NBER WORKING PAPER SERIES

TECHNOLOGY AND SKILL:
TWIN ENGINES OF GROWTH

Nancy L. Stokey

Working Paper 24570
http://www.nber.org/papers/w24570

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2018

An early version of this paper titled “Technology, Skill and Long Run Growth” was given as the Lionel McKenzie Lecture at the SAET Meetings, Rio de Janeiro, 2016. I am grateful to Mark Aguiar, Marios Angeletos, Gadi Barlevy, Paco Buera, Ariel Burstein, Bob Lucas, Jesse Perla, Rob Shimer, Chris Tonetti, and participants at the M&M Development Workshop at the Chicago Fed, the Minnesota Summer Workshop in Macroeconomics, the 2017 SED Meetings, and the IMF for useful comments. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Technology and Skill: Twin Engines of Growth
Nancy L. Stokey
NBER Working Paper No. 24570
May 2018
JEL No. O30,O33,O34,O40

ABSTRACT

A model is developed in which two complementary forms of investment contribute to growth—technology and skill acquisition, and growth takes two forms—TFP and variety growth. The rate of TFP growth depends more heavily on the parameters governing skill accumulation, while variety growth depends, roughly, on the difference between the parameters governing technology and skill accumulation. Conditions for the existence of a BGP are established, and the effects of various parameters are characterized. In an example, subsidies to skill acquisition (technology acquisition) are powerful tools for stimulating TFP growth (variety growth). Investment incentives off the BGP are also explored.

Nancy L. Stokey
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
and NBER
nstokey@uchicago.edu

1. INTRODUCTION

This paper develops a model in which two factors contribute to growth: investments in technology by heterogeneous firms and investments in human capital by heterogeneous workers. Income growth in turn takes two forms: growth in the quantity produced of each differentiated good and growth in the number of goods available. Call these two forms total factor productivity (TFP) growth and variety growth.

Both types of investment affect both forms of growth, but the contributions are not symmetric. Improvements in the parameters governing investment in skill raise the rate of TFP growth and reduce the rate of variety growth. Improvements in the parameters governing investment in technology raise the rate of variety growth, while the effect on TFP growth is positive, zero, or negative as the elasticity of intertemporal substitution (EIS) is greater than, equal to or less than unity.

This asymmetry appears despite the fact that skill and technology are modeled as symmetric in many respects, and on balanced growth paths (BGPs) the rate of TFP growth is also the (common) growth rate of technology and human capital. But the factors are fundamentally different in two respects. First, human capital is a rival input while technology is nonrival. That is, an increase in a worker's human capital affects only his own productivity, while an improvement in a firm's technology can be exploited by all its workers.

In addition, the two factors differ in the way entry occurs. Growth in the size of the workforce is exogenous. Entry by new firms is endogenous, and entering firms must invest to obtain technologies for new goods. Thus, the expected profitability of a new product affects the incentives of entrants, and the entry rate is governed by a zero-profit condition.

Analyzing both types of investment together is important because there is strategic complementarity in the incentives to invest. Incumbent workers invest in skill to increase their wages. But without continued improvement in the set of technologies used by firms, the returns to workers' investments would decline and, eventually, be too small to justify further investment. Similarly, incumbent firms invest in better technologies to increase their profits, but without continued improvement in the skill distribution of the workforce, their returns would eventually be too small to justify further investments. Sustained growth requires continued investment in both factors, and the contribution of this paper is to characterize the interplay between the two types of investment. It also suggests why empirical work based on regression analysis is likely to fail: if two factors are highly complementary, a linear framework will have difficulty picking up their effects.

The rest of the paper is organized as follows. Related literature is discussed in section 2. Section 3 sets out the production technologies and characterizes the (static) production equilibrium. The production function for differentiated intermediates has two inputs, technology and human capital, and it is log-supermodular. Hence the competitive equilibrium features positively assortative matching between technology and skill. Proposition 1 establishes the existence, uniqueness and efficiency of a production equilibrium, describing the allocation of labor across technologies and the resulting prices, wages, output levels, and profits. Lemmas 2 and 3 establish some homogeneity properties. The first main result, Proposition 4, shows that if the technology and skill distributions are Pareto, with locations that are appropriately aligned, then the equilibrium allocation of skill to technology is linear, and the wage, price, output, and profit functions are isoelastic.

?

Section 4 treats dynamics: the investment decisions of incumbent firms, new entrants, and workers; the evolution of the technology and skill distributions; and the interest rate and consumption growth. Section 5 provides formal definitions of a competitive equilibrium and a balanced growth path. A balanced growth path features stationary, nondegenerate distributions of relative technology and relative human capital, with both growing at a common, constant rate.

Section 6 specializes to the case where technology and skill have Pareto distributions, showing that the isoelastic forms for the profit and wage functions are inherited by the value functions for producers and workers. This fact leads to a tractable set of conditions describing investment and the evolution of the technology and skill distributions on a BGP. The second main result, Proposition 5, provides conditions that ensure the existence of a BGP.

Section 7 looks at the effects of various parameters and policies. Proposition 6, the third main result, describes the effects of parameter changes on TFP and variety growth. Because investments of both types have important positive external effects, (!) the competitive equilibrium investment rates are presumably too low, and the effects of subsidies to investment are described in Corollary 1. In a roughly calibrated example, such subsidies are very powerful.

Section 8 examines the incentives to invest off the BGP. Proposition 8 shows, in a simple setting, that the incentive increases (decreases) for the lagging (leading) factor, suggesting that growth in one factor alone cannot be sustained in the long run.

Section 9 concludes. Proofs and technical derivations and arguments are gathered in Appendices.

2. RELATED LITERATURE

This paper is related to two literatures: on diffusion of ideas/technologies and on technology-skill complementarity. Each has both theoretical and empirical components.

The idea that technology diffusion is an important factor for growth has a long history, and in an early contribution Nelson and Phelps (1966) emphasize the role of educated labor in facilitating such diffusion. Jovanovic and Rob (1989) look at an aggregate model with heterogeneous agents, where meetings between agents can generate both new ideas and imitation—diffusion of an idea from the initially more productive agent to one that is less productive. Jovanovic and MacDonald (1994) look at innovation and diffusion in a competitive industry, where firm-level incentives are influenced by the distribution of technologies across their rivals.

Kortum (1997) introduces the notion of a technology frontier as a description of the state of knowledge in a society, and Lucas (2009, 2015) develops models where that frontier evolves as a consequence of meetings between agents that result in the transfer of knowledge. Lucas and Moll (2014) extend those models to allow agents to divide their available time between production and knowledge acquisition, and Caicedo, Lucas and Rossi-Hansberg (2016) embed idea diffusion into a model of hierarchies. Benhabib, Perla, and Tonetti (2017) analyze a model with innovation as well as diffusion. Models of learning by doing at the firm, industry, or society level, as in Arrow (1962), Stokey (1988), and Bahk and Gort (1993) can also be interpreted as a type of diffusion.

There is also an extensive literature looking at the diffusion of technologies or ideas across international boundaries as engines of growth, including Eaton and Kortum (1999), Parente and Prescott (1999), Alvarez, Buera and Lucas (2014), Perla, Tonetti and Waugh (2015), Stokey (2015), Buera and Oberfield (2017). See Buera and Lucas (2017) for a description of some of the key models of technology diffusion.

The model here is most closely related to Perla and Tonetti (2014). Indeed, the investment model used here, separately for workers and firms, is exactly the one developed there, where the single type of agent can equally well be interpreted as a firm or a worker. The change here is to introduce a production function that uses both inputs and displays complementarity.

Empirical studies of technology diffusion have taken several approaches, looking at geographic diffusion within a country, diffusion across firms within an industry, and cross-country diffusion. Geographic diffusion of agricultural innovations, where inputs and outputs are easily measured, include Griliches' (1957) early study of hybrid corn in the U.S., and Foster and Rosenzweig's studies of the introduction of high yielding varieties in India during the green revolution (1996) and of adoption in a broader set of countries (2010). In both cases they find that schooling is an important factor in explaining differences. Manuelli and Seshadri (2014) look at the slow diffusion of tractors in U.S. agriculture, while Allen, Bilir, and Tonetti (2017) look at the geographic diffusion of prescriptions for statins across physicians in the U.S.

Mansfield (1961) looks at the diffusion of twelve major innovations across large firms in four industries, finding wide variation in the speed of adoption, much of which can be explained by the profitability of the innovation and the cost of adoption. Gort and Klepper (1982) look at the diffusion of 46 new products through entry and exit of producers, identifying various stages in the product life-cycle that include an early period of rapid entry and a later period of substantial net exit.

Comin and Hobijn look at cross-country diffusion for specific technologies over two centuries, in 23 industrial economies (2004) and in 166 countries covering the full range of income levels (2010). In the former they find that human capital is an important factor influencing speed of adoption. In Comin and Hobijn (2017) they build on the latter, looking at adoption lags and intensity of use. They conclude that differences in adoption lags account for much of the cross-country divergence in incomes during the nineteenth century, while differences in intensity of use account for further divergence in the twentieth century.

Since many new technologies are 'embodied' in new capital goods, technology-skill and capital-skill complementarity are to a large extent two labels for the same phenomenon, difficult to distinguish either conceptually or empirically. Griliches (1969) introduced the notion of capital-skill complementarity in a three-input demand model with capital and two types of labor, and found evidence in its favor in cross-industry U.S. data. Since then the idea has been incorporated into many types of models, including models of long run growth.[1] Some of these drop the physical capital component, looking at technology-skill complementarity, and some retain it, often in the form of vintage capital.

Jovanovic and Nyarko (1996) explore a single-agent model of learning-by-doing and technology choice, where the agent switches to better technologies as her skill improves, while Greenwood and Yorokoglu (1997) investigate the hypothesis that adoption of new technologies requires significant learning, and that skilled workers facilitate that learning. They find evidence going back to the period of the Industrial Revolution that supports this idea. Caselli (1999) studies a vintage capital model with overlapping generations, looking at the adoption of a single new technology that is more productive but requires more training for workers. In his model use of the new technology increases gradually. If it is a sufficiently large improvement, the new technique eventually displaces the old one, but only in the long run. In a similar vein Jovanovic (2009) looks at the spread of technology in a frictionless market with heterogenous labor, where low-skill agents may prefer to use older technologies because they are less expensive.

Acemoglu (1998) suggests that reverse causation may also play a role, as over time newer technologies are designed to complement the rising skills of the workforce. See Acemoglu (2002) for a more comprehensive discussion of the extensive literature on

[1] In other applications, capital-skill complementarity has been used by Stokey (1996) to analyze the wage effects of lowering trade barriers between countries with different aggregate input mixes; by Krusell et al. (2000) in a dynamic model with investment-specific technical change to explain the increasing skill premium; and by Costinot and Vogel (2010) and Stokey (2016) in static models to look at the general equilibrium wage effects of various types of technical changes. Labeled as skill-biased technical change, it has also been offered as the main source of the rising skill premium in the 1970's, 1980's and 1990's, as in Berman, Bound, and Griliches (1994); Dunne, Haltiwanger, and Troske (1996); Bresnahan, Brynjolfsson and Hitt (2002), Violante (2002), and Autor, Levy and Murnane (2003)).

skill-biased technical change.

Nelson, et. al. (1967) suggest that the educational requirement of a technology declines as the technology ages, emphasizing the role of high-skill labor in diffusion. The idea is that experience—transferable learning-by-doing—allows substitution toward less skilled workers. Thus, short-run and long-run effects may differ. Nevertheless, sectors with faster rates of technical change should employ relatively more highly educated labor, an implication that has been studied empirically.

Bartel and Lichtenberg (1987) examine it using cross-industry data for 61 U.S. manufacturing industries in 1960, 1970, 1980. They find that the share of labor cost accounted for by workers with more than a high school education declines with the average age of the equipment in that industry. Doms, Dunne and Troske (1997) find in cross-section that high-tech plants employ more skilled workers, and in time series that plants with more skilled workers are more likely to adopt new technologies. Bartel, Ichniowski and Shaw (2007) find that adoption of IT-enhanced capital equipment is associated with increases in skill requirements for the operators of that equipment.

Goldin and Katz (1998) report evidence of capital-skill and technology-skill complementarity in manufacturing in the U.S. in the early nineteenth century, where they were related to the adoption of electric motors and to the transition from artisanal shops to factories and assembly lines.

More recently, two particular innovations have been examined for evidence of technology-skill complementarity, computers and internet access. Studies of the former have used cross-industry data as in Autor, Katz and Krueger (1998) and Beaudry, Doms, and Lewis (2010) for the U.S. and Machin and Van Reenen (1998) for seven OECD countries. Internet access has been studied using geographic variation, as in Forman et al. (2012) looking at county-level U.S. data, and Akerman et al. (2015) using data from a broadband rollout in Norway. All find evidence of technology-skill complementarity.

Only a few papers to date look at multiple engines of growth, all focusing on skill and technology as the two factors. In an early contribution, Jovanovic (1998) develops a vintage capital model, where later vintages are more productive, workers are heterogeneous in terms of skill, and each worker is assigned to one machine.

The model of long-run growth in Lloyd-Ellis and Roberts (2002) features overlapping generations, who acquire skill to implement and invent new technologies. Like the model here, it features complementarity between investments in skill and technology. But those investments are modeled as schooling and R&D decisions, in contrast to the diffusion model here.

The model in Stokey (2018) is similar in many respects to the one here. The main difference is the way investments in technical change and human capital are treated. There, firms have technologies that evolve as geometric Brownian motions, as in Luttmer (2007), and the firm's investment rate controls the drift of the process. Workers' skills evolve in the same way, with an investment decision controlling the drift. Fairly strong joint restrictions on the investment cost functions are needed to obtain existence a balanced growth path. vs. here?

Goldin and Katz (2008) provide a comprehensive review of the evidence on the co-evolution of education and technology in the U.S. over the twentieth century.

3. PRODUCTION AND PRICES

The single final good is produced by competitive firms using intermediate goods as inputs. Intermediate goods are produced by heterogeneous, monopolistically competitive firms. Each intermediate firm produces a unique variety, and all intermediates enter symmetrically into final good production. But intermediate firms differ in their technology level x, which affects their productivity. Let N_p be the number (mass) of intermediate good producers, and let $F(x)$, with with continuous density f, denote the distribution function for technology.

Intermediate good producers use heterogeneous labor, differentiated by its human capital level h, as the only input. Let L_w, be the size of the workforce, and let $\Psi(h)$, with continuous density ψ, denote the distribution function for human capital. This section looks at the the allocation of labor across producers, and wages, prices, output levels, and profits, given N_p, F, L_w, Ψ.

A. Technologies

Each final good producer has the CRS technology

$$y_F = \left[N_p^{1-\chi} \int y(x)^{(\rho-1)/\rho} f(x) dx \right]^{\rho/(\rho-1)}, \tag{1}$$

where $\rho > 1$ is the substitution elasticity and $\chi \in (0, 1/\rho]$ measures diminishing returns to increased variety. Let $p(x)$ denote the price charged by a producer with technology x. Then input demands are

$$y^d(x) = N_p^{-\rho\chi} p(x)^{-\rho} y_F, \qquad \text{all } x,$$

where the price of the final good is normalized to unity,

$$1 = p_F = \left[N_p^{1-\rho\chi} \int p(x)^{1-\rho} f(x) dx \right]^{1/(1-\rho)}. \tag{2}$$

The output of a firm depends on the size and quality of its workforce, as well as its technology. In particular, if a producer with technology x employs ℓ workers with human capital h, then its output is

$$y = \ell \phi(h, x),$$

where $\phi(h, x)$ is the CES function

$$\phi(h, x) \equiv \left[\omega h^{(\eta-1)/\eta} + (1-\omega) x^{(\eta-1)/\eta} \right]^{\eta/(\eta-1)}, \qquad \eta, \omega \in (0, 1). \tag{3}$$

The elasticity of substitution between technology and human capital is assumed to be less than unity, $\eta < 1$. Firms could employ workers with different human capital levels, and in this case their outputs would simply be summed. In equilibrium firms never choose to do so, however, and for simplicity the notation is not introduced.

B. Intermediate goods: price, output, labor

Let $w(h)$ denote the wage function. For a firm with technology x, the cost of producing a unit of output with labor of quality h is $w(h)/\phi(h,x)$. Optimal labor quality $h^*(x)$ minimizes this expression, so h^* satisfies

$$\frac{w'(h^*)}{w(h^*)} = \frac{\phi_h(h^*,x)}{\phi(h^*,x)}. \tag{4}$$

It is straightforward to show that if the (local, necessary) second order condition for cost minimization holds, then $\eta < 1$ implies h^* is strictly increasing in x. The labor market is competitive, and since the production function in (3) is log-supermodular, efficiency requires positively assortative matching (Costinot, 2009).

Unit cost

$$c(x) = \frac{w(h^*(x))}{\phi(h^*(x),x)},$$

is strictly decreasing in x,

$$\frac{c'(x)}{c(x)} = -\frac{\phi_x(h^*(x),x)}{\phi(h^*(x),x)} < 0.$$

As usual, profit maximization by intermediate good producers entails setting a price that is a markup of $\rho/(\rho-1)$ over unit cost. Output is then determined by demand, and labor input by the production function. Hence price, quantity, labor input, and operating profits for the intermediate firm are

$$\begin{aligned}
p(x) &= \frac{\rho}{\rho-1}\frac{w(h^*(x))}{\phi(h^*(x),x)}, \\
y(x) &= y_F N_p^{-\rho\chi} p(x)^{-\rho}, \\
\ell(x) &= \frac{y(x)}{\phi(h^*(x),x)}, \\
\pi(x) &= \frac{1}{\rho}p(x)y(x), \qquad \text{all } x,
\end{aligned} \tag{5}$$

where the price normalization requires (2). Firms with higher technology levels x have lower prices, higher sales, and higher profits. They may or may not employ more labor.

Each worker inelastically supplies one unit of labor. Let x_m and h_m denote the lower bounds for the supports of F and Ψ. Then markets clear for all types of labor if

$$\begin{aligned} h_m &= h^*(x_m), & (6) \\ L_w\left[1-\Psi(h^*(x))\right] &= N_p \int_x^\infty \ell(\xi) f(\xi) d\xi, \qquad \text{all } x \geq x_m. & (7) \end{aligned}$$

C. Production equilibrium

At any instant, the economy is described by its production parameters, the number of firms and workers, and the distributions of technology and skill.

Definition: A *production environment* $\mathcal{E}_p$ is described by

i. parameters $(\rho, \chi, \omega, \eta)$, with $\rho > 1$, $\chi \in (0, 1/\rho]$, $\omega \in (0,1)$, $\eta \in (0,1)$;

ii. numbers of producers and workers $N_p > 0$ and $L_w > 0$;

iii. distribution functions $F(x)$ with continuous density $f(x)$ and lower bound $x_m \geq 0$ on its support, and $\Psi(h)$ with continuous density $\psi(h)$ and lower bound $h_m \geq 0$ on its support.

A production equilibrium consists of price functions and an allocation that satisfy profit maximization and labor market clearing.

