+u} e^{\mu\beta s} \left\{h\left(\Phi\left[r_{b+s+u} e^{(1-\beta)\mu s}, u\right], u\right) - r_{b+s+u} e^{(1-\beta)\mu s} \Phi\left[r_{b+s+u} e^{(1-\beta)\mu s}, u\right]\right\}.$$

Since B_t grows at rate γ_L and on a BGP r_t declines at rate γ_K it follows that along a BGP the wage function can be written as

$$W_{b+s+u}(s,u) = B_b e^{\gamma_L(s+u)} e^{\mu\beta s} \left\{ h\left(\Phi\left[e^{-\gamma_K u} x_b, u\right], u\right) - e^{-\gamma_K u} x_b \Phi\left[e^{-\gamma_K u} x_b, u\right] \right\},$$

where $x_b \equiv r_b e^{[(1-\beta)\mu - \gamma_K]s}$.

Now consider the optimal choice of schooling. From substituting the wage equation above into (10) it follows that maximizing the expected present discounted value of lifetime earnings is equivalent to choosing x_b to maximize $v(x_b)$ where

$$v(x_b) = x_b^{\psi} \int_0^{\bar{u}} e^{-(\iota+\upsilon-\gamma_L)u} \left\{ h\left(\Phi\left[e^{-\gamma_K u} x_b, u\right], u\right) - e^{-\gamma_K u} x_b \Phi\left[e^{-\gamma_K u} x_b, u\right] \right\} du, \tag{30}$$

and

$$\psi \equiv \frac{\gamma_L + \mu\beta - \iota - \upsilon}{(1-\beta)\mu - \gamma_K}. \tag{31}$$

Note that this maximization problem is independent of time of birth b. Therefore, the solution is the same for every cohort.

Differentiating (30) yields

$$v'(x_b) = (\psi+1) x_b^{\psi-1} \int_0^{\bar{u}} e^{-(\iota+\upsilon-\gamma_L)u} h\left(\Phi\left[e^{-\gamma_K u} x_b, u\right], u\right) \left\{ \frac{\psi}{\psi+1} - \mathcal{E}_{h,z}\left(\Phi\left[e^{-\gamma_K u} x_b, u\right], u\right) \right\} du,$$

and

$$\begin{aligned} v''(x_b) \quad = \quad & (\psi-1)\frac{v'(x_b)}{x_b} + (\psi+1) x_b^{\psi-1} \int_0^{\bar{u}} e^{-(\iota+\upsilon+\gamma_K-\gamma_L)u} \Phi_z\left[e^{-\gamma_K u} x_b, u\right] \left\{ e^{-\gamma_K u} x_b \right. \\ & \left. \left[\frac{\psi}{\psi+1} - \mathcal{E}_{h,z}\left(\Phi\left[e^{-\gamma_K u} x_b, u\right], u\right) \right] - h\left(\Phi\left[e^{-\gamma_K u} x_b, u\right], u\right) \frac{\partial \mathcal{E}_{h,z}\left(\Phi\left[e^{-\gamma_K u} x_b, u\right], u\right)}{\partial z} \right\} du. \end{aligned}$$

We assume that if the first order condition $v'(x_b) = 0$ has a solution x^* then the second order condition $v''(x^*) < 0$

holds. Since $v'(x_b)$ is continuous it follows that if a solution to the first order condition exists, this solution is unique and maximizes $v(x_b)$.

Substituting $\iota = \eta(g_Y - n) + \rho$ and (17) into (31) shows that on a balanced growth path

$$\psi = \frac{1}{(1-\beta)\mu - \gamma_K}\left[n - \rho + (1-\eta)\gamma_L + (\mu\beta - \lambda)\left(1 - \frac{\eta\gamma_K}{(1-\beta)\mu}\right)\right] = \Omega,$$

where Ω is defined in Assumption 5. Consequently, part (ii) of Assumption 5 implies $d_{\min} < \psi/(\psi+1) < d_{\max}$. Since $h(z,u)$ is strictly concave in z we must have $\Phi_z(z,u) < 0$. Recalling the definitions of $d_{\min}$ and $d_{\max}$ it then follows that $\frac{\psi}{\psi+1} > \mathcal{E}_{h,z}\left(\Phi\left[e^{-\gamma_K u}x_b, u\right], u\right)$ for all $u \in [0, \bar{u})$ when x_b is chosen sufficiently small and $\frac{\psi}{\psi+1} < \mathcal{E}_{h,z}\left(\Phi\left[e^{-\gamma_K u}x_b, u\right], u\right)$ for all $u \in [0, \bar{u})$ when x_b is chosen sufficiently large. Therefore, continuity of $v'(x_b)$ guarantees the first order condition has a solution. This solution x^* satisfies (12) for all b.

Capital Share

The capital share, θ_K, is given by (19). Using the first order condition $v'(x^*) = 0$ we can rewrite this equation as

$$\theta_K = \frac{\psi}{\psi+1}\frac{\int_0^{\bar{u}} e^{-(\iota-g_Y)u}e^{-\left[\lambda+\frac{(\mu\beta-\lambda)\gamma_K}{(1-\beta)\mu}\right]u}h\left\{\Phi\left[e^{-\gamma_K u}x^*, u\right], u\right\}du}{\int_0^{\bar{u}} e^{-(\iota-g_Y)u}e^{-\left[\lambda+\frac{(\mu-\lambda)\gamma_K}{(1-\beta)\mu}\right]u}x^*\Phi\left[e^{-\gamma_K u}x^*, u\right]du}\frac{\int_0^{\bar{u}} e^{-\left[\lambda+\frac{(\mu-\lambda)\gamma_K}{(1-\beta)\mu}\right]u}x^*\Phi\left[e^{-\gamma_K u}x^*, u\right]du}{\int_0^{\bar{u}} e^{-\left[\lambda+\frac{(\mu\beta-\lambda)\gamma_K}{(1-\beta)\mu}\right]u}h\left\{\Phi\left[e^{-\gamma_K u}x^*, u\right], u\right\}du},$$

where ψ is given by (31). Finally, taking the limit as $\iota - g_Y$ converges to zero from above, we have

$$\lim_{\iota - g_Y \searrow 0}\theta_K = \lim_{\iota - g_Y \searrow 0}\frac{\psi}{\psi+1} = \frac{\mu\beta - \lambda}{\mu - \lambda}.$$