NBER WORKING PAPER SERIES

KNOWLEDGE GROWTH AND THE ALLOCATION OF TIME

Robert E. Lucas, Jr.
Benjamin Moll

Working Paper 17495
http://www.nber.org/papers/w17495

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2011

We thank Roland Benabou, Paco Buera, Olivier Gueant, Gene Grossman, Boyan Jovanovic, Jean-Michel Lasry, Stephen Morris, Ezra Oberfield, Nancy Stokey, and seminar participants at Minnesota, Notre Dame, Princeton, Penn State, NYU, Washington, the Minneapolis Fed, the Chicago Fed, the 2011 SED and the NBER Summer Institute for helpful suggestions and criticism. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Knowledge Growth and the Allocation of Time
Robert E. Lucas, Jr. and Benjamin Moll
NBER Working Paper No. 17495
October 2011
JEL No. O0,O15

ABSTRACT

We analyze a model economy with many agents, each with a different productivity level. Agents divide their time between two activities: producing goods with the production-related knowledge they already have, and interacting with others in search of new, productivity-increasing ideas. These choices jointly determine the economy's current production level and its rate of learning and real growth. Individuals' time allocation decisions depend on the knowledge distribution because the productivity levels of others determine their own chances of improving their productivities through search. The time allocations of everyone in the economy in turn determine the evolution of its knowledge distribution. We construct the balanced growth path for this economy, thereby obtaining a theory of endogenous growth that captures in a tractable way the social nature of knowledge creation. We also study the allocation chosen by an idealized planner who takes into account and internalizes the external benefits of search, and tax structures that implement an optimal solution. Finally, we provide two examples of alternative learning technologies, as concrete illustrations of other directions that might be pursued.

Robert E. Lucas, Jr.
Department of Economics
The University of Chicago
1126 East 59th Street
Chicago, IL 60637
and NBER
relucas@midway.uchicago.edu

Benjamin Moll
106 Fisher Hall Department of Economics
Princeton University
Princeton, NJ 08544
moll@princeton.edu

1 Introduction

We analyze a new model of endogenous growth, driven by sustained improvements in individual knowledge. Agents in this economy divide their time between two activities: producing goods with the production-related knowledge they already have, and interacting with others in search of new, productivity-increasing ideas.[1] These choices jointly determine the economy's current production level and its rate of learning and real growth.

The production technology in this model is kept very simple: Each person produces at a rate that is the product of his personal productivity level and the fraction of time that he chooses to spend producing goods. There are no factors of production other than labor and there are no complementarities between workers. There are no markets, no prices, and no public or private property other than individuals' knowledge—their human capital. The learning technology involves random meetings: Each person meets others at a rate that depends on the fraction of time he spends in search. For us, a meeting means simply an observation of someone else's productivity. If that productivity is higher than his own, he adopts it in place of the productivity he came in with. Everyone's productivity level is simply the maximum of the productivities of all the people he has ever met. To ensure that the growth generated by this process can be sustained, we add an assumption to the effect that the stock of good ideas waiting to be discovered is inexhaustible.[2]

The state of the economy is completely described by the distribution of productivity levels. An individual's time allocation decisions will depend on this distribution because the productivity levels of others determine his own chances of improving his productivity through search. Individuals' time allocation decisions in turn determine learning rates and thus the evolution of the productivity distribution. One of the two equilibrium conditions of the model is the Bellman equation for the time allocation problem of a single atomistic agent who takes the productivity distribution