Definition: Given a production environment $\mathcal{E}_p$, the prices $w(h), p(x)$, and allocation $h^*(x)$, $y(x)$, $\ell(x)$, $\pi(x)$, y_F, are a *production equilibrium* if (2) and (4)-(7) hold.

The following result is then straightforward.

Proposition 1: For any production environment $\mathcal{E}_p$, an equilibrium exists, and it is unique and efficient.

D. Homogeneity properties

The analysis of BGPs will exploit the fact that production equilibria have certain homogeneity properties. Lemma 2 deals with proportionate shifts in the two distribution functions.

LEMMA 2: Fix $\mathcal{E}_p$, and let $\mathcal{E}_{pA}$ be a production environment with the same parameters $(\rho, \chi, \omega, \eta)$ and numbers N_p, L_w, but with distribution functions F_A, Ψ_A satisfying

$$\begin{aligned} F_A(X) &= F(X/Q), \qquad \text{all } X, \\ \Psi_A(H) &= \Psi(H/Q), \qquad \text{all } H. \end{aligned}$$

If $[w, p, h^*, y, \ell, \pi, y_F]$ is the production equilibrium for $\mathcal{E}_p$, then the equilibrium for $\mathcal{E}_{pA}$ is

$$\begin{array}{ll} w_A(H) = Qw(H/Q), & p_A(X) = p(X/Q) \\ h_A^*(X) = Qh^*(X/Q), & y_A(X) = Qy(X/Q), \\ \ell_A(X) = \ell(X/Q), & \pi_A(X) = Q\pi(X/Q), \\ y_{FA} = Qy_F, & \text{all } X, H. \end{array}$$

Price and employment for any firm depend only on its relative technology $x = X/Q$, while its labor quality, output, and profits are scaled by Q. Wages and final output are also scaled by Q.

Lemma 3 deals with the effects of changes in the numbers of producers and workers. The impact of variety growth depends on

?

$$\Omega \equiv \frac{1 - \rho\chi}{\rho - 1}, \tag{8}$$

where $\Omega \in [0, 1/(\rho - 1))$. In the limiting case $\chi = 1/\rho$, growth in variety is not valued and $\Omega = 0$.

Lemma 3: Fix $\mathcal{E}_p$, and let $\mathcal{E}_{pB}$ be a production environment with the same parameters and distribution functions, but with $L_{wB} = e^{\upsilon} L_w$ and $N_{pB} = e^n N_p$. If $[w, p, h^*, y, \ell, \pi, y_F]$ is the production equilibrium for $\mathcal{E}_p$, then the equilibrium for $\mathcal{E}_{pB}$ is

$$\begin{aligned} w_B &= e^{\Omega n} w, & p_B &= e^{\Omega n} p, \\ h_B^* &= h^*, & y_B &= e^{\upsilon - n} y, \\ \ell_B &= e^{\upsilon - n} \ell, & \pi_B &= e^{\upsilon + (\Omega - 1) n} \pi, \\ y_{FB} &= e^{\upsilon + \Omega n} y_F, & &\text{all } X, H. \end{aligned}$$

A change in L_w leads to proportionate changes in employment, output and profits at each firm and in final output, with wages, prices and the allocation of skill to technology unaffected.

An increase in N_p leads to proportionate decreases in employment and output at each firm. Final output, the price of each intermediate, and all wage rates change with an elasticity of $\Omega \geq 0$. Thus, all increase if variety is valued, if $\Omega > 0$, and all are unchanged if it is not, if $\Omega = 0$.

Profits per firm—which reflect both the increase in price and decrease in scale—can change in either direction. If $\Omega > 1$, then the love of variety is strong enough so that an increase in the number of producers actually increases the profit of each incumbent. This case occurs only if $\rho < 2$ and, in addition, the parameter χ is not too large. In the analysis of BGPs we will impose the restriction $\rho \geq 2$, to rule out this case.

E. Pareto distributions

In this section we will show that if the distribution functions F and Ψ are Pareto, with shape parameters that are not too different and location parameters that are appropriately aligned, the production equilibrium has a linear assignment of skill to technology, and wage, price, and profit functions that are isoelastic.

PROPOSITION 4: Let $\mathcal{E}_p$ be a production environment for which F and Ψ are Pareto distributions with parameters (α_x, x_m) and (α_h, h_m). Assume that $\alpha_x > 1$, $\alpha_h > 1$, and

$$-1 < \alpha_x - \alpha_h < \rho - 1. \tag{9}$$

Define

$$\begin{aligned} \varepsilon &\equiv \frac{1}{\rho}\left(1 + \alpha_x - \alpha_h\right), && (10)\\ \zeta &\equiv \varepsilon + \alpha_h - \alpha_x, && (11)\\ a_h &\equiv \left(\frac{1-\varepsilon}{\varepsilon}\frac{1-\omega}{\omega}\right)^{\eta/(\eta-1)}, && (12) \end{aligned}$$

and in addition, assume

$$h_m = a_h x_m. \tag{13}$$

The production equilibrium for $\mathcal{E}_p$ has price and allocation functions

$$\begin{aligned} h^*(x) &= a_h x, && \text{all } x, && (14)\\ w(h) &= w_2\left(\frac{h}{h_m}\right)^{1-\varepsilon}, && \text{all } h, && (15)\\ y_F &= L_w N_p^{\Omega} p_0^{\rho}\phi(a_h, 1)\frac{\alpha_h}{\alpha_x}x_m, && && (16)\\ p(x) &= N_p^{\Omega} p_0\left(\frac{x}{x_m}\right)^{-\varepsilon}, && \text{all } x, && (17)\\ y(x) &= \ell_2\phi(a_h, 1)x\left(\frac{x}{x_m}\right)^{\alpha_x-\alpha_h}, && \text{all } x, &&\\ \ell(x) &= \ell_2\left(\frac{x}{x_m}\right)^{\alpha_x-\alpha_h}, && \text{all } x, &&\\ \pi(x) &= \pi_2\left(\frac{x}{x_m}\right)^{1-\zeta}, && \text{all } x, && \end{aligned}$$

where

$$w_2 \equiv \frac{\rho-1}{\rho}L_w^{-1}y_F\frac{\alpha_x}{\alpha_h}p_0^{1-\rho}, \tag{18}$$

$$\begin{aligned}
\ell_2 &\equiv L_w N_p^{-1} \frac{\alpha_h}{\alpha_x}, \\
\pi_2 &\equiv \frac{1}{\rho} N_p^{-1} y_F p_0^{1-\rho}, \\
p_0^{\rho-1} &\equiv \mathrm{E}_F \left(\frac{x}{x_m} \right)^{1-\zeta}.
\end{aligned}$$

The shape parameters α_x and α_h need not be the same, but (9) puts a restriction on how different they can be. It implies that $1-\varepsilon \in (0,1)$ and $1-\zeta \in (0, \rho-1)$, so both the wage and profit functions are strictly increasing.

Pareto distributions for skill and technology imply that wages and profits also have Pareto distributions, with tail parameters $\alpha_w \equiv \alpha_h / (1-\varepsilon)$ and $\alpha_\pi = \alpha_x / (1-\zeta)$, respectively. Thus, α_w is increasing in α_x and is increasing or decreasing in α_h as $\rho > 1+\alpha_x$ or $\rho < 1+\alpha_x$, while α_π is increasing in α_h and decreasing in α_x. Employment has a Pareto distribution if and only if $\alpha_x > \alpha_h$, and in this case it has tail parameter $\alpha_\ell = \alpha_x / (\alpha_x - \alpha_h)$. If $\alpha_h = \alpha_x$, employment is uniform across firm types, and if $\alpha_x < \alpha_h$ employment declines with technology level. Thus, evidence on the size distribution of firms by employment suggests $\alpha_x > \alpha_h$ is the relevant case.

CONT HERE

4. DYNAMICS

In this section the dynamic aspects of the model are described: investment decisions of incumbent producers and workers, the entry decisions of new firms, the evolution of the distribution functions for technology and skill, and the consumption/saving decisions of households. As in Perla and Tonetti (2014) investment is imitative, and it is a zero-one decision. The Pareto shape for the technology and skill distributions this investment technology requires for balanced growth fits well with the production environment here.

It is useful to start with a brief overview. Time is continuous and the horizon is infinite. At any date $t \geq 0$, there are three groups of firms: producers, process innovators and product innovators. A producer can at any time abandon its current technology and become a process innovator, attempting to acquire a new technology. The only cost is the opportunity cost: process innovators do not produce. Success is stochastic, with a fixed hazard rate, and conditional on success the process innovator receives a technology that is a random draw from those of current producers. Hence producers become process innovators if and only if their technology lies below an endogenously determined threshold.

New firms, product innovators, arrive at an endogenously determined rate. Each entrant chooses a one-time (sunk) investment level, which determines its hazard rate for success. After paying the sunk cost, product innovators are like process innovators except that their hazard rate is different.

Similarly, at any date the labor force has three groups: workers, retoolers and trainees. A worker can at any time become a retooler, attempting to acquire a new skill, and the only cost is an opportunity cost—retoolers do not work. Success is stochastic, with a fixed hazard rate, and conditional on success the retooler receives a skill that is a random draw from those of current workers. Hence workers become retoolers if and only if their skill lies below an endogenously determined threshold.

The workforce grows at an exogenously fixed rate υ. New entrants, trainees, are like retoolers except that their hazard rate for success may be different.

Note that there is an important asymmetry between firms and workers: the workforce grows at an exogenously fixed rate, while entry by new firms is endogenous, satisfying a free entry condition.

At date t, $N_p(t), N_i(t), N_e(t)$ are the numbers of producers, process innovators, and product innovators, with sum $N(t)$; $L_w(t), L_i(t), L_e(t)$ are the numbers of workers, retoolers, and entrants, with sum $L(t)$; $F(X,t)$, $\Psi(H,t)$ are the distribution functions for technology among producers and skill among workers; $W(H,t)$, $H^*(X,t)$, $P(X,t)$, $Y(X,t)$, $\mathfrak{L}(X,t)$, $\Pi(X,t)$, $Y_F(t)$, $t \geq 0$, are the wage function, skill allocation, and so on; and $r(t)$ is the interest rate. Note that only producers and workers are identified by a technology or skill level.

A. Firms: process and product innovation

Let $V^f(X,t)$ denote the value of a producer with technology X at date t. A firm that chooses to become a process innovator abandons its current technology and waits to acquire a new one. A process innovator pays no direct costs: there is only the opportunity cost of forgone profits. Abandoned technologies cannot be reclaimed, so all process innovators at date t are in the same position. Let $V_{fi}(t)$ denote their (common) value.

Success is stochastic, arriving at rate λ_{xi}, and conditional on success at date t, the innovator gets a new technology that is random draw from the distribution $F(\cdot,t)$ among current producers. Hence $V_{fi}(t)$ satisfies the Bellman equation

$$[r(t)+\delta_x]\, V_{fi}(t) = \lambda_{xi}\left\{\mathrm{E}_{F(\cdot,t)}[V^f(X,t)] - V_{fi}(t)\right\} + V'_{fi}(t), \qquad \text{all } t,$$

where the term in braces is the expected gain in value conditional on success, $r(t)$ is the interest rate, and $\delta_x \geq 0$ is an exogenous exit rate.

The value $V^f(X,t)$ of a producer is the expected discounted value of its future profit flows. Clearly V^f is nondecreasing in its first argument: a better technology can only raise the firm's value. Hence at any date t, producers with technologies below some threshold $X_m(t)$ become process innovators, while those with technologies above the threshold continue to produce. It follows that at date t, the value of a producer with technology X is $V_{fi}(t)$ if $X \leq X_m(t)$, and the irreversibility of investment means that $X_m(t)$ is nondecreasing. While a firm produces, its technology X grows (or declines) at a constant rate μ_x. Hence the value $V^f(X,t)$ of a producer, a firm with $X > X_m(t)$, satisfies the Bellman equation[2]

$$[r(t) + \delta_x]\, V^f(X,t) = \Pi\,(X,t) + \mu_x X V^f_X(X,t) + V^f_t(X,t), \qquad \text{all } t.$$

Value matching provides a boundary condition for this ODE, and the optimal choice about when to invest implies that smooth pasting holds. Hence

$$\begin{aligned} V^f[X_m(t),t] &= V_{fi}(t), \\ V^f_X[X_m(t),t] &= 0, \qquad\qquad \text{all } t. \end{aligned}$$

Entering firms—product innovators—have a similar investment technology, except that they make a one-time (sunk) investment I_e. The success rate of an innovator depends on his own investment I_e relative to the average spending $\overline{I}_e$ of others in his cohort, scaled by the ratio of new entrants to existing products. In particular, let

[2]At this stage, it would be easy to assume that the technology X of an incumbent evolves as a geometric Brownian motion. The cross-sectional distribution of technologies among initially identical firms, within each age cohort, would be lognormal, with a growing variance, and the overall distribution would be a mixture of lognormals. When the solution to the model is actually characterized in section 6, however, the argument relies on technologies across incumbents having a Pareto distribution. At that point the mixture of lognormals would be incompatible with the requirement of a Pareto distribution overall, and the variance term would have to be dropped.

$E(t)$ denote the flow of entrants at t, and define the entry rate $\epsilon(t) = E(t)/N_p(t)$. The success rate of an entrant who invests I_e is

$$\Lambda_{xe}(I_e/\overline{I}_e, \epsilon) = \frac{\phi_e}{\epsilon} \times \begin{cases} 0, & \text{if } I_e/\overline{I}_e < 1 - \varepsilon_x, \\ 1 - \left(1 - I_e/\overline{I}_e\right)/\varepsilon_x, & \text{if } I_e/\overline{I}_e \in [1 - \varepsilon_x, 1], \\ 1 + \varepsilon_x \left(I_e/\overline{I}_e - 1\right), & \text{if } I_e/\overline{I}_e > 1, \end{cases}$$

where $\phi_e > 0$ and where $\varepsilon_x > 0$ is small. Thus, $\Lambda_{xe}(., \epsilon)$ is kinked at $I_e/\overline{I}_e = 1$, reflecting a 'patent race' with intense competition among entrants, and scaling by $1/\epsilon$ reflects the reduced chances for success when the field is crowded. The value $V_{fe}(\cdot; \Lambda)$ of a product innovator with success rate Λ who is waiting for a technology at date τ, gross of the investment cost, satisfies the Bellman equation

$$[r(\tau) + \delta_x]\, V_{fe}(\tau; \Lambda) = \Lambda \left\{ \mathrm{E}_{F(\cdot,\tau)} \left[V^f(X, \tau) \right] - V_{fe}(\tau; \Lambda) \right\} + \frac{\partial V_{fe}(\tau; \Lambda)}{\partial \tau}, \quad \tau \geq 0.$$

Note that it does not depend on the investment date t.

An entrant takes $\overline{I}_e(t)$ and $\epsilon(t)$ as given, chooses I_e to solve

$$\max_{I_e} \left\{ V_{fe} \left[t; \Lambda_{xe}(I_e/\overline{I}_e, \epsilon) \right] - I_e \right\},$$

and is willing to enter if and only if the maximized value is nonnegative. Since Λ_{xe} diverges as $\epsilon \to 0$, in equilibrium there is positive entry at all dates, $E(t) > 0$. All entering firms choose the same investment level, and their common success rate is $\lambda_{xe}(t) = \phi_e/\epsilon(t)$. Free entry, together with the form of the function Λ_{xe}, imply that their common expenditure level is bid up to exhaust profits,

$$I_e(t) = V_{fe}(t; \lambda_{xe}(t)), \qquad \text{all } t,$$

and aggregate spending by entrants at t is $E(t)I_e(t)$.

B. Workers: investment in human capital

Workers invest to maximize the expected discounted value of their lifetime earnings. An individual who chooses to invest—a retooler—stops working, abandons his old skill and waits to acquire a new one. Let $V^w(H,t)$ denote the value of a worker with skill H at date t, and let $V_{wi}(t)$ denote the value of a retooler.

Success for retoolers is stochastic, arriving at rate $\lambda_{hi} > 0$, and conditional on success at date t, the individual gets a skill level drawn from the distribution $\Psi(\cdot,t)$ across current workers. Hence $V_{wi}(t)$ satisfies the Bellman equation

$$[r(t)+\delta_h]\,V_{wi}(t) = \lambda_{hi}\left\{\mathrm{E}_{\Psi(\cdot,t)}[V^w(H,t)] - V_{wi}(t)\right\} + V'_{wi}(t), \qquad \text{all } t,$$

where the term in braces is the expected gain in value conditional on success, and $\delta_h \geq 0$ is an exogenous exit rate.

Clearly $V^w(H,t)$ is nondecreasing in its first argument: higher human capital can only raise the worker's expected lifetime income. Hence at any date t, all individuals with skill below some threshold $H_m(t)$ become retoolers, while those with skill above the threshold continue working. It follows that at date t, the value of a worker with skill $H \leq H_m(t)$ is $V_{wi}(t)$, and the irreversibility of investment implies that $H_m(t)$ is nondecreasing.

While an individual works, his human capital H grows (or declines) at a constant rate μ_h, which can be interpreted as on-the-job learning. Hence the value $V^w(H,t)$ for a worker with skill $H > H_m(t)$, satisfies the Bellman equation

$$[r(t)+\delta_h]\,V^w(H,t) = W(H,t) + \mu_h H V^w_H(H,t) + V^w_t(H,t), \qquad \text{all } t.$$

As for firms, value matching and smooth pasting hold at the threshold $H_m(t)$, so

$$\begin{aligned} V^w[H_m(t),t] &= V_{wi}(t), \\ V^w_H[H_m(t),t] &= 0, \qquad\qquad \text{all } t. \end{aligned}$$

New entrants into the workforce—trainees, have an investment technology like the one for retoolers, except that their hazard rate for success, call it λ_{he}, may be different. They pay no costs, so their value function $V_{we}(t)$ satisfies the Bellman equation

$$[r(t)+\delta_h]\,V_{we}(t) = \lambda_{he}\left\{\mathrm{E}_{\Psi(\cdot,t)}[V^w(H,t)] - V_{wi}(t)\right\} + V'_{we}(t), \qquad \text{all } t.$$

Trainees arrive at the rate $(\upsilon + \delta_h)\, N(t)$.

C. The evolution of technology and skill

Next consider the evolution of the group sizes (N_p, N_i, N_e) and (L_w, L_i, L_e) and the distribution functions F and Ψ.

The number of producers $N_p(t)$ grows because of success by innovators of both types, and declines because of exit and decisions to switch to process innovation. The producers that switch to innovating around date t are those with technologies $X(t)$ that are close enough to the threshold $X_m(t)$ so that growth in that threshold overtakes them. Since technologies for producers grow at the rate μ_x, there is a positive level of switching at date t if and only if

$$X_m'(t) - \mu_x X_m(t) > 0, \qquad \text{all } t, \tag{19}$$

and

$$\begin{aligned} N_p'(t) \;&=\; \lambda_{xi} N_i(t) + \lambda_{xe} N_e(t) - \delta_x N_p(t) \\ &\quad - \max\left\{0,\ [X_m'(t) - \mu_x X_m(t)]\, f[X_m(t), t] N_p(t)\right\}, \qquad \text{all } t. \end{aligned}$$

The number of process innovators N_i grows because producers switch to innovating, while the number of product innovators N_e grows because new entrants join. Each declines because of exit and success, so

$$\begin{aligned} N_i'(t) \;&=\; \max\left\{0,\ [X_m'(t) - \mu_x X_m(t)]\, f(X_m(t), t) N_p(t)\right\} - (\delta_x + \lambda_{xi})\, N_i(t), \\ N_e'(t) \;&=\; E(t) - (\delta_x + \lambda_{xe})\, N_e(t), \qquad \text{all } t. \end{aligned}$$

The distribution function F for technology among producers evolves because their technologies grow at rate μ_x, they exit at the rate δ_x, innovators of both types succeed, and firms at the threshold $X_m(t)$ switch to process innovations. As shown in the

Appendix, $F(X,t)$ satisfies

$$-F_t(X,t) = f(X,t)\mu_x X + [1 - F(X,t)]\left[-\frac{N_p'(t)}{N_p(t)} - \delta_x + \lambda_{xi}\frac{N_i(t)}{N_p(t)} + \lambda_{xe}\frac{N_e(t)}{N_p(t)}\right],$$
$$\text{all } X \geq X_m(t), \quad t \geq 0.$$

The dynamics for the labor force are analogous. Workers who switch to retooling around date t are those whose human capital $H(t)$ falls below the (moving) threshold $H_m(t)$, despite growth at the rate μ_h. Hence workers are switching at date t if and only if

$$H_m'(t) - \mu_h H_m(t) > 0, \qquad \text{all } t, \tag{20}$$

and

$$\begin{aligned}
L_w'(t) &= \lambda_{hi}L_i(t) + \lambda_{he}L_e(t) - \delta_h L_w(t) \\
&\quad - \max\{0,\ [H_m'(t) - \mu_h H_m(t)]\,\psi[H_m(t),t]L_w(t)\}, \\
L_i'(t) &= \max\{0,\ [H_m'(t) - \mu_h H_m(t)\psi[H_m(t),t]L_w(t)]\} - (\delta_h + \lambda_{hi})\,L_i(t), \\
L_e'(t) &= (\upsilon + \delta_h)\,L(t) - (\delta_h + \lambda_{he})\,L_e(t), \qquad \text{all } t,
\end{aligned}$$

and $\Psi(H,t)$ satisfies

$$-\Psi_t(H,t) = \psi(H,t)\mu_h H + [1 - \Psi(H,t)]\left[-\frac{L_w'(t)}{L_w(t)} - \delta_h + \lambda_{hi}\frac{L_i(t)}{L_w(t)} + \lambda_{he}\frac{L_e(t)}{L_w(t)}\right],$$
$$\text{all } H \geq H_m(t), \quad t \geq 0.$$

D. Consumption

Individuals are organized into a continuum of identical, infinitely lived households of total mass one, where each dynastic household comprises a representative cross-section of the population. New entrants to the workforce arrive at the fixed rate $\delta_h + \upsilon$, so each household grows in size at the constant rate $\upsilon \geq 0$, and total population at date t is $L(t) = L_0 e^{\upsilon t}$.

Members of the household pool their earnings and they own the profit streams from firms. The investment decisions of firms and workers, both incumbents and entrants, maximize, respectively, the expected discounted value of net profits and wages. Hence there are no further investment decisions at the household level. Since there is no aggregate uncertainty, the household faces no consumption risk.

The household's income consists of the wages of its workers plus the profits from its portfolio, which sum to output of the final good,

$$Y_F(t) = L_w(t)\mathrm{E}_{\Psi(\cdot,t)}\left[W(H,t)\right] + N_p(t)\mathrm{E}_{F(\cdot,t)}\left[\Pi(X,t)\right], \qquad \text{all } t.$$

That income is used for consumption and to finance the investment (entry) costs of new firms. Hence the household's net income at date t is $Y_F(t) - E(t)I_e(t)$.

All household members share equally in consumption, and the household has the constant-elasticity preferences

$$U = \int_0^\infty L_0 e^{\upsilon t} e^{-\hat{r}t} \frac{1}{1-\theta} c(t)^{1-\theta} dt,$$

where $\hat{r} > 0$ is the rate of pure time preference, $1/\theta > 0$ is the elasticity of intertemporal substitution, and $c(t)$ is per capita consumption.

The household chooses $c(t)$, $t \geq 0$, to maximize utility, subject to the budget constraint,

$$\int_0^\infty e^{-R(t)} \left\{ L_0 e^{\upsilon t} c(t) - \left[Y_F(t) - E(t)I_e(t)\right] \right\} dt \leq 0,$$

where

$$R(t) = \int_0^t r(s) ds, \qquad \text{all } t.$$

The condition for an optimum implies that per capita consumption grows at the rate

$$\frac{c'(t)}{c(t)} = \frac{1}{\theta}\left[r(t) - \hat{r}\right], \qquad \text{all } t,$$

with $c(0)$ determined by budget balance.

Final output is used for consumption and for the investment costs of entering firms. Hence market clearing for goods requires

$$Y_F(t) = L_0 e^{\upsilon t} c(t) + E(t) I_e(t), \qquad \text{all } t.$$

5. COMPETITIVE EQUILIBRIA, BGPS

This section provides formal definitions of a competitive equilibrium and a BGP. We start with the definition of a (dynamic) economy.

DEFINITION: An *economy* $\mathcal{E}$ is described by

i. parameters $(\rho, \chi, \omega, \eta, \theta, \hat{r}, \upsilon)$, with $\rho > 1$, $\chi \in (0, 1/\rho]$, $\omega \in (0,1)$, $\eta \in (0,1)$, $\theta > 0$, $\hat{r} > 0$, $\upsilon \geq 0$;

ii. parameters $\delta_j \geq 0, \lambda_{ji} > 0$ and μ_j, for $j = h, x$, and $\lambda_{he} > 0$;

iii. parameters $\phi_e > 0$ and $\varepsilon_x > 0$ for the function Λ_{xe};

iv. initial conditions $N_{p0}, N_{i0}, N_{e0} > 0$, $L_{w0}, L_{i0}, L_{e0} > 0$;

v. initial distribution functions $F_0(X)$ with continuous density $f_0(X)$ and lower bound X_{m0} on its support, and $\Psi_0(H)$ with continuous density $\psi_0(H)$ and lower bound $H_{m0} \geq 0$ on its support.

A. Competitive equilibrium

The definition of a competitive equilibrium is standard.

DEFINITION: A *competitive equilibrium* of an economy $\mathcal{E}$ consists of the following, for all $t \geq 0$:

a. the numbers of producers, process innovators, product innovators, workers, retoolers, and trainees, $[N_p(t), N_i(t), N_e(t), L_w(t), L_i(t), L_e(t)]$; and the inflow rate $E(t)$ of product innovators;

b. distribution functions $F(X;t), \Psi(H;t)$;

c. prices and allocations $[W(H;t), P(X,t), H^*(X;t), Y(X,t), \mathfrak{L}(X,t), \Pi(X,t), Y_F(t)]$;

d. value functions $\left[V^f(X,t),\ V_{fi}(t)\right]$ for producers and process innovators and an investment threshold $X_m(t)$; and a value function $V_{fe}(t;\Lambda)$, investment level $I_e(t)$, success rate, $\lambda_{xe}(t)$ for product innovators;

e. value functions $[V^w(H;t), V_{wi}(t), V_{we}(t)]$, for each workers, retooler, and trainees, and an investment threshold $H_m(t)$ for retoolers;

f. per capita consumption $c(t)$, and the interest rate $r(t)$;

such that for all $t \geq 0$,

i. $[W, P, H^*, Y, \mathfrak{L}, \Pi, Y_F]$, is a production equilibrium, given $[N_p, L_w, F, \Psi]$;

ii. $\left[V^f, X_m\right]$ solve the investment problem of producers, given $[r, \Pi, V_{fi}]$; V_{fi} and V_{fe} are consistent with $\left[r, V^f, F\right]$; I_e satisfies the optimization and entry conditions; and the success rate for product innovators is $\lambda_{xe} = \Lambda_{xe}(1, E/N_p)$;

iii. $[V^w, H_m]$ solve the investment problem of workers, given $[r, W, V_{wi}]$; and $[V_{wi}, V_{we}]$ are consistent with $[r, V^w, \Psi]$;

iii. $[N_p, N_i, N_e, F]$ are consistent with $[X_m, E]$, and the initial conditions $[N_{p0}, N_{i0}, N_{e0}, F_0]$; and $X_m(0) = X_{m0}$;

v. $[L_w, L_i, L_e, \Psi]$ are consistent with H_m and the initial conditions $[L_{w0}, L_{i0}, L_{e0}, \Psi_0]$; and $H_m(0) = H_{m0}$;

vi. c solves the consumption/savings problem of households, given $[r, Y_F - EI_e]$; and

vii. the goods market clears.

B. Balanced growth

The rest of the analysis focuses on balanced growth paths, competitive equilibria with the property that quantities grow at constant rates, and the normalized distributions of technology and skill are time invariant.

Let $Q(t) \equiv \mathrm{E}_{F(\cdot,t)}(X)$, $t \geq 0$, denote average technology at date t. On a BGP Q grows at a constant rate, call it g, and the distributions of relative technology

$x = X/Q(t)$ and relative human capital $h = H/Q(t)$ are constant. By assumption total population L grows at the fixed rate υ. On a BGP the number of firms N also grows at a constant rate, call it n, and the shares of firms and individuals in each category, $[N_p/N, N_i/N, N_e/N]$ and $[L_w/L, L_i/L, L_e/L]$ are constant. The growth rates g and n are endogenous.

It follows from Lemma 2 that on a BGP the labor allocation in terms of relative technology and relative skill is time invariant. The growth rates for wages, prices, output levels, and so on are then described by Lemma 3, where Ω, defined in (8), measures the impact of variety growth. In particular, average product price grows at rate Ωn, average output per firm at rate $g + \upsilon - n$, and average employment per firm at rate $\upsilon - n$. Output per capita (wages), aggregate output, and profits per firm, grow at rates

$$g_w = g + \Omega n, \qquad g_Y = g_w + \upsilon, \qquad g_\pi = g_Y - n. \tag{21}$$

Total investment costs $Ei_e \mathrm{E}_F[\Pi]$, also grow at rate g_Y. If $\Omega > 1$, then love of variety is so strong that an increase in the number of producers actually raises the profits of each incumbent. In a dynamic model with free entry, this feature poses obvious problems. In the rest of the analysis we will assume that $\rho \geq 2$, which implies $\Omega < 1$.

These observations lead to the following definition.

DEFINITION: A competitive equilibrium for $\mathcal{E}$ is a *balanced growth path* (BGP) if for some $g > 0$ and n, with g_Y, g_w and g_π as in (21), the equilibrium has the property that for all $t \geq 0$:

a. the numbers of firms and individuals satisfy

$$\begin{aligned} N_p(t) &= e^{nt} N_{p0}, \qquad N_i(t) = e^{nt} N_{i0}, \qquad N_e(t) = e^{nt} N_{e0}, \\ L_w(t) &= e^{\upsilon t} L_{w0}, \qquad L_i(t) = e^{\upsilon t} L_{i0} \qquad L_e(t) = e^{\upsilon t} L_{e0}; \end{aligned}$$

and for some $E_0 > 0$, the flow of entrants satisfies

$$E(t) = e^{nt} E_0;$$

b. for $Q_0 \equiv \mathrm{E}_{F_0}[X]$, average technology satisfies

$$Q(t) \equiv \mathrm{E}_{F(\cdot,t)}[X] = e^{gt}Q_0;$$

and for some $\left[\hat{F}(x), \hat{\Psi}(h)\right]$, the distribution functions satisfy

$$\begin{aligned} F(X,t) &= \hat{F}(X/Q(t)), && \text{all } X, \\ \Psi(H,t) &= \hat{\Psi}(H/Q(t)), && \text{all } H; \end{aligned}$$

c. for some $[w, p, h^*, y, \ell, \pi, y_F]$, the production equilibria satisfy

$$\begin{aligned} W(H;t) &= e^{g_w t}Q_0 w(H/Q(t)), && \text{all } H; \\ P(X,t) &= e^{\Omega n t}Q_0 p(X/Q(t)), \\ H^*(X;t) &= e^{gt}Q_0 h^*(X/Q(t)), \\ Y(X,t) &= e^{(g+\upsilon-n)t}Q_0 y(X/Q(t)), \\ \mathfrak{L}(X,t) &= e^{(\upsilon-n)t}\ell(X/Q(t)), \\ \Pi(X,t) &= e^{g_\pi t}Q_0 \pi(X/Q(t)), && \text{all } X; \\ Y_F(t) &= e^{g_Y t}Q_0 y_F; \end{aligned}$$

d. for some $[v_{fp}(x), v_{fi}, x_m]$, the value function and optimal policy for producers and the value for process innovators satisfy

$$\begin{aligned} V^f(X,t) &= e^{g_\pi t}Q_0 v_{fp}(X/Q(t)), && \text{all } X, \\ X_m(t) &= e^{gt}Q_0 x_m, \\ V_{fi}(t) &= e^{g_\pi t}Q_0 v_{fi}; \end{aligned}$$

for some $[v_{fe}(\lambda), i_e]$, the value function and investment level for product innovators satisfy

$$\begin{aligned} V_{fe}(\lambda;t) &= e^{g_\pi t}Q_0 v_{fe}(\lambda), \\ I_e(t) &= e^{g_\pi t}Q_0 i_e, \end{aligned}$$

and their success rate

$$\lambda_{xe}(t) = \lambda_{xe0} = \Lambda_{xe}\left(1, \frac{E_0}{N_{p0}}\right)$$

is constant;

e. for some $[v_w(h), v_{wi}, v_{we}, h_m]$, the values and optimal policies for individuals satisfy

$$\begin{aligned}
V^w(H;t) &= e^{g_w t} Q_0 v_w(H/Q(t)), \qquad \text{all } H, \\
V_{wi}(t) &= e^{g_w t} Q_0 v_{wi}, \\
V_{we}(t) &= e^{g_w t} Q_0 v_{we}, \\
H_m(t) &= e^{g t} Q_0 h_m;
\end{aligned}$$

f. for some $c_0 > 0$ aggregate consumption satisfies

$$C(t) = e^{g_Y t} Q_0 L_0 c_0;$$

and the interest rate satisfies

$$r(t) = r \equiv \hat{r} + \theta g_w.$$

BGPs arise—if at all—only for initial conditions $[N_{p0}, N_{i0}, N_{e0}, L_{w0}, L_{i0}, L_{e0}, F_0(X), \Psi_0(H)]$ that satisfy certain restrictions. The rest of the analysis focusses on a class of economies for which BGPs exist, and studies the determinants of the growth rates g and n of TFP and variety

6. CONDITIONS FOR BALANCED GROWTH

In this section we will show that if an economy $\mathcal{E}$ has initial distribution functions F_0 and Ψ_0 that are Pareto, with shape and location parameters that satisfy the requirements of Proposition 4, then the normalized value functions $v_{fp}(x)$ and $v_w(h)$

for producers and workers inherit the isoelastic forms of the normalized profit and wage functions, and simple closed form solutions can be found. Moreover, the growth rates g and n, as well as other features of the BGP, can be solved for explicitly. The arguments are summarized in Proposition 5, which provides sufficient conditions for existence and uniqueness of a BGP.

A. Production equilibrium

Suppose the initial distributions F_0 and Ψ_0 are Pareto, with parameters (α_x, X_{m0}) and (α_h, H_{m0}). Assume (9) holds, define ε, ζ and a_h by (10)-(12), and assume that $H_{m0} = a_h X_{m0}$. Average technology under the initial distribution is

$$Q_0 \equiv \mathrm{E}_{F_0}[X] = \frac{\alpha_x}{\alpha_x - 1} X_{m0}. \tag{22}$$

Use Q_0 to define the normalized distribution functions

$$\begin{aligned} \hat{F}(X/Q_0) &\equiv F_0(X), \qquad \text{all } X \geq X_m, \\ \hat{\Psi}(H/Q_0) &\equiv \Psi_0(H), \qquad \text{all } H \geq H_m. \end{aligned} \tag{23}$$

By construction $\mathrm{E}_{\hat{F}}(x) = 1$, and the location parameters for $\hat{F}$ and $\hat{\Psi}$ are

$$x_m = \frac{X_{m0}}{Q_0} = \frac{\alpha_x - 1}{\alpha_x}, \qquad h_m \equiv \frac{H_{m0}}{Q_0} = a_h x_m. \tag{24}$$

Hence the hypotheses of Proposition 4 hold for $\hat{F}, \hat{\Psi}$, and given N_{p0}, L_{w0}, (14)-(17) describe the production equilibrium $[w, p, h^*, y, \ell, \pi, y_F]$.

B. Firms, investment in technology

Next we will characterize the normalized value functions $[v_{fp}(x), v_{fi}, v_{fe0}]$ and investment cost i_{e0} for entrants as functions of (g, n), and derive one additional equation relating (g, n). Recall that a BGP requires positive process innovation, which in turn requires $g > \mu_x$, so that the investment threshold grows faster than producers' technologies. We will assume that this condition holds.

Consider the investment decision and value of a producer. As shown in the Appendix, if Π and V^f have the forms required for a BGP, and $\pi(x)$ has the isoelastic form in Proposition 4, then the normalized value function $v_{fp}(x)$ for a producer satisfies

$$(r - g_\pi + \delta_x)\, v_{fp}(x) = \pi_2 \left(x/x_m\right)^{1-\zeta} - (g - \mu_x)\, x v'_{fp}(x), \qquad x \geq x_m,$$

where π_2 is as defined in (18), with $L_w = L_{w0}$ and $N_p = N_{p0}$. This equation is a first-order ODE. Value matching provides a boundary condition, yielding the solution

$$v_{fp}(x) = \pi_2 B_x \left(x/x_m\right)^{1-\zeta} + \left(v_{fi} - \pi_2 B_x\right) \left(x/x_m\right)^{R_x}, \qquad x \geq x_m, \tag{25}$$

where the constant $B_x > 0$ and characteristic root $R_x < 0$ are known functions of the parameters. The first term in (25) is the value of a producer who operates forever, never investing. The second term represents the additional value from the option to invest in process innovation. The smooth pasting condition

$$\begin{aligned} v_{fi} &= \pi_2 B_x \left(1 - \frac{1-\zeta}{R_x}\right) \\ &= \frac{\pi_2}{r - g_\pi + \delta_x}, \end{aligned} \tag{26}$$

determines v_{fi}, the normalized value of a process innovator, from the optimal choice of the investment threshold by a producer switching to process innovation.

But from the perspective of a firm that has already switched and is waiting for a new technology, its value v_{fi} is the expected discounted value across current producers, adjusted for exit, growth, and waiting time. That is, on a BGP v_{fi} also satisfies

$$(r - g_\pi + \delta_x)\, v_{fi} = \lambda_{xi} \left\{ \mathrm{E}_{\hat{F}} \left[v_{fp}(x)\right] - v_{fi} \right\}.$$

To simplify this condition, substitute for $\mathrm{E}_{\hat{F}}\left[v_{fp}(x)\right]$ and v_{fi} from (25) and (26) and factor out $\pi_2 B_x$ to get

$$r - g_\pi + \delta_x = \frac{-R_x\,(1-\zeta)\,\lambda_{xi}}{(\alpha_x - 1 + \zeta)\,(\alpha_x - R_x)}. \tag{27}$$

Since r, g_π and R_x involve g and n, while all of the other parameters in this expression are exogenous, (27) is an equation in the pair (g, n), a restriction that the Bellman equation for process innovators places on the ratio $\mathrm{E}_{\hat{F}}\left[v_{fp}(x)\right]/v_{fi}$.

For any success rate λ, on a BGP the normalized value $v_{fe}(\lambda)$ of a product innovator, gross of the sunk cost, satisfies

$$(r - g_\pi + \delta_x + \lambda)\, v_{fe}(\lambda) = \lambda \mathrm{E}_{\hat{F}}\left[v_{fp}(x)\right], \tag{28}$$

which determines the function v_{fe}. The flow of new entrants is determined by the rate of variety growth, $E = (n + \delta_x)\, N$, so the entry rate is $\epsilon = (n + \delta_x)\, N/N_p$. As shown below, the ratio N/N_p is constant on a BGP, so the entry rate $\epsilon > 0$ is also constant. In any equilibrium $i_e/\bar{i}_e = 1$, so the success rate for product innovators is also constant,

$$\lambda_{xe0} = \Lambda_{xe}(1, \epsilon) = \phi_e/\epsilon > 0.$$

From (28), the value of an entrant is then

$$v_{fe0} = \frac{\lambda_{xe0}}{r - g_\pi + \delta_x + \lambda_{xe0}} \mathrm{E}_{\hat{F}}\left[v_{fp}(x)\right], \tag{29}$$

and expected profits just cover entry costs if

$$i_{e0} = v_{fe0}. \tag{30}$$

We must also check the incentive to invest. Given the function $v_{fe}(\cdot)$ in (28), the entry rate ϵ, and the average investment $\bar{i}_e$ of others in its cohort, the normalized problem of a product innovator is

$$\max_{i_e} \left\{ v_{fe}[\Lambda_{xe}(i_e/\bar{i}_e, \epsilon)] - i_e \right\},$$

As shown in the Appendix, the optimum is at $i_e/\bar{i}_e = 1$ if and only if

$$\varepsilon_x \le \frac{r - g_\pi + \delta_x}{(r - g_\pi + \delta_x + \lambda_{xe0})^2} \frac{\phi_e}{\epsilon} \le \frac{1}{\varepsilon_x}, \tag{31}$$

which holds for $\varepsilon_x > 0$ sufficiently small.

C. Workers, investment in skill

The argument for the labor force is analogous except that the entry rate of trainees is exogenous, as is their success rate λ_{he}. Hence the normalized value function v_w for a worker satisfies

$$(r - g_w + \delta_h)\, v_w(h) = w_2 \left(h/h_m\right)^{1-\varepsilon} + (\mu_h - g)\, h v_w'(h), \qquad h \geq h_m.$$

Suppose $g > \mu_h$, so there is positive retooling, as required on a BGP. Using value matching for the boundary condition, the solution to this ODE is

$$v_w(h) = B_h w_2 \left(h/h_m\right)^{1-\varepsilon} + (v_{wi} - B_h w_2) \left(h/h_m\right)^{R_h}, \qquad h \geq h_m, \tag{32}$$

where the constant $B_h > 0$ and characteristic root $R_h < 0$ are known functions of the parameters. The first term in (32) is the value of a worker who never invests, and the second represents the additional value from the option to retool. The value of a retooler v_{wi} is determined by the smooth pasting condition

$$\begin{aligned} v_{wi} &= B_h w_2 \left(1 - \frac{1-\varepsilon}{R_h}\right) \\ &= \frac{w_2}{r - g_w + \delta_h}. \end{aligned} \tag{33}$$

The value of a retooler also satisfies the Bellman equation

$$(r - g_w + \delta_h)\, v_{wi} = \lambda_{hi} \left\{ \mathrm{E}_{\Psi} \left[v_w(h) \right] - v_{wi} \right\}.$$

Using (32) and (33) to substitute for $\mathrm{E}_{\hat{\Psi}}\left[v_w\right]$ and v_{wi}, and factoring out $w_2 B_h$, gives

$$r - g_w + \delta_h = \frac{-R_h (1-\varepsilon)\, \lambda_{hi}}{(\alpha_h - 1 + \varepsilon)(\alpha_h - R_h)}, \tag{34}$$

a second equation in the pair (g, n), a restriction that the Bellman equation for retoolers places on the ratio $\mathrm{E}_{\Psi}\left[v_w(h)\right]/v_{wi}$.

The value v_{we} of a trainee is determined by

$$(r - g_\pi + \delta_h + \lambda_{he})\, v_{we} = \lambda_{he} \mathrm{E}_{\hat{\Psi}} \left[v_w(h) \right]. \tag{35}$$

D. Flows of firms and workers, the evolution of technology and skill

On a BGP the number of firms grows at a constant rate n, and the shares of firms engaged in production and the two kinds of innovation are constant. Hence the flow of new entrants at any date is

$$E = (n + \delta_x)\, N, \tag{36}$$

and the entry rate is constant,

$$\epsilon = \frac{E}{N_p} = (n + \delta_x)\, \frac{N}{N_p}.$$

The laws of motion for N_p, N_i and N_e determine the ratios of process and product innovators to producers,

$$\begin{aligned} \frac{N_i}{N_p} &= \frac{\alpha_x (g - \mu_x)}{n + \delta_x + \lambda_{xi}}, \\ \frac{N_e}{N_p} &= \frac{n + \delta_x}{\lambda_{xe}} \left[1 + \frac{\alpha_x (g - \mu_x)}{n + \delta_x + \lambda_{xi}} \right]. \end{aligned} \tag{37}$$

The success rate for product innovators is

$$\lambda_{xe0} = \Lambda_{xe}(1, \epsilon) = \frac{\phi_e}{n + \delta_x} \frac{N_p}{N}, \tag{38}$$

where N_p/N is determined by (37).

The shares of the labor force in each group are also constant, and the laws of motion for L_w, L_i, and L_e imply that the ratios of retoolers and trainees to workers are

$$\begin{aligned} \frac{L_i}{L_w} &= \frac{\alpha_h (g - \mu_h)}{\upsilon + \delta_h + \lambda_{hi}}, \\ \frac{L_e}{L_w} &= \frac{\upsilon + \delta_h}{\lambda_{he}} \left[1 + \frac{\alpha_h (g - \mu_h)}{\upsilon + \delta_h + \lambda_{hi}} \right]. \end{aligned} \tag{39}$$

It is easy to check that if X_m and H_m grow at rate g, as required on a BGP, then the distribution functions $F(\cdot, t)$ and $\Psi(\cdot, t)$ evolve as required.

E. Consumption, the interest rate

On a BGP per capita consumption grows at the rate g_w, so the interest rate is

$$r = \hat{r} + \theta g_w. \tag{40}$$

Aggregate income grows at the rate g_Y, so its present discounted value is finite if and only if $r > g_Y$, or

$$\hat{r} > g_Y - \theta g_w = \upsilon + (1-\theta)(g + \Omega n).$$

Using (30) for i_{e0}, market clearing for goods determines c_0,

$$y_F = L_0 c_0 + E_0 i_{e0}. \tag{41}$$

F. Existence of BGPs

The growth rates (g, n) are determined by (27) and (34). Substituting for g_w, g_π, r, and the roots R_x and R_h, gives a pair of linear equations in the two unknowns,

$$\begin{aligned} g &= \frac{1}{\xi_h}\left(Z_h + A_h\lambda_{hi} - W_n n\right), \\ n &= \upsilon + (\xi_h - \xi_x)\, g + (Z_x - Z_h) + (A_x\lambda_{xi} - A_h\lambda_{hi}), \end{aligned} \tag{42}$$

where

$$\begin{aligned} \xi_h &\equiv \alpha_h - 1 + \theta > 0, & \xi_x &\equiv \alpha_x - 1 + \theta > 0, \\ Z_h &\equiv \alpha_h\mu_h - \delta_h - \hat{r}, & Z_x &\equiv \alpha_x\mu_x - \delta_x - \hat{r}, \\ A_h &\equiv \frac{1-\varepsilon}{\alpha_h - 1 + \varepsilon} > 0, & A_x &\equiv \frac{1-\zeta}{\alpha_x - 1 + \zeta} > 0, \\ W_n &\equiv (\theta - 1)\,\Omega. \end{aligned}$$

Propositions 5 and 6 both use the assumption $\rho \geq 2$, which implies $\Omega < 1$, so love of variety is not too strong. One additional joint restriction on ρ, α_h is also imposed if $\theta < 1$. Although stronger than required for existence, it will be needed for the comparative statics results.

Proposition 5: Let $\mathcal{E}$ be an economy with:

a. $\rho \geq 2$, and $\rho > \alpha_h / (\alpha_h - 1)$ if $\theta < 1$;

b. initial distributions F_0, Ψ_0 that are Pareto, with shape and location parameters (α_x, X_{m0}), (α_h, H_{m0}) satisfying the hypotheses of Proposition 4.

Define ε and ζ by (10) and (11). Then the pair of equations in (42) has a unique solution (g, n), and there are unique $[Q_0, \hat{F}, \hat{\Psi}]$, $[w, p, h^*, y, \ell, \pi, y_F]$ satisfying conditions (b)-(c) for a BGP; $[v_{fp}(x), x_m, v_{fi}, v_{fe}(\lambda), \lambda_{xe0}, i_{e0}]$ satisfying (d); $[v_w(h), h_m, v_{wi}, v_{we}]$ satisfying (e); and $[c_0, r]$ satisfying (f).

If in addition:

c. the initial ratios $[N_{i0}/N_{p0}, N_{e0}/N_{p0}]$ and $[L_{i0}/L_{w0}, L_{e0}/L_{w0}]$ satisfy (37) and (39); and

d. $g > \mu_x$, $g > \mu_h$, $r > g_Y$, ε_x satisfies (31), and $c_0 > 0$,

then $\mathcal{E}$ has a unique competitive equilibrium that is a BGP.

7. GROWTH RATES ON THE BGP

In this section we will examine the effect of various parameters on the growth rates g and n. We will also look at subsidy policies that increase growth and provide a roughly calibrated numerical example. In the example, subsidies to retooling have a powerful effect on growth.

A. Growth rates

If preferences are logarithmic, $\theta = 1$, or if $\Omega = 0$, so variety is not valued, then $W_n = 0$ and the first equation in (42), by itself, determines g. Thus g is a weighted sum of the skill parameters $\mu_h, \lambda_{hi}, \delta_h$, and the rate of time preference $\hat{r}$, and the technology parameters $\mu_x, \lambda_{xi}, \delta_x$, do not enter. Faster on-the-job skill growth μ_h raises g, as does a higher success rate λ_{hi} for retoolers. A higher exit rate δ_h or discount rate $\hat{r}$ reduces g. If $\theta > 1$ and $\Omega > 0$, then TFP growth has a downward adjustment for variety growth.

From the second equation in (42), the rate of variety growth n increases one-for-one with population growth, with adjustments for differences in the parameters governing technology and skill acquisition. As noted in section 4, data on the distribution of firm or establishment size by employment suggests $\alpha_x > \alpha_h$. In this case $\xi_h - \xi_x < 0$, so variety growth is adjusted downward for TFP growth.

Proposition 6 has a precise summary of the comparative statics results.

Proposition 6: Let $\mathcal{E}$ be as in Proposition 5. Then

a. an increase in μ_h, λ_{hi} or a decrease in δ_h raises g and reduces n;

b. an increase in $\upsilon, \mu_x, \lambda_{xi}$ or a decrease in δ_x raises n, and

—raises g if $(\theta - 1)\,\Omega < 0$,

—has no effect on g if $(\theta - 1)\,\Omega = 0$, and

—reduces g if $(\theta - 1)\,\Omega > 0$;

c. a decrease in $\hat{r}$ raises g if $(\theta - 1)\,\Omega \leq 0$, and has otherwise ambiguous effects.

The initial population size and number firms, L_0 and N_0, do not affect the growth rates, although they do affect the levels for wages, profits and the distribution of employment across technologies. In particular, since N_{p0} and L_{w0} are directly proportional to N_0 and L_0, it follows from the definitions of w_2, ℓ_2, and π_2, that wages are proportional to N_0^{Ω}, employment is proportional to L_0. and profits per firm are proportional to $L_0 N_0^{\Omega-1}$. Similarly, the success rates $\lambda_{he}, \lambda_{xe}$ for entrants acquiring initial skill/productivity affect the population shares but not the growth rate.

B. Policies to increase growth

Investments by incumbent producers and workers have positive external effects, since they improve the pools from which later investors—both incumbents and entrants—

draw their technologies and skills. Hence investment is below the efficient level, as in Perla and Tonetti (2014, Propositions 3 and 4). Subsidies to investment are obvious policies to overcome this inefficiency.

A complete analysis of optimal policies, which would require characterizing the transition path between old and new BGPs, is beyond the scope of this paper. But it is easy to assess the effect of subsidies on long-run growth rates.

Consider subsidies to process innovators and retoolers at constant rates σ_x and σ_h, scaled by the wage/profit flow of the the marginal producer/worker. Under such a policy the normalized Bellman equations for process innovators and retoolers are

$$
\begin{aligned}
(r - g_\pi + \delta_x) \, v_{fi} &= \sigma_x \pi_2 + \lambda_{xi} \left\{ \mathrm{E}_F[v^f(x)] - v_{fi} \right\}, \\
(r - g_w + \delta_h) \, v_{wi} &= \sigma_h w_2 + \lambda_{hi} \left\{ \mathrm{E}_\Psi[v^w(x)] - v_{wi} \right\},
\end{aligned}
$$

and the constants A_x and A_h in (42) become

$$
\hat{A}_x = \frac{1}{1-\sigma_x} A_x, \qquad \hat{A}_h = \frac{1}{1-\sigma_h} A_h.
$$

Thus, subsidies increase the coefficients on the hazard rates λ_{xi} and λ_{hi}. The next result is then immediate.

Corollary 1: Under the hypotheses of Proposition 6, a subsidy $\sigma_h > 0$ to retoolers raises g and reduces n, while a subsidy $\sigma_x > 0$ to process innovators increases n and increases, leaves unchanged, or decreases g as $(\theta - 1) \, \Omega < 0$, $= 0$, or > 0.

C. An example

We conclude with a roughly calibrated numerical example, using data for establishments to calibrate parameters for products (technologies). In particular,

$$\begin{aligned} &\alpha_x = 7, &&\alpha_h = 1.2, &&\rho = 8, \\ &\theta = 2, &&\hat{r} = 0.04, &&\upsilon = 0.0107, \\ &\delta_x = 0.10, &&\mu_x = 0.008, &&\lambda_{xe} = 0.33, \\ &\delta_h = 0.03, &&\mu_h = 0.008, &&\lambda_{he} = 0.20, \end{aligned}$$

$$\begin{aligned} &\chi = [1/8,\ 1/16,\ 1/1000], &&\lambda_{hi} = [0.576,\ 0.572,\ 0.567], \\ &\Omega = [0,\ 0.0714,\ 0.1417], &&\lambda_{xi} = [0.0283,\ 0.0276,\ 0.0270], \end{aligned}$$

which implies

$$\begin{aligned} &\alpha_R = 1.18, &&\alpha_\ell = 1.21, &&\alpha_w = 8, \\ &g_w = 0.00995, &&n = 0.00741, &&r = 0.0599, \\ &g_Y = 0.02065, &&g_\pi = 0.0132, &&g = [0.00995,\ 0.00942,\ 0.00890]. \end{aligned}$$

The parameters α_x, α_h, ρ, are chosen to roughly target the tail parameters $\alpha_R, \alpha_\ell, \alpha_w$, for the distributions of revenue, employment and wages. Zipf's law for establishment size, measured by either revenue or employment, suggests values not too far above unity, while the distribution for wages is approximately log-normal, suggesting a large value for α_w.

The values for $\theta, \hat{r}$ are standard in the macro literature, υ is set at the average rate of employment growth over the period 1988-2015, δ_x is the empirical exit rate for firms, and δ_h implies an average working lifetime of 33 years. The evidence for on-the-job-training/learning-by-doing in Hansen and Imrohoroglu (2009) suggests a high value for μ_h. Here it is set near the upper bound imposed by Proposition 5, and μ_x is set at the same level.

Since there is little evidence on returns-to-variety, several values—spanning the entire feasible range—are used for χ. For each value, the success rates for incumbent investors λ_{hi} and λ_{xi} are chosen to target wage growth g_w and growth n in the number of establishments, which are set at their historical averages for the period 1988-2015. These determine the interest rate r, aggregate output growth g_Y, growth in profits per establishment g_π, and TFP growth g. Of these, only g varies with χ. The success rates λ_{he} and λ_{xe} do not affect g or n. The growth rates satisfy the inequalities required by Proposition 5, $c_0 > 0$, ε_x can be chosen so (31) holds, and the ratios N_{i0}/N_{p0}, etc. can be chosen so condition (c) holds.

The model also has predictions for wage growth for individuals and growth in revenue and employment for establishments. For either type of agent growth has two components, continuous growth while working/producing and jumps from successful investment. Since the jumps are hard to match with data, we will focus on age cohorts of individuals and establishments, which display smooth growth.

Each age cohort of individuals has a mix of workers, retoolers, and entrants, with proportions that change as the cohort ages. Among survivors at age $a \geq 0$, the shares of the three groups, $\sigma_w(a), \sigma_r(a), \sigma_{he}(a)$, are

$$\begin{aligned}
\sigma_w(a) &= \frac{\lambda_{hi}}{b_h} - \frac{\lambda_{he}}{b_h}\frac{b_h - \lambda_{hi}}{b_h - \lambda_{he}} e^{-b_h a} - \frac{\lambda_{hi} - \lambda_{he}}{b_h - \lambda_{he}} e^{-\lambda_{he} a}, \\
\sigma_r(a) &= 1 - \sigma_w(a) - \sigma_{he}(a), \\
\sigma_{he}(a) &= e^{-\lambda_{he} a}, \qquad\qquad a \geq 0,
\end{aligned}$$

where $b_h \equiv \lambda_{hi} + \alpha_h (g - \mu_h)$. The share of workers is zero at $a = 0$ and grows monotonically as the cohort ages, converging to λ_{hi}/b_h.

Similarly, each age cohort of establishments has a mix of producers, process innovators, and entrants. Among survivors at age $a \geq 0$, the shares of the three groups, $\sigma_p, \sigma_i, \sigma_{xe}$, are

$$\begin{aligned}
\sigma_p(a) &= \frac{\lambda_{xi}}{b_x} - \frac{\lambda_{xe}}{b_x}\frac{b_x - \lambda_{xi}}{b_x - \lambda_{xe}} e^{-b_x a} - \frac{\lambda_{xi} - \lambda_{xe}}{b_x - \lambda_{xe}} e^{-\lambda_{xe} a}, \\
\sigma_i(a) &= 1 - \sigma_p(a) - \sigma_{xe}(a), \\
\sigma_{xe}(a) &= e_0^{-\lambda_{xe} a}, \qquad\qquad a \geq 0,
\end{aligned}$$

where $b_x \equiv \lambda_{xi} + \alpha_x (g - \mu_x)$.

Figure 1 displays the demographics for the calibration with $\Omega = 0$ (no love-of-variety). Panel (a) shows the shares of survivors at each age for entering cohorts of labor and establishments, panel (b) shows the population shares for the labor force, and panel (c) shows the population shares for establishments. Note that there are almost no retoolers, at any age, but there is a substantial amount of process innovation. Indeed the share of producers peaks and then declines slightly, as producers switch to process innovation, and in the long run about a third of cohort of establishments are investing in process innovation. The difference in long run shares comes from the difference in the tail parameters for the technology and skill distributions, with the technology distribution having a much thicker density near its lower threshold.

Only workers receive wages, and at any date, the workers in every age cohort have average wages equal to the economy-wide average. Since average wages grow at the rate g_w, average earnings among survivors in the cohort of age a, relative to the average wage in the economy when they entered, are

$$e_{Av}(a) = e^{g_w a} \sigma_w(a), \qquad a \geq 0.$$

This average grows rapidly rapidly when the cohort is young, since σ_w grows rapidly as entrants transition to workers. As the cohort ages, $\sigma_w(a)$ converges to a constant and growth in average earnings slows, approaching g_w as a grows without bound. The number of survivors declines over time through exit, so total earnings for the cohort are

$$e_T(a) = e^{-\delta_h a}{}_{Av}(a), \qquad a \geq 0,$$

which grows when the cohort is young and declines when it is old.

Similarly, only producers have revenue and employees, and at any date producers in every age cohort have average revenue and employment equal to the economy-wide averages. Since the averages grow at rates g_π (revenue) and $\upsilon - n$ (employment), the averages among survivors in the cohort, relative to the levels when they entered, are

$$\begin{aligned} R_{Av}(a) &= e^{g_\pi a}\sigma_p(a), \\ \ell_{Av}(a) &= e^{(\upsilon-n)a}\sigma_p(a), \qquad a \geq 0. \end{aligned}$$

Both grow rapidly when the cohort is young, since σ_p grows rapidly as entrants transition to producers. As the cohort ages, $\sigma_p(a)$ converges to a constant, and average revenue (employment) among survivors grows in the long run if and only if $g_\pi > 0$ (revenue) or $\upsilon - n > 0$ (employment). The number of survivors in the cohort declines over time as establishments exit, so cohort totals are

$$\begin{aligned} R_T(a) &= e^{-\delta_x a}R_{Av}(a), \\ \ell_T(a) &= e^{-\delta_x a}\ell_{Av}(a), \qquad a \geq 0. \end{aligned}$$

Both have an inverse U-shape.

Figure 2 displays cohort earnings and employment. Figure 2a shows (log) cohort earnings, both the average per survivor and the cohort total. In the model, "entry" is the date at which cohort members start attempting to acquire, so at age $a = 0$, there are no workers and the cohort has no earnings. To better compare the model with data, it is useful to view $a = 4$ as the age when the cohort "enters." At this point about 55% of the cohort is working. Both curves are normalized to zero at that point. Earnings per survivor grow rapidly for a little less than a decade, and after that more slowly. The cohort total peaks quickly and then declines, due to exit.

Figures 2b shows (log) employment for a cohort of entering establishments.[3] At age $a = 4$, about 71% of the cohort is producing. Average employment per survivor grows rapidly for the next five years. It then declines slightly while the cohort is in middle age, and many survivors are switching from production to process innovation. In the cohort's old age the population shares are approximately constant. Since $\upsilon - n$

[3]The figure for revenue is very similar, since g_π and $\upsilon - n$ are both very small.

is very small, employment per survivor is approximately constant for very old cohorts. The cohort total starts declining almost immediately.

In the example, a modest subsidy to retooling has a large effect on TFP growth. Specifically, a subsidy of 20% of the marginal worker's wage increases TFP growth by almost one percentage point all three cases. It also reduces variety growth by 2.1 to 2.8 percentage points, depending on χ. The net effect on per capita income (wage) growth is between 0.9 and 0.5 percentage points, depending on χ,

$$\begin{aligned}
\Delta g &= [\ \ 0.00935, \quad 0.00928, \quad 0.00921]\,, \\
\Delta n &= [-0.0206, \quad -0.0243, \ -0.0278]\,, \\
\Delta g_w &= [\ \ 0.00935, \quad 0.00755, \quad 0.00526]\,.
\end{aligned}$$

In each case growth in profits per variety rises by about three percentage points, because of faster TFP growth and slower growth in varieties.

By contrast, a subsidy to process innovation of 5% of a marginal firm's profits has a zero or slightly negative effect on TFP growth and a strong positive effect on variety growth. The net effect on growth in per capita income is zero or slightly positive, depending on χ,

$$\begin{aligned}
\Delta g = &\quad [0.0, \qquad -0.000267, \quad -0.000519], \\
\Delta n = &\quad [0.00843, \quad 0.00979, \qquad 0.0111], \\
\Delta g_w = &\quad [0.0, \qquad 0.000432, \qquad 0.00105].
\end{aligned}$$

In each case growth in profits per variety falls by close to a percentage point, a result of the large increase in variety growth.

8. INCENTIVES OFF THE BGP

Complementarity between investments in technology and skill suggests that the incentive for continued investment in either factor relies on continued growth in the other. A complete analysis of dynamics away from the BGP is complicated, but the incentives for the leading or lagging factor to invest can be explored in a simple setting.

Lemma 7 provides conditions for existence of an equilibrium with no investment and no growth. Proposition 8 then specializes to economies in which marginal agents are just indifferent about investing, and looks at the incentives to invest in a perturbation of that environment. It shows that the incentive increases for the lagging factor and decreases for the leading factor.

For simplicity we consider economies where entry is impossible, $\phi_{xe} = 0$, and there is no population growth, no exit, and no on-the-job learning ($\upsilon = 0$, $\delta_h = \delta_x = 0$, $\mu_h = \mu_x = 0$). Thus, there is no growth in variety, $n = 0$, and both N and L are constant over time. TFP growth, if there is any, must come from investment by incumbents. Retooling and process innovation are possible, $\lambda_{hi} > 0$ and $\lambda_{xi} > 0$, but the pools of investors are initially empty, $N_i = L_i = 0$. Since $N_e = L_e = 0$ as well, $N_p = N$ and $L_w = L$. If there is no investment by incumbents, then the technology and skill distributions F and Ψ, and wages, prices, and the interest rate w, p, r, are constant over time.

In this setting the values $v_h(h)$ and $v_f(x)$ for workers and firms are simply the present discounted values of their wages and profits. If λ_{hi} and λ_{xi} are sufficiently small, then the expected values from retooling and process innovation are less than the values $v_h(h_m)$ and $v_f(x_m)$ of continuing to work or produce, and there is no TFP growth: $g = 0$. The interest rate is then the rate of time preference, $r = \hat{r}$. Formally, we have the following result.

Lemma 7: Let $\mathcal{E}$ be an economy with:

a. $\phi_{xe} = 0$, $\upsilon = 0$, $\delta_h = \delta_x = 0$, $\mu_h = \mu_x = 0$; and

b. initial conditions $N_{i0} = N_{e0} = 0$, $N_{p0} = N$, $L_{i0} = L_{e0} = 0$, $L_{w0} = L$.

Let $w(h), \pi(x)$ denote the wage and profit functions at $t = 0$, and let $v_h(h) = w(h)/\hat{r}$,

$v_f(x) = \pi(x)/\hat{r}$, denote the associated value functions. If

$$\begin{aligned} \hat{r} v_f(x_m) &\geq \lambda_{xi} \mathrm{E}_F \left[v_f(x) - v_f(x_m) \right], \\ \hat{r} v_w(h_m) &\geq \lambda_{hi} \mathrm{E}_\Psi \left[v_w(h) - v_w(h_m) \right], \end{aligned} \tag{43}$$

then $\mathcal{E}$ has a competitive equilibrium with no investment and no growth.

The two inequalities in (43) say that the firm with the worst technology and the worker with the lowest skill prefer, at least weakly, to continue producing and working. Hence there is an equilibrium in which the allocation and prices that prevail at $t = 0$ prevail at every date, and the interest rate and values are as claimed.

The next proposition uses this result, specialized in two ways. First, F and Ψ are assumed to be Pareto distributions that satisfy the hypotheses of Proposition 4. Thus, the wage and profit functions are isoelastic. In addition, the two conditions in (43) are assumed to hold with equality, so the marginal agents are indifferent about investing.

Fix the parameters α_x, x_m for F and α_h for Ψ. For the baseline economy suppose the location parameter for skill is $h_m = a_h x_m$, so Proposition 4 holds, and let $h^*(x)$, $w(h)$, $v_h(h)$, etc. denote the equilibrium skill allocation, wage function, value function and so on in the baseline economy. For the perturbed economy suppose the skill distribution is shifted to the right, with

$$\hat{h}_m \equiv h_m (1 + \Delta), \tag{44}$$

where $\Delta > 0$ is small. Let $\hat{\Psi}(h)$, $\hat{h}^*(x)$, $\hat{w}(h)$, $\hat{v}_h(h)$, etc. denote the distribution function for skill, the equilibrium skill allocation, and so on in the perturbed economy.

Proposition 8 shows that if a producer with technology x_m is indifferent about investing in process innovation in the baseline economy, then it strictly prefers to invest in the perturbed economy; and if a worker with skill h_m is indifferent about retooling in the baseline economy, then a worker with skill $\hat{h}_m$ strictly prefers *not* to retool in the perturbed economy. Formally, we have the following result.

PROPOSITION 8: Let $\mathcal{E}$ be an economy satisfying the hypotheses of Lemma 7. Assume in addition that F_0, Ψ_0, that are Pareto distributions, with shape and location parameters (α_x, x_m), (α_h, h_m) satisfying the hypotheses of Proposition 4, and that the two conditions in (43) hold with equality.

Let $\mathcal{E}_\Delta$ be an economy that is the same as $\mathcal{E}$ except that it has distribution $\hat{\Psi}$ with location parameter $\hat{h}_m$ in (44). Let $\hat{h}^*(x)$, $\hat{w}(h)$, etc. denote equilibrium for $\mathcal{E}_\Delta$, and let $\hat{v}_h(h) = \hat{w}(h)/\hat{r}$ and $\hat{v}_f(x) = \hat{\pi}(x)/\hat{r}$. Then

$$\begin{aligned} \hat{r}\hat{v}_f(x_m) &< \lambda_{xi}\mathrm{E}_F\left[\hat{v}_f(x) - \hat{v}_f(x_m)\right]; \\ \hat{r}\hat{v}_w(\hat{h}_m) &> \lambda_{hi}\mathrm{E}_{\hat{\Psi}}\left[\hat{v}_w(h) - \hat{v}_w(\hat{h}_m)\right]. \end{aligned} \tag{45}$$

The proof involves first-order approximations to the equilibrium allocation and prices in the perturbed economy, so the same argument applies for a small leftward shift in the skill distribution or for small shifts in the technology distribution.

Proposition 8 suggests that in economies where the initial skill and technology distributions are (sufficiently) misaligned, there may be a period where there is investment only in the lagging factor. It also suggests that in the long run, sustained investment in either factor requires sustained investment in its complement.

9. CONCLUSION

This paper develops a model in which investments in both technology and skill acquisition are required for long run growth. Growth, in turn, takes two forms: TFP growth and growth in product variety. The main results are to provide conditions for the existence of a BGP, to show how the rates of TFP and variety growth depend on the parameters governing technology and skill acquisition, and to show that the incentive to invest in either factor depends on the availability of its complement.

On a BGP skill and technology grow at a common rate, which is also the rate of TFP growth. Nevertheless, the parameters governing skill accumulation are more important than those governing technological change in determining that rate. The parameters for skill and technology enter more symmetrically—but with opposite signs—in determining growth in product variety. Thus, improvements in the parameters for technological change encourage entry, while improvements in the parameters for skill accumulation encourage investment in both skill and technology, but discourage growth in variety.

In a roughly calibrated numerical example, a subsidy to retooling has a large effect on TFP growth, and hence on growth in per capital output and consumption. The reason for this is interesting as well: here a subsidy to retooling is powerful not because of the direct effect on the workers receiving the training, but rather because it changes the distribution from which new entrants to the labor force draw their skills. Role models are important here: they are they templates that new workers imitate. Removing the lowest-skill role models means that young workers who would have imitated them instead draw from skills higher in the distribution.

In equilibrium, continued investment in either factor remains worthwhile only because the other grows: the technology and skill distributions must shift together. Although transitional dynamics are not studied in detail, an example shows that the incentive for either factor to invest depends on the relative level of development of its complement. In particular, if one factor has a distribution that gets 'ahead' of the other, the incentive to invest in the leading factor declines while the incentive to invest in the lagging factor gets stronger. This result suggests that the economy converges to a BGP for any initial distributions that are Pareto, or at least have Pareto tails.

The model here also suggests that growth rates depend on various types of investment in a highly nonlinear way. Thus, nonparametric techniques may be more useful than linear regression models for empirical studies of these relationships.

REFERENCES

[1] Acemoglu, D. 1998. Why do new technologies complement skills? Directed technical change and wage inequality. *Quarterly Journal of Economics*, 113(4), 1055-1089.

[2] Acemoglu, D. 2002. Technical change, inequality, and the labor market. *Journal of Economic Literature*, 40(1), 7-72.

[3] Akerman, Anders, Ingvil Gaarder, and Magne Mogstad. 2015. The skill complementarity of broadband internet, *Quarterly Journal of Economics*, 130(4), 1781-1824.

[4] Allen, Treb, Kamran Bilir, and Christopher Tonetti. 2017. The spatial diffusion of knowledge, Working paper.

[5] Alvarez, Fernando E., Francisco J. Buera, and Robert E. Lucas, Jr. 2014. Idea flows, economic growth, and trade, N.B.E.R. Working Paper 19667.

[6] Arrow, Kenneth J. 1962. The economic implications of learning by doing. Review of Economic Studies, 29(3): 155-173.

[7] Autor, David H., Lawrence F. Katz and Alan B. Krueger. 1998. Computing inequality: have computers changed the labor market? *Quarterly Journal of Economics*, 113(4), 1169-1213.

[8] Autor, David H., Frank Levy and Richard J. Murnane. 2003. The skill content of recent technological change: an empirical exploration, *Quarterly Journal of Economics*, 118(4): 1279-1333.

[9] Bahk, Byong-hyong and Michael Gort. 1993. Decomposing learning by doing in new plants. *Journal of Political Economy*, 101(4), 561-583.

[10] Bartel, Anne, Casey Ichniowski and Kathryn Shaw. 2007. How does information technology affect productivity? Plant-level comparisons of product innovation,

process improvement, and worker skills. *Quarterly Journal of Economics*, 122(4), 1721-1758.

[11] Bartel, Anne, and Frank Lichtenberg. 1987. The comparative advantage of educated workers in implementing new technology, *The Review of Economics and Statistics*, 69(1): 1-11.

[12] Beaudry, Paul, Mark Doms and Ethan Lewis. 2010. Should the personal computer be considered a technological revolution? Evidence from U.S. metropolitan areas. *Journal of Political Economy*, 118(5), 988-1036.

[13] Benhabib, Jess, Jesse Perla, and Christopher Tonetti. 2017. Reconciling models of diffusion and innovation: a theory of the productivity distribution and technology frontier. Working paper.

[14] Berman, Eli, John Bound and Zvi Griliches. 1994. Changes in the demand for skilled labor within U.S. manufacturing: evidence from the Annual Survey of Manufacturers. *Quarterly Journal of Economics*, 109(2), 367-397.

[15] Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt. 2002. Information technology, workplace organization, and the demand for skilled labor: firm-level evidence, *Quarterly Journal of Economics*, 117(1): 297-338.

[16] Buera, Francisco J. and Robert E. Lucas, Jr. 2017. Idea flows and economic growth, *Annual Review of Economics*, forthcoming.

[17] Buera, Francisco J. and Ezra Oberfield. 2017. The global diffusion of ideas. Working paper.

[18] Caicedo, Santiago, Robert E. Lucas, Jr., and Esteban Rossi-Hansberg. 2016. Learning, career paths, and the distribution of wages, NBER Working Paper 22151. (forthcoming, *American Economic Journal, Macroeconomics*)

[19] Caselli, Francesco. 1999. Technological revolutions, *America Economic Review*, 89(1): 78-102.

[20] Comin, Diego and Bart Hobijn. 2004. Cross-country technology adoption: making the theories face the facts. *Journal of Monetary Economics*, 51(1): 39-83

[21] ______. 2010. An exploration of technology diffusion. *American Economic Review*, 100(5):2031-59.

[22] ______. 2017. If technology has arrived everywhere, why has income diverged? Working paper.

[23] Costinot, Arnaud, and Jonathan Vogel. 2010. Matching and inequality in the world economy. *Journal of Political Economy*, 118(4): 747-786.

[24] Doms, Mark, Timothy Dunne, and Kenneth R. Troske. 1997. Workers, wages, and technology, *Quarterly Journal of Economics* 112(1): 253-90.

[25] Dunne, Timothy, John Haltiwanger, and Kenneth R. Troske. 1997. Technology and jobs: secular changes and cyclical dynamics, *Carnegie-Rochester Conference Series on Public Policy*, 46: 107-178.

[26] Dunne, Timothy, Mark J. Roberts and Larry Samuelson. 1989. The growth and failure of U. S. manufacturing plants. *Quarterly Journal of Economics*, 104(4), 671-698.

[27] Eaton, Jonathan and Samuel Kortum. 1999. International technology diffusion: theory and measurement, *International Economic Review*, 40: 537-570.

[28] Forman, Chris, Avi Goldfarb and Shane Greenstein. 2012. The internet and local wages: a puzzle. *American Economic Review*, 102(1), 556-575.

[29] Foster, Andrew D. and Mark R. Rosenzweig. 1996. Technical change and human-capital returns and investments: evidence from the Green Revolution. *American Economic Review*, 86(4), 931-953.

[30] ______. 2010. Microeconomics of technology adoption, *Annual Review of Economics*, Annual Reviews, vol. 2(1), pages 395-424, September.

[31] Goldin, Claudia, and Lawrence F. Katz. 1998. The origins of technology-skill complementarity. *Quarterly Journal of Economics*, 113(3), 693-732.

[32] ______. 2008. *The Race between Education and Technology*, Harvard University Press.

[33] Gort, Michael and Steven Klepper. 1982. Time paths in the diffusion of product innovations. *Economic Journal*, 92(367), 630-653.

[34] Greenwood, Jeremy and Mehmet Yorukoglu. 1997. "1974," *Carnegie-Rochester Series on Public Policy*, 46: 49-95.

[35] Griliches, Z. 1957. Hybrid corn: an exploration in the economics of technological change. *Econometrica*, 25(4), 501-522.

[36] ______. 1969. Capital-skill complementarity. *The Review of Economics and Statistics*, 51(4), 465-468.

[37] Hansen, Gary D. and Selahattin Imrohoroglu. 2009. Business cycle fluctuations and the life cycle: How important is on-the-job skill accumulation? *Journal of Economic Theory*, 144(6): 2293-2309.

[38] Jovanovic, Boyan. 1998. Vintage capital and inequality. *Review of Economic Dynamics*, 1: 497-530.

[39] ______. 2009. The technology cycle and inequality. *Review of Economic Studies*, 76: 707-729.

[40] Jovanovic, Boyan, and Glenn M. MacDonald. 1994. Competitive diffusion, *Journal of Political Economy*, 102: 24-52.

[41] Jovanovic, Boyan and Yaw Nyarko. 1996. Learning by doing and the choice of technology. *Econometrica*, 64(6), 1299-1310.

[42] Jovanovic, Boyan, and Rafael Rob. 1989. The growth and diffusion of knowledge, *Review of Economic Studies*, 56: 569-582.

[43] Kortum, Samuel. 1997. Research, patenting, and technological change, *Econometrica*, 65: 1389-1419.

[44] Krusell, Per, et. al. 2000. Capital-skill complementarity and inequality: a macroeconomic analysis, *Econometrica*, 68(5): 1029-1053.

[45] Lloyd-Ellis, Huw, and Joanne Roberts. 2002. Twin engines of growth: skill and technology as equal partners in balanced growth, *Journal of Economic Growth*, 7(2):87-115.

[46] Lucas, Robert E. 2009. Ideas and growth, *Economica*, 76: 1-19.

[47] ______. 2015. Human capital and growth. *American Economic Review, Papers and Proceedings*, 105(5) 85-88.

[48] Lucas, Robert E., and Benjamin Moll. 2014. Knowledge growth and the allocation of time, *Journal of Political Economy*, 122(1): 1-51.

[49] Luttmer, Erzo J.G. 2007. Selection, growth and the size distribution of firms, *Quarterly Journal of Economics* 122(3) 1103–44.

[50] Machin, Stephen and John Van Reenen. 1998. Technology and changes in skill structure: evidence from seven OECD countries, *Quarterly Journal of Economics,* 113(4): 1215-1244.

[51] Mansfield, Edwin. 1961. Technical change and the rate of imitation. *Econometrica*, 29(4), 741-766.

[52] Manuelli, Rodolfo E., and Ananth Seshadri. 2014. Frictionless technology diffusion: the case of tractors. *American Economic Review*, 104(4): 1368-91.

[53] Michaels, Guy Ashwini Natraj, and John Van Reenen. 2014. Has ICT polarized skill demand? Evidence from eleven countries over twenty-five years, *Review of Economics and Statistics*, 96(1): 60-77.

[54] Nelson, Richard R., Merton Peck, and Edward Kalachek. 1967, *Technology, Economic Growth, and Public Policy*. Washington, D.C., Brookings.

[55] Nelson, Richard R. and Edmund S. Phelps. 1966. Investment in humans, technological diffusion, and economic growth. *American Economic Review*, 56 (1/2): 69-75.

[56] Parente, Stephen L. and Edward C. Prescott. 1999. Monopoly rights: a barrier to riches, *American Economic Review,* 89: 1216-1233.

[57] Perla, Jesse, and Chris Tonetti. 2014. Equilibrium imitation and growth, *Journal of Political Economy*, 122(1): 52-76.

[58] Perla, Jesse, Christopher Tonetti, and Michael E. Waugh. 2015. Equilibrium technology diffusion, trade, and growth. Working paper.

[59] Stokey, Nancy L. 1988. Learning-by-doing and the introduction of new goods. *Journal of Political Economy*, 96: 369-405.

[60] ______. 1996. Free trade, factor returns, and factor accumulation, *Journal of Economic Growth*, 1: 421-447.

[61] ______. 2015. Catching up and falling behind, *Journal of Economic Growth*, 20: 1-36.

[62] ——. 2016. Technology, skill and the wage structure, NBER Working Paper 22176. (forthcoming, *Journal of Human Capital*)

[63] ——. 2018. The race between technology and human capital. Working paper.

[64] Violante, Giovanni L. 2002. Technological acceleration, skill transferability, and the rise in residual inequality, *Quarterly Journal of Economics*, 117(1): 339-376.

APPENDIX A: PRODUCTION AND PRICES

A. Production equilibrium

Proof of Proposition 1: Use (5) to write labor demand as

$$\ell(x) = N_p^{-\rho\chi}\left(\frac{\rho-1}{\rho}\right)^{\rho}\frac{\phi(h^*(x),x)^{\rho-1}}{w(h^*(x))^{\rho}}y_F, \qquad \text{all } x \geq x_m, \tag{46}$$

and differentiate (7) to write labor market clearing as

$$\begin{aligned} L_w\psi(h^*(x))h^{*\prime}(x) &= N_p\ell(x)f(x), \qquad \text{all } x \geq x_m, & (47)\\ L_w &= N_p\int_{x_m}^{\infty}\ell(\xi)f(\xi)d\xi. & (48)\end{aligned}$$

Then (4) and (47) are a pair of ODEs in $w(h)$ and $h^*(x)$, with $\ell(x)$ given by (46). The price normalization (2) serves as a boundary condition for w, and (6) is the boundary condition for h^*, with y_F determined by (48). The other equations in (5) determine p, y, π. ■

Proof of Proposition 4: The functions in (14)-(17) must satisfy (2), (4)-(6) and (47)-(48).

For any wage function of the form $w(h) = w_0\left(h/h_m\right)^{1-\varepsilon}$, as in (15), the linear function $h^*(x) = a_h x$ in (14), with a_h in (12), satisfies (4). Moreover, for $w(h)$ of this form the first line in (5) implies that

$$p(x) = \frac{\rho}{\rho-1}\frac{w_2}{x_m}\frac{\left(x/x_m\right)^{-\varepsilon}}{\phi(a_h,1)}, \qquad \text{all } x \geq x_m.$$

Then (2) implies w_2 is as in (18), with p_0 as claimed.

Then from the second and third lines of (5), $y(x)$ and $\ell(x)$ have the form

$$
\begin{aligned}
y(x) &= y_2 \left(\frac{x}{x_m}\right)^{\rho\varepsilon}, && \text{all } x \geq x_m. \\
\ell(x) &= \ell_2 \left(\frac{x}{x_m}\right)^{\rho\varepsilon-1}, && \text{all } x \geq x_m.
\end{aligned}
$$

Using the Pareto distribution for F, (48) holds if and only if

$$
\begin{aligned}
L_w &= N_p \ell_2 \mathrm{E}_F \left[\left(\frac{x}{x_m}\right)^{\alpha_x-\alpha_h}\right] \\
&= N_p \ell_2 \frac{\alpha_x}{\alpha_h},
\end{aligned}
$$

so ℓ_2 is as claimed. Then $y(x)$ and y_F follow from the production technologies, and $\pi(x)$ is straightforward from (5). The ODE in (47) requires

$$
\psi(a_h x) a_h = \frac{\alpha_h}{\alpha_x} \left(\frac{x}{x_m}\right)^{\alpha_x-\alpha_h} f(x), \qquad \text{all } x \geq x_m,
$$

which holds for the Pareto densities ψ and f, and clearly (13) implies (6), completing the proof. ■

B. The evolution of technology

The distribution function for technology among producers evolves as follows. As noted above, $X_m(t)$ is nondecreasing. Let $\Delta_t > 0$ be a small time increment. For any $t \geq 0$ and any $X \geq X_m(t+\Delta_t)$, the number of producers with technology above X at $t+\Delta_t$ consists of incumbents at t, adjusted for exit, plus successful innovators of both types, selected to include only those with technology greater than $(1-\mu_x \Delta_t) X$ at date t,

$$
\begin{aligned}
& [1 - F(X, t+\Delta_t)] \, N_p(t+\Delta_t) \\
\approx \; & \{1 - F[(1-\mu_x\Delta_t) X, t]\} \, [(1-\delta_x \Delta_t) N_p(t) + \lambda_{xi}\Delta_t N_i(t) + \Lambda_{xe}(t) \Delta_t N_e(t)].
\end{aligned}
$$

Taking a first-order approximation gives

$$[1 - F(X,t)] \left[N_p(t) + N_p'(t)\Delta_t\right] - F_t(X,t)\Delta_t N_p(t)$$

$$\approx \quad [1 - F(X,t)] \left[(1 - \delta_x \Delta_t) N_p(t) + \lambda_{xi}\Delta_t N_i(t) + \Lambda_{xe}(t)\Delta_t N_e(t)\right]$$
$$+ f(X,t)\mu_x \Delta_t X N_p(t).$$

Collecting terms and dividing by $\Delta_t N_p(t)$ gives the equation in the text.

APPENDIX B: BGPS WITH PARETO DISTRIBUTIONS

A. Firms: process and product innovation

If Π and V^f have the forms required for a BGP, then factoring out $e^{g_\pi t}Q_0$, the Bellman equation for a producing firm is

$$\begin{aligned}(r + \delta_x) v_{fp}(X/Q(t)) \quad &= \quad \pi(X/Q(t)) + \mu_x \frac{X}{Q(t)} v_{fp}'(X/Q(t)) \\ &\quad + g_\pi v_{fp}(X/Q(t)) - v_{fp}'(X/Q(t)) \frac{X}{Q(t)} \frac{\dot{Q}(t)}{Q(t)},\end{aligned}$$

or

$$(r - g_\pi + \delta_x) v_{fp}(x) = \pi(x) - (g - \mu_x) x v_{fp}'(x),$$

where $x = X/Q$ and $\dot{Q}/Q = g$. For π as in (17), the normalized Bellman equation is as claimed. Define

$$\begin{aligned} R_x \quad &\equiv \quad -\frac{r - g_\pi + \delta_x}{g - \mu_x} < 0, \\ B_x \quad &\equiv \quad \left[(g - \mu_x)(1 - \zeta - R_x)\right]^{-1} > 0, \end{aligned} \tag{49}$$

where R_x is the characteristic root of the ODE. It is straightforward to verify that $v_P(x) = B_x \pi_2 (x/x_m)^{1-\zeta}$ is a particular solution, and $v_H(x) = c_x x^{R_x}$ is the homogeneous solution. The coefficient c_x is determined by the value matching condition,

$$c_x = x_m^{-R_x} (v_{fi} - B_x \pi_2) > 0,$$

and $v_{fp}(x)$ is as in (25). Differentiate (25) to get the smooth pasting condition, $v'_{fp}(x_m) = 0$, so v_{fi} is as in (26).

On a BGP, the Bellman equation for v_{fi} is as claimed. To obtain (27), substitute from (26) and take the expectation in (25) to get

$$\mathrm{E}_{\hat{F}}\left[v_{fp}(x)\right] = B_x \pi_2 \left(\frac{\alpha_x}{\alpha_x - 1 + \zeta} - \frac{1-\zeta}{R_x} \frac{\alpha_x}{\alpha_x - R_x} \right). \tag{50}$$

Use this expression and (26) in the Bellman equation for v_{fi} to get (27),

$$\begin{aligned} r - g_\pi + \delta_x &= \lambda_{xi} \left(\frac{\mathrm{E}_{\hat{F}}\left[v_{fp}(x)\right]}{v_{fi}} - 1 \right) \\ &= \frac{\lambda_{xi} R_x}{R_x - 1 + \zeta} \left(\frac{\alpha_x}{\alpha_x - 1 + \zeta} - \frac{1-\zeta}{R_x} \frac{\alpha_x}{\alpha_x - R_x} - 1 + \frac{1-\zeta}{R_x} \right) \\ &= \frac{\lambda_{xi} R_x}{R_x - 1 + \zeta} \left(\frac{1-\zeta}{\alpha_x - 1 + \zeta} - \frac{1-\zeta}{\alpha_x - R_x} \right) \\ &= \frac{\lambda_{xi} R_x (1-\zeta)}{R_x - 1 + \zeta} \frac{-R_x + 1 - \zeta}{(\alpha_x - 1 + \zeta)(\alpha_x - R_x)} \\ &= \frac{-R_x (1-\zeta) \lambda_{xi}}{(\alpha_x - 1 + \zeta)(\alpha_x - R_x)}. \end{aligned}$$

To show that the optimal investment for a product innovator is at $i_e/\bar{i}_e = 1$, use the definition of v_{fe} in (28) and the definition of Λ_{xe} in the neighborhood of $i_e/\bar{i}_e = 1$ to find that

$$\begin{aligned} v'_{fe}(\lambda) &= \frac{r - g_\pi + \delta_x}{(r - g_\pi + \delta_x + \lambda)^2} \mathrm{E}_{\hat{F}}\left[v_{fp}(x)\right], \\ \frac{\partial}{\partial i_e} \Lambda_{xe}(i_e/\bar{i}_e, \epsilon) &= \frac{\phi_e}{\epsilon} \frac{1}{\bar{i}_e} \cdot \begin{cases} 1/\varepsilon_x, & \text{if } i_e/\bar{i}_e < 1, \\ \varepsilon_x, & \text{if } i_e/\bar{i}_e > 1. \end{cases} \end{aligned}$$

Hence an optimum at $i_e/\bar{i}_e = 1$, for $\bar{i}_e = \mathrm{E}_{\hat{F}}\left[v_{fp}(x)\right]$, requires (31).

For any $q < \alpha_x$, integrating w.r.t. the density $f(x) = \alpha_x x_m^{\alpha_x} x^{-\alpha_x - 1}$ gives $\mathrm{E}_{\hat{F}}\left[(x/x_m)^q\right] = \alpha_x / (\alpha_x - q)$.

B. Workers: investment in human capital on the BGP

The analysis for the labor force is analogous, except that trainees pay no entry cost. Hence (32)-(35) have the same form as (25)-(28), with

$$\begin{aligned} R_h &\equiv -\frac{r - g_w + \delta_h}{g - \mu_h} < 0, \\ B_h &\equiv \left[(g - \mu_h)(1 - \varepsilon - R_h)\right]^{-1} > 0. \end{aligned} \tag{51}$$

C. Flows of firms and workers, the DFs for technology and skill

On a BGP $X_m(t)$ grows at the rate g; $N_p(t), N_i(t)$, and $N_e(t)$ grow at the rate n; and there is strictly positive process innovation, so (19) holds. Hence the law of motion for N_p requires

$$\begin{aligned} nN_p &= \lambda_{xi} N_i + \lambda_{xe} N_e - \delta_x N_p - (g - \mu_x) \frac{X_m(t)}{Q(t)} f(x_m) N_p \\ &= \lambda_{xi} N_i + \lambda_{xe} N_e - \left[\delta_x + \alpha_x (g - \mu_x)\right] N_p, \end{aligned}$$

where the second line the Pareto density for $f(x)$. Hence

$$\left[n + \delta_x + \alpha_x (g - \mu_x)\right] N_p = \lambda_{xi} N_i + \lambda_{xe} N_e.$$

The laws of motion for N_i and N_e require

$$\begin{aligned} (n + \delta_x + \lambda_{xi}) N_i &= \alpha_x (g - \mu_x) N_p, \\ (n + \delta_x + \lambda_{xe}) N_e &= E. \end{aligned}$$

Sum the three laws of motion to get (36), which determines the entry rate E. The population shares for firms are

$$\begin{aligned} \frac{N_p}{N} &= \frac{n + \delta_x + \lambda_{xi}}{(n + \delta_x + \lambda_{xi}) + \alpha_x (g - \mu_x)} \frac{\lambda_{xe}}{n + \delta_x + \lambda_{xe}}, \\ \frac{N_i}{N} &= \frac{\alpha_x (g - \mu_x)}{(n + \delta_x + \lambda_{xi}) + \alpha_x (g - \mu_x)} \frac{\lambda_{xe}}{n + \delta_x + \lambda_{xe}}, \\ \frac{N_e}{N} &= \frac{n + \delta_x}{n + \delta_x + \lambda_{xe}}, \end{aligned}$$

and the ratios N_i/N_p and N_e/N_p are as in (37).

If $F(X,t)$ has the form required for a BGP and $Q(t)$ grows at the rate g, then

$$\begin{aligned} f(X,t) &= \hat{f}(X/Q(t))/Q(t), \\ -F_t(X,t) &= \hat{f}\left(X/Q(t)\right) gX/Q(t), \qquad \text{all } X \geq X_m(t), \ \ \text{all } t. \end{aligned}$$

Use these expressions in the law of motion for F to get

$$(g-\mu_x)\,\hat{f}\left(X/Q(t)\right)X/Q(t) = \left[1-\hat{F}(X/Q(t))\right]\alpha_x\left(g-\mu_x\right),$$

or

$$(g-\mu_x)\,x\hat{f}\left(x\right) = \left(g-\mu_x\right)\alpha_x\left[1-\hat{F}\left(x\right)\right], \qquad \text{all } x \geq x_m,$$

which holds since F is a Pareto distribution with parameters (α_x, x_m).

The arguments for the workforce are analogous.

D. Proofs of Propositions 5 and 6

Proof of Proposition 5: Write the equations in (27) and (34) as

$$\begin{aligned} g &= \frac{1}{\xi_h}\left[\qquad\qquad Z_h + A_h\lambda_{hi} - W_n\Omega n\right], \\ g &= \frac{1}{\xi_x}\left[\upsilon - n + Z_x + A_x\lambda_{xi} - W_n n\right], \end{aligned} \tag{52}$$

where $Z_x, Z_h, A_x.A_h, W_n$, are as defined in (42). For existence and uniqueness of a solution to (52), it suffices to show that the two equations are not collinear. Here we will prove a slightly stronger result, that

$$\frac{1}{\xi_x}\left[1+(\theta-1)\,\Omega\right] > \frac{1}{\xi_h}\left(\theta-1\right)\Omega,$$

or

$$\alpha_h > (\theta-1)\left[\left(\alpha_x-\alpha_h\right)\Omega-1\right]. \tag{53}$$

Since $\chi \in (0, 1/\rho]$ implies $\Omega \in [0, 1/(\rho - 1))$, and (9) implies $\alpha_x - \alpha_h \in (-1, \rho - 1)$, it follows that

$$(\alpha_x - \alpha_h)\,\Omega - 1 \in \left(-\frac{\rho}{\rho - 1},\ 0\right).$$

Hence if $\theta \geq 1$, the term on the right in (53) is zero or negative. If $\theta < 1$, then by assumption $\alpha_h > \rho/(\rho - 1)$. In either case (53) holds, and there exists a unique (g, n) satisfying (42).

Define $[Q_0, \hat{F}, \hat{\Psi}]$, x_m, h_m by (22)-(24). Since (24) implies (13) holds, by Proposition 4 the normalized production equilibrium $[w, p, h^*, y, \ell, \pi, y_F]$ is described by (14)-(17), so the price and allocation functions are isoelastic.

Then (25), (26) and (28) determine $[v_{fp}, v_{fi}, v_{fe}]$; (29) determines v_{fe0} as a function of λ_{xe0}; (30) determines i_{e0}; (32), (33) and (35) determine $[v_w, v_{wi}, v_{we}]$; (36) determines E_0; (38) determines λ_{xe0}; (40) and (41) determine r and c_0; and (d) implies that the the required inequalities hold. Hence the solution describes a BGP. ■

Proof of Proposition 6: Plotted in n-$g-$space, the pair of equations in (52) are as shown Figure A1: the line defined by the first equation is downward sloping; the line defined by the second equation has a positive, zero, or negative slope as $(\theta - 1)\,\Omega < 0$, $= 0$, or > 0; and in all case the second line crosses the first from below.

For the first claim, note that $\rho \geq 2$ implies $\Omega \in [0, 1)$, so $[(\theta - 1)\,\Omega + 1] > 0$. The second claim is obvious, and the third follows from (53).

Then claims (a) - (c) follow directly from Figure A1. As shown in panel (a), an increase in μ_h or λ_{hi}, or a decrease in δ_h, shifts the second line upward, increasing g and decreasing n. As shown in panel (b), an increase in υ, μ_x or λ_{xi}, or a decrease in δ_x, shifts the first line to the right, increasing n. The effect on g depends on the slope of the second line. A decrease in $\hat{r}$ does both, as shown in panel (c). Hence it raises g if $(\theta - 1)\,\Omega \leq 0$, and otherwise the effects depend on the relative slopes of the two lines. ■

E. Proof that $c_0 > 0$ in the example

From (41), $c_0 > 0$ if and only if $E_0 i_{e0}/y_F < 1$. Use (29) and (30) to find that

$$i_{e0} = \frac{\lambda_{xe}}{r - g_\pi + \delta_x + \lambda_{xe}} \mathrm{E}_{\hat{F}}\left[v_{fp}(x)\right].$$

Since $(\mu_x - g)\, x v_{fp}'(x) < 0$, the Bellman equation for v_{fp} implies

$$\begin{aligned} \mathrm{E}_{\hat{F}}\left[v_{fp}(x)\right] &< \frac{1}{r - g_\pi + \delta_x} \mathrm{E}_{\hat{F}}\left[\pi_2 \left(\frac{x}{x_m}\right)^{1-\zeta}\right] \\ &= \frac{1}{r - g_\pi + \delta_x} \frac{1}{\rho} \frac{y_F}{N_{p0}}, \end{aligned}$$

where the second line uses the fact that $1/\rho$ is the factor share of profits in income. Use these two expression and $E_0 = (n + \delta_x)\, N_0$ to get

$$\frac{E_0 i_{e0}}{y_F} < \frac{\lambda_{xe}}{r - g_\pi + \delta_x + \lambda_{xe}} \frac{1}{r - g_\pi + \delta_x} \frac{1}{\rho} \frac{(n + \delta_x)\, N_0}{N_{p0}}. \tag{54}$$

Hence $E_0 i_{e0}/y_F < 1$ if the term on the right in (54) is less than unity. Hence it suffices if

$$\begin{aligned} \frac{\lambda_{xe}}{r - g_\pi + \delta_x + \lambda_{xe}} \frac{n + \delta_x}{r - g_\pi + \delta_x} \frac{1}{\rho} &< \frac{N_{p0}}{N_0} \\ &= \frac{n + \delta_x + \lambda_{xi}}{(n + \delta_x + \lambda_{xi}) + \alpha_x (g - \mu_x)} \frac{\lambda_{xe}}{n + \delta_x + \lambda_{xe}} \end{aligned}$$

where the second line uses N_p/N from above, or

$$\frac{n + \delta_x + \lambda_{xe}}{r - g_\pi + \delta_x + \lambda_{xe}} \frac{n + \delta_x}{r - g_\pi + \delta_x} \frac{1}{\rho} < \frac{n + \delta_x + \lambda_{xi}}{(n + \delta_x + \lambda_{xi}) + \alpha_x (g - \mu_x)}.$$

In the example, $n < r - g_\pi$, so the left side is less than $1/\rho$, while since $\alpha_x (g - \mu_x)$ is small, the right side is close to unity.

F. Dynamics of cohort composition

As a function of age, the shares of firms of various types satisfy

$$\begin{pmatrix} \sigma_p'(a) \\ \sigma_i'(a) \\ \sigma_e'(a) \end{pmatrix} = \begin{pmatrix} -\alpha_x (g-\mu_x) & \lambda_{xi} & \lambda_{xe} \\ \alpha_x (g-\mu_x) & -\lambda_{xi} & 0 \\ 0 & 0 & -\lambda_{xe} \end{pmatrix} \begin{pmatrix} \sigma_p(a) \\ \sigma_i(a) \\ \sigma_e(a) \end{pmatrix}.$$

The characteristic roots are $R_j = 0, -b_x, -\lambda_{xe}$, where $b_x \equiv \alpha_x (g-\mu_x) + \lambda_{xi}$. It is then straightforward to find that

$$\begin{aligned} \sigma_p(a) &= \frac{\lambda_{xi}}{b_x} - \frac{\lambda_{xe}}{b_x}\frac{b_x-\lambda_{xi}}{b_x-\lambda_{xe}} e^{-b_x a} - \frac{\lambda_{xi}-\lambda_{xe}}{b_x-\lambda_{xe}} e^{-\lambda_{xe} a}, \\ \sigma_i(a) &= 1 - \frac{\lambda_{ix}}{b_x} + \frac{\lambda_{xe}}{b_x}\frac{b_x-\lambda_{xi}}{b_x-\lambda_{xe}} e^{-b_x a} - \frac{b_x-\lambda_{xi}}{b_x-\lambda_{xe}} e^{-\lambda_{xe} a}, \\ \sigma_e(a) &= 0 + 0 + e^{-\lambda_{xe} a}. \end{aligned}$$

The argument for workers is analogous.

APPENDIX C: NO-GROWTH RESULT

PROOF OF PROPOSITION 8: If wages and prices are constant over time, a firm's value is proportional to its profit, which in turn is proportional to its total wage bill, and a worker's value is proportional to his wage. Hence v_f and v_w in (43) and (45) can be replaced with $w(h^*)\ell$ and w, respectively. Then, since by hypothesis the conditions in (43) hold with equality, the conditions in (45) are equivalent to

$$0 < \mathrm{E}_F \left[\frac{\hat{w}[\hat{h}^*(x)]}{\hat{w}[\hat{h}^*(x_m)]} \frac{\hat{\ell}(x)}{\hat{\ell}(x_m)} - \frac{w[h^*(x)]}{w[h^*(x_m)]} \frac{\ell(x)}{\ell(x_m)} \right], \tag{55}$$

$$0 > \mathrm{E}_{\hat{\Psi}} \left[\frac{\hat{w}(h)}{\hat{w}(h_m)} \right] - \mathrm{E}_{\Psi} \left[\frac{w(h)}{w(h_m)} \right], \qquad \text{for } \Delta > 0. \tag{56}$$

By assumption Proposition 4 holds for the baseline economy, so the functions $w(h) = w_0 h^{1-\varepsilon}, h^*(x) = a_h x$ and $\ell(x) = \ell_0 x^{\alpha_x - \alpha_h}$ are isoelastic. The proof uses approximations to $\hat{w}, \hat{h}^*$ and $\hat{\ell}$ that are linear in Δ,

$$\begin{aligned}
\hat{w}(h) &\approx w(h)\left[1+\sigma(h)\Delta\right], \qquad h \geq h_m, \\
\hat{h}^*(x) &\approx h^*(x)\left[1+m(x)\Delta\right], \\
\hat{\ell}(x) &\approx \ell(x)\left[1+q(x)\Delta\right], \qquad x \geq x_m,
\end{aligned}$$

with

$$\begin{aligned}
\frac{d\ln w}{d\ln h} &\approx (1-\varepsilon)+h\sigma'(h)\Delta, \qquad h > h_m, \\
\frac{d\ln h^*}{d\ln x} &\approx 1+xm'(x)\Delta, \\
\frac{d\ln \ell}{d\ln x} &\approx (\alpha_x-\alpha_h)+xq'(x)\Delta, \qquad x > x_m,
\end{aligned}$$

where the functions $\sigma(h), m(x)$ and $q(x)$ are determined by equilibrium conditions.

By direct calculation,

$$\begin{aligned}
\left.\frac{\partial \ln\phi(h,x)}{\partial \ln h}\right|_{h=\hat{h}^*(x)} &\approx (1-\varepsilon)\left[1-\varepsilon b_3 m(x)\Delta\right], \qquad (57)\\
\left.\frac{\partial \ln\phi(h,x)}{\partial \ln x}\right|_{h=\hat{h}^*(x)} &\approx \varepsilon\left[1+(1-\varepsilon)\,b_3 m(x)\Delta\right], \qquad x > x_m,
\end{aligned}$$

where $b_3 \equiv 1/\eta - 1 > 0$. Recall from optimization by firms over choice of skill level in (4) that

$$\left.\frac{d\ln\hat{w}(h)}{d\ln h}\right|_{h=\hat{h}^*(x)} = \left.\frac{\partial \ln\phi(h,x)}{\partial \ln h}\right|_{h=\hat{h}^*(x)}, \qquad x > x_m,$$

so

$$h\sigma'(h)|_{h=\hat{h}^*(x)} \approx -\left(1-\varepsilon\right)\varepsilon b_3 m(x), \qquad x > x_m. \qquad (58)$$

In addition, recall from optimal price setting by firms in (5) that

$$\hat{p}(x) = \frac{\rho}{\rho-1}\frac{\hat{w}[\hat{h}^*(x)]}{\phi[\hat{h}^*(x),x]}, \qquad x > x_m,$$

so

$$\begin{aligned}
\frac{d\ln\hat{p}(x)}{d\ln x} &= -\left.\frac{\partial \ln\phi(h,x)}{\partial \ln x}\right|_{h=\hat{h}^*(x)} \\
&\approx -\varepsilon\left[1+(1-\varepsilon)\,b_3 m(x)\Delta\right], \qquad x > x_m. \qquad (59)
\end{aligned}$$

A. The functions m, q

The market clearing conditions for labor and differentiated goods can be used to characterize the functions m and q.

First note that $\hat{\ell}(x)$ satisfies overall clearing in the labor market if and only if

$$\int_{x_m}^{\infty} \hat{\ell}(\xi) f(\xi) d\xi = \frac{L_w}{N_p}, \tag{60}$$

while markets clear for each skill level if and only if the pair $\left(\hat{h}^*, \hat{\ell}\right)$ satisfies

$$\int_{x}^{\infty} \hat{\ell}(\xi) f(\xi) d\xi = \frac{L_w}{N_p} \left\{1 - \hat{\Psi}[\hat{h}^*(x)]\right\}, \qquad x \geq x_m. \tag{61}$$

Use (60) in (61) evaluated at $x = x_m$ to find that $\hat{h}^*(x_m) = \hat{h}_m$. Hence (44) implies $m(x_m) = 1$. Differentiate (61) to get

$$-\hat{\ell}(x) f(x) = -\frac{L_w}{N_p} \hat{\psi}[\hat{h}^*(x)] \frac{\hat{h}^*(x)}{x} \frac{d \ln \hat{h}^*(x)}{d \ln x}, \qquad x \geq x_m,$$

or

$$\begin{aligned} \frac{d \ln \hat{h}^*(x)}{d \ln x} &= [1 + q(x)\Delta] \left(\frac{1 + m(x)\Delta}{1 + \Delta}\right)^{\alpha_h} \\ &\approx 1 + \{q(x) - \alpha_h [1 - m(x)]\} \Delta, \qquad x > x_m, \end{aligned}$$

where the first line uses $\ell_0 = (L_w/N_p)(\alpha_h/\alpha_x) x_m^{\alpha_h - \alpha_x}$, the Pareto densities for f and ψ, and $m(x_m) = 1$. Hence

$$x m'(x) = q(x) - \alpha_h [1 - m(x)], \qquad x > x_m. \tag{62}$$

For q', note that markets in the perturbed economy clear for each differentiated good if and only if

$$y_F N_p^{-\rho\chi} \hat{p}(x)^{-\rho} = \hat{\ell}(x) \phi[\hat{h}^*(x), x], \qquad x \geq x_m.$$

Differentiate this equation and use (59) to get

$$\begin{aligned}\frac{d\ln\hat{\ell}(x)}{d\ln x} &= (\rho-1)\frac{\partial\ln\phi(\hat{h}^*,x)}{\partial\ln x}-\frac{\partial\ln\phi[\hat{h}^*(x),x]}{\partial\ln h}\frac{d\ln\hat{h}^*(x)}{d\ln x}\\ &\approx \alpha_x-\alpha_h+(1-\varepsilon)\left[\rho\varepsilon b_3 m(x)-xm'(x)\right]\Delta, \quad x>x_m,\end{aligned}$$

where the second line uses (57) and $\rho\varepsilon-1=\alpha_x-\alpha_h$, and use (62) to find that

$$xq'(x)\approx(1-\varepsilon)\left\{\rho\varepsilon b_3 m(x)-q(x)+\alpha_h\left[1-m(x)\right]\right\}, \quad x>x_m. \tag{63}$$

Then (62) and (63) are a pair of ODEs for (m,q).

As noted above, $m(x_m)=1$, which provides one boundary conditions. For the other, overall market clearing (60) requires

$$\mathrm{E}_F\left[x^{\alpha_x-\alpha_h}q(x)\right]=x_m^{\alpha_x}\int_{x_m}^{\infty}x^{-\alpha_h}q(x)x^{-1}dx=0. \tag{64}$$

With $m(x_m)=1$, there is a family of solutions to (62) and (63) indexed by the initial value $q(x_m)=q_{m0}$. To see that there exists a unique q_{m0} for which (64) holds, consider the phase diagram for (62) and (63).

The locus in (m,q) space where $m'=0$ is a straight line with slope $-\alpha_h$. The locus where $q'=0$ is a straight line, with slope $\rho\varepsilon b_3-\alpha_h$. The slope can be positive or negative, but in either case it exceeds $-\alpha_h$, and the two lines cross at $(m,q)=(0,\alpha_h)$. The stable arms, paths that converge to $(0,\alpha_h)$, lie between the locus where $m'=0$ and the vertical axis. The solution requires $m(x_m)=1$, so the relevant stable arm lies in the southeast quadrant, to the left of the $m'=0$ locus and hence below the $q'=0$ locus. Along other trajectories with $m(x_m)=1$, $q(x)$ diverges and (64) is violated. Thus $(m,q)\to(0,\alpha_h)$.

Hence m is strictly positive and everywhere decreasing, asymptotically converging to zero, while q is everywhere increasing. For the path that converges, (64) holds, so clearly $q(x)$ takes negative values for x below some threshold and positive values above the threshold.

Thus, the skill allocation $\hat{h}^*(x)$ matches each technology x with a higher skill level. The wedge is a decreasing function of x, and as $x \to \infty$, the wedge declines to zero. Correspondingly, labor is shifted away from lower technologies, toward higher ones. Labor input declines for technologies below some threshold $\hat{x}$, and rise for those above $\hat{x}$.

In the limit as $x \to \infty$, the increase is the one required by the shift in the density function. Since $\lim_{x\to\infty} \hat{h}^*(x) = a_h x$, and

$$\hat{\psi}(h) \approx \psi(h)\,(1+\Delta)^{\alpha_h}\,, \qquad h > \hat{h}_m,$$

it follows that $\lim_{x\to\infty} \hat{\ell}(x) = \ell(x)\,(1+\alpha_h\Delta)\,.$

B. Incentives for process innovation

Write (55) as

$$\begin{aligned} 0 \quad &< \quad \mathrm{E}_F \left\{ \left[\frac{1+\sigma[\hat{h}^*(x)]\Delta}{1+\sigma[\hat{h}^*(x_m)]\Delta} \frac{1+q(x)\Delta}{1+q(x_m)\Delta} - 1 \right] \frac{w\,[h^*(x)]}{w\,[h^*(x_m)]} \frac{\ell(x)}{\ell(x_m)} \right\} \\ &\approx \quad \Delta \mathrm{E}_F \left\{ \left[\sigma[\hat{h}^*(x)] - \sigma[\hat{h}^*(x_m)] + q(x) - q(x_m) \right] \left(\frac{x}{x_m} \right)^{1-\varepsilon+\alpha_x-\alpha_h} \right\}. \end{aligned}$$

Note that

$$\begin{aligned} \sigma[\hat{h}^*(x)] \quad &= \quad \sigma\left\{h^*(x)\,[1+m(x)\Delta]\right\} \\ &\approx \quad \sigma(a_h x) + \sigma'(a_h x) m(x)\Delta. \end{aligned}$$

Dropping terms of order Δ^2, the inequality condition is

$$\begin{aligned} 0 \quad &< \quad \int_{x_m}^{\infty} \left[\sigma(a_h x) - \sigma(a_h x_m) + q(x) - q(x_m)\right] \alpha_x \left(\frac{x}{x_m}\right)^{1-\varepsilon-\alpha_h} x^{-1} dx \\ &= \quad \int_{x_m}^{\infty} \int_{x_m}^{x} \left[a_h \sigma'(a_h \xi) + q'(\xi)\right] d\xi \;\alpha_x \left(\frac{x}{x_m}\right)^{1-\varepsilon-\alpha_h} x^{-1} dx \\ &= \quad \alpha_x \int_{x_m}^{\infty} \int_{\xi}^{\infty} \left(\frac{x}{x_m}\right)^{1-\varepsilon-\alpha_h} x^{-1} dx \;\; \left[a_h \sigma'(a_h \xi) + q'(\xi)\right] d\xi \end{aligned}$$

$$= \frac{\alpha_x}{\alpha_h - 1 + \varepsilon} \int_{x_m}^{\infty} \left(\frac{\xi}{x_m}\right)^{1-\varepsilon-\alpha_h} \left[a_h \xi \sigma'(a_h \xi) + \xi q'(\xi)\right] \xi^{-1} d\xi.$$

Since $a_h \xi = h^*(\xi; 0)$, from (58),

$$a_h \xi \sigma'(a_h \xi) = -(1-\varepsilon)\, \varepsilon b_3 m(\xi).$$

Then use (63) for $\xi q'(\xi)$ to find that the term in brackets in the integral is

$$\begin{aligned} & (1-\varepsilon)\left\{(\rho-1)\,\varepsilon b_3 m(\xi) - q(\xi) + \alpha_h\left[1 - m(\xi)\right]\right\} \\ = \ & (1-\varepsilon)\left[(\rho-1)\,\varepsilon b_3 m(\xi) - \xi m'(\xi)\right], \end{aligned}$$

where the second line uses (62). Since $m(\xi) \in [0,1]$ and $m'(\xi) < 0$, this term is everywhere positive, and the inequality holds.

C. Incentives for retooling

The argument for retooling is similar. Note that

$$\hat{\psi}(h) = (1+\Delta)^{\alpha_h}\, \psi(h), \qquad h > a_h x_m (1+\Delta),$$

and write (56) as

$$\begin{aligned} 0 \ &> \int_{h_m(1+\Delta)}^{\infty} \left[(1+\Delta)^{\alpha_h} \frac{\hat{w}(h)}{\hat{w}(h_m)} - \frac{w(h)}{w(h_m)}\right] \alpha_h \left(\frac{h}{h_m}\right)^{-\alpha_h} h^{-1} dh \\ &\quad - \int_{h_m}^{h_m(1+\Delta)} \frac{w(h)}{w(h_m)} \alpha_h \left(\frac{h}{h_m}\right)^{-\alpha_h} h^{-1} dh \\ &\approx \int_{h_m(1+\Delta)}^{\infty} \left[(1+\alpha_h \Delta) \frac{1+\sigma(h)\Delta}{1+\sigma(h_m)\Delta} - 1\right] \alpha_h \left(\frac{h}{h_m}\right)^{1-\varepsilon-\alpha_h} h^{-1} dh - \alpha_h \Delta \\ &\approx \Delta \int_{h_m}^{\infty} \left[\alpha_h + \sigma(h) - \sigma(h_m)\right] \left(\frac{h}{h_m}\right)^{1-\varepsilon-\alpha_h} h^{-1} dh - \alpha_h \Delta \\ &\approx \Delta \int_{h_m}^{\infty} \left[\sigma(h) - \sigma(h_m)\right] \left(\frac{h}{h_m}\right)^{1-\varepsilon-\alpha_h} h^{-1} dh. \end{aligned}$$

From (58) we see that σ is everywhere decreasing, so the term in brackets is negative and the inequality holds.

Figure 1a: shares of survivors

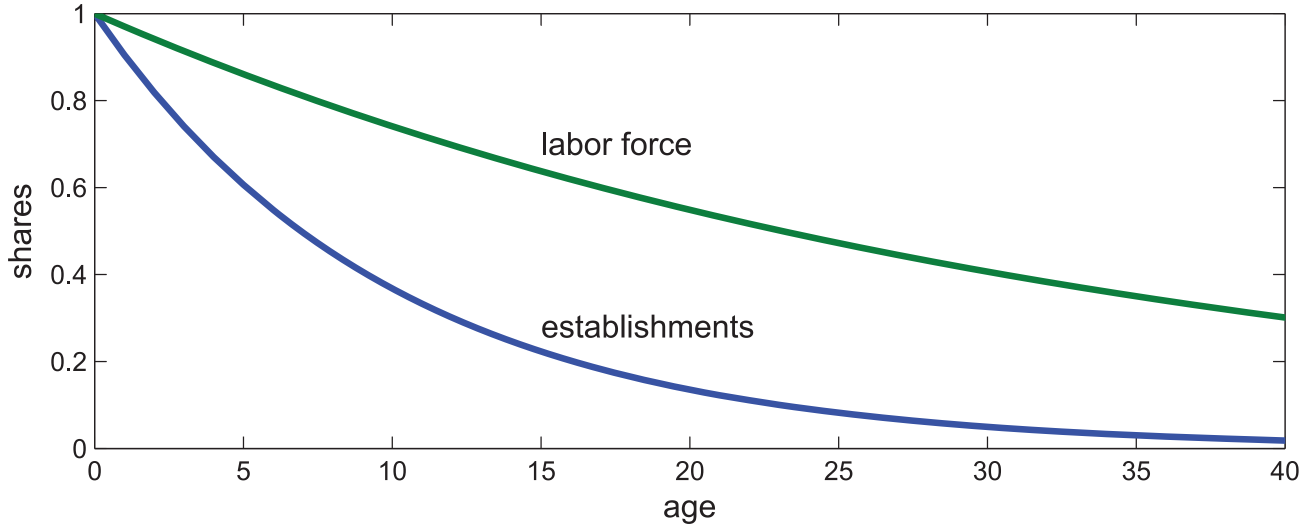

Figure 1b: population shares in surviving labor force

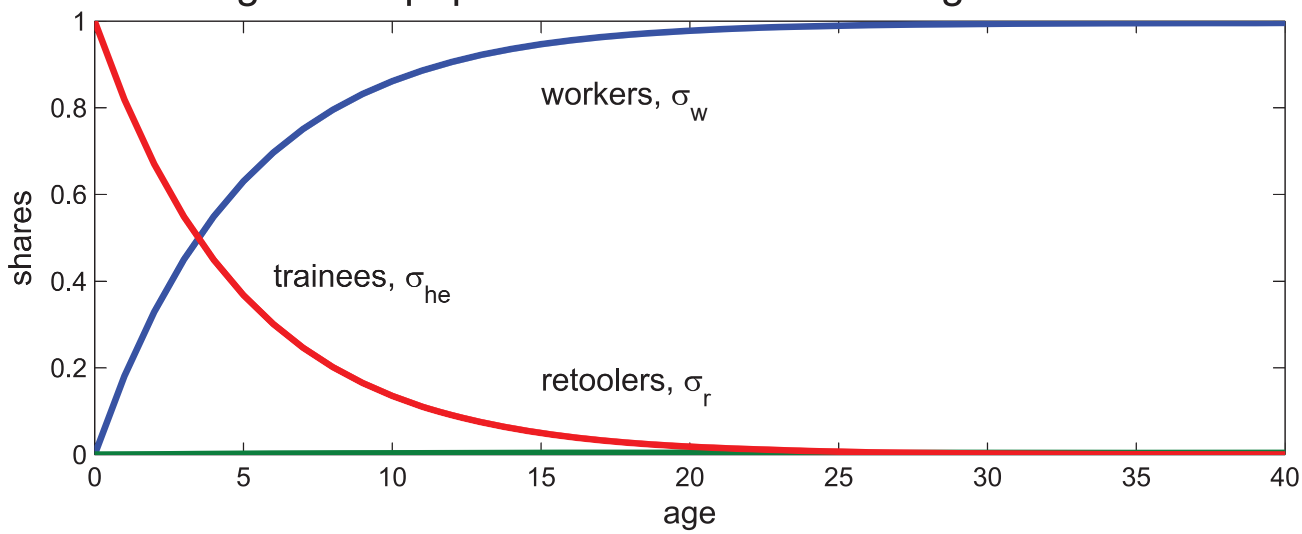

Figure 1c: population shares among surviving establishments

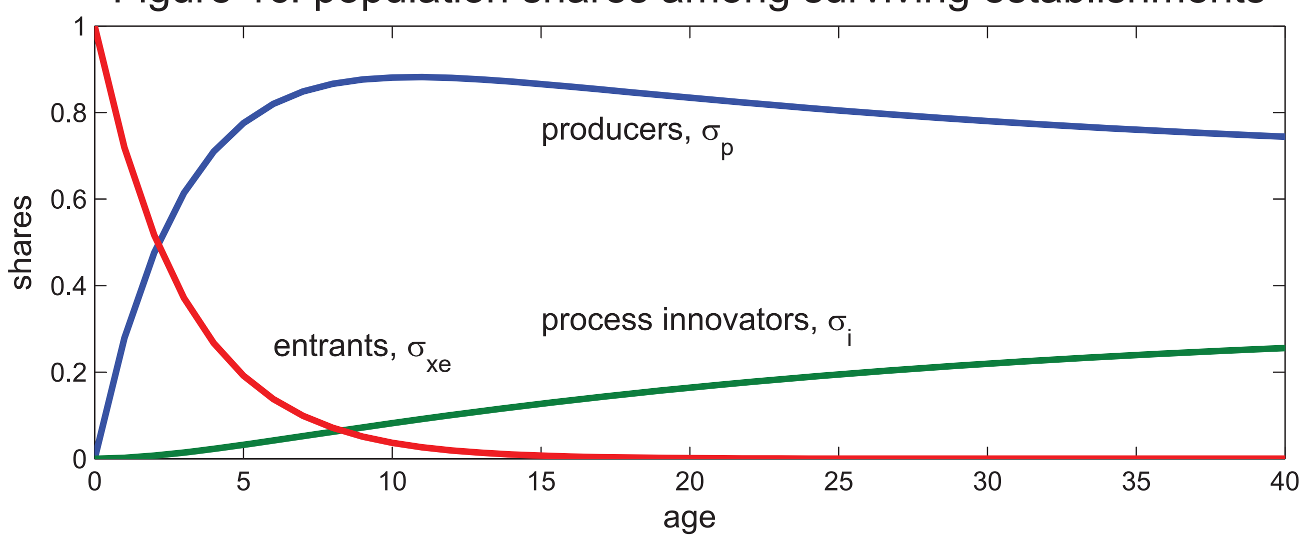

Figure 2a: cohort earnings for workers

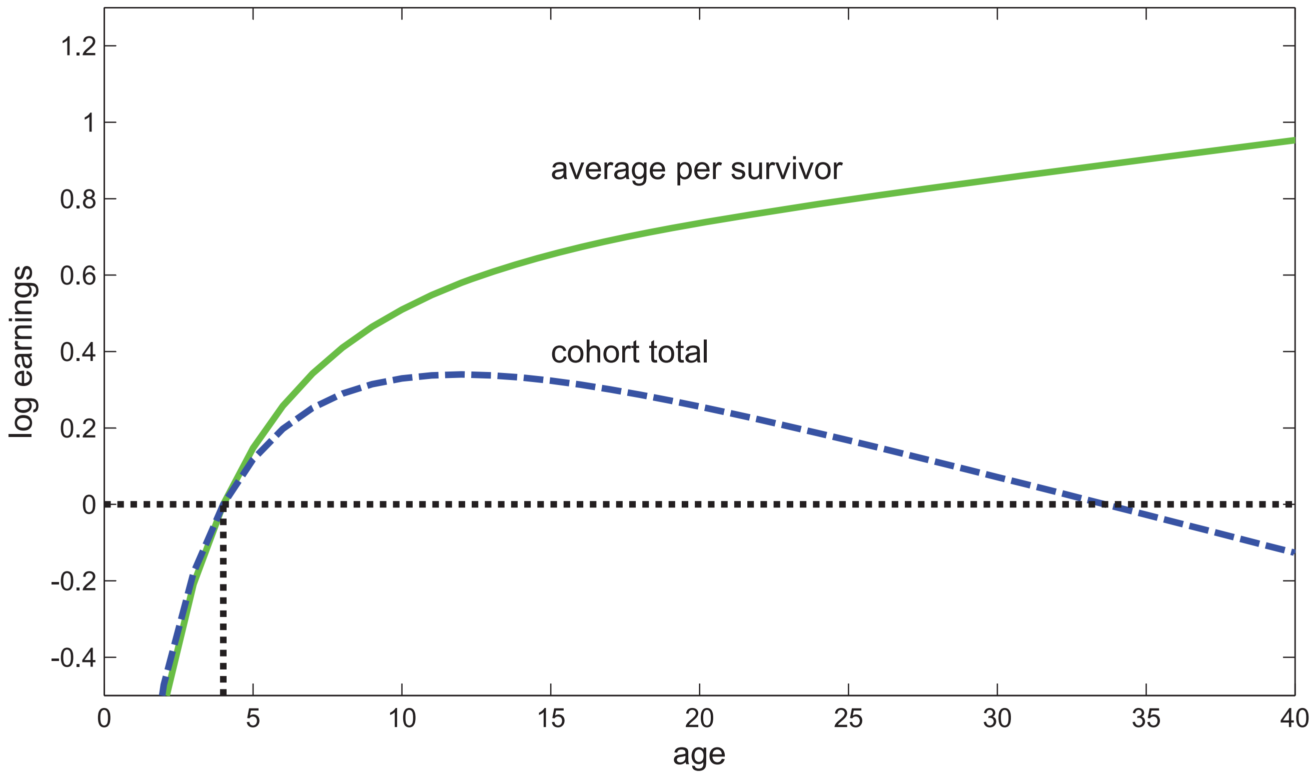

Figure 2b: cohort employment for establishments

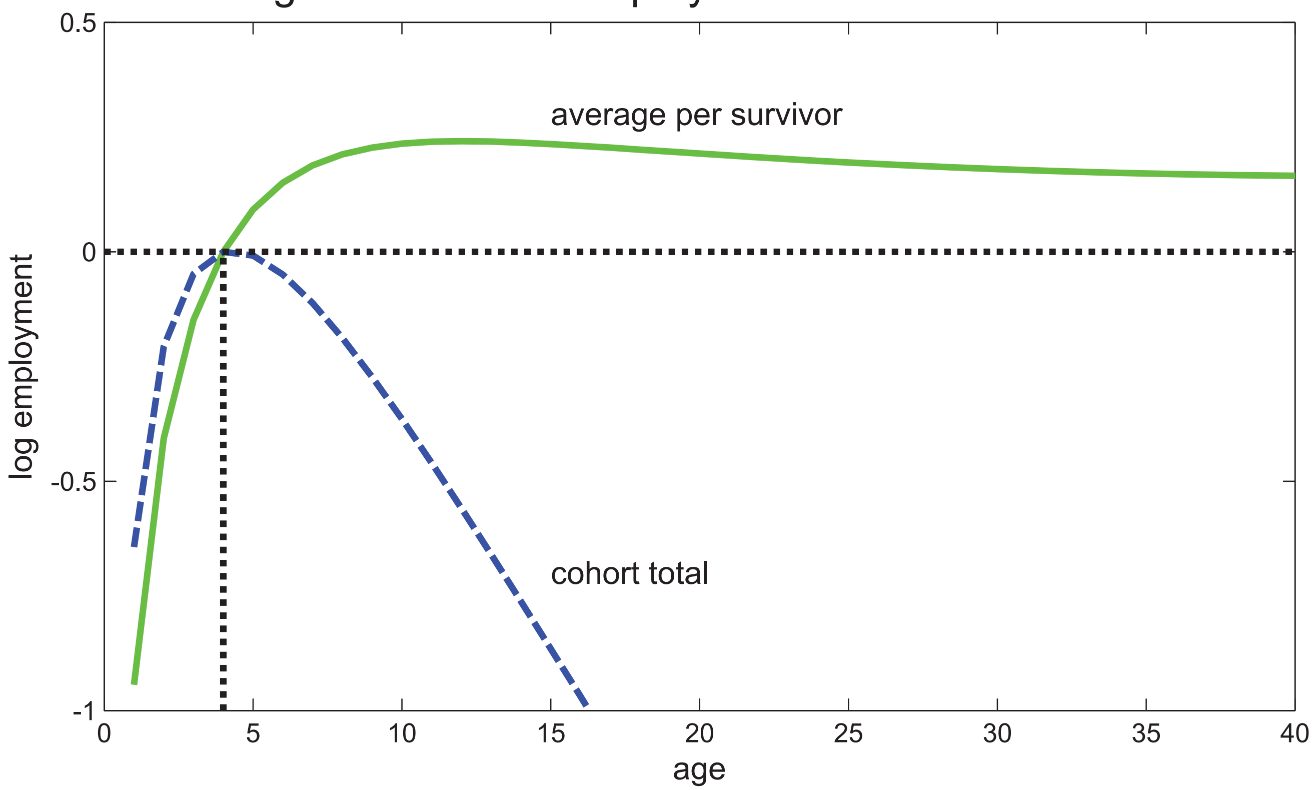

Figure A1a: comparative static (a)

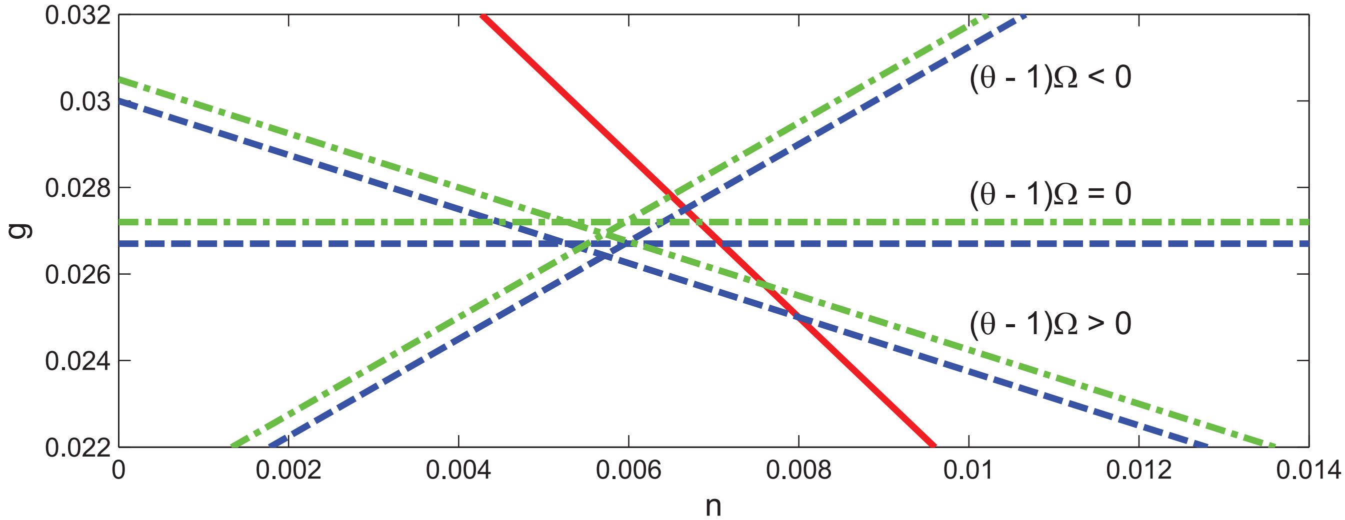

Figure A1b: comparative static (b)

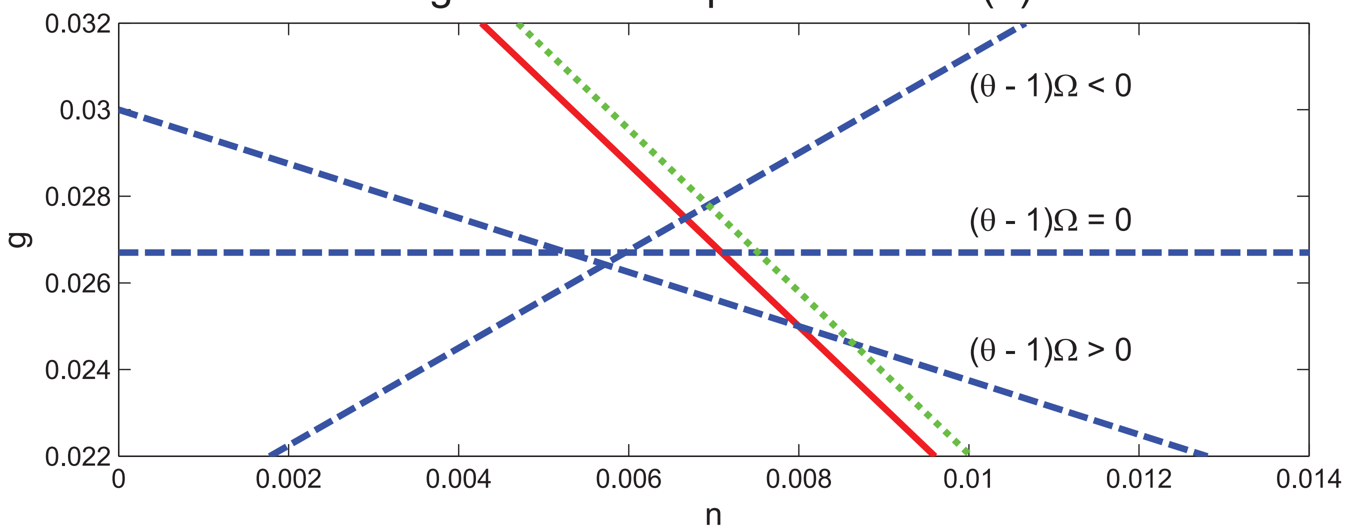

Figure A1c: comparative static (c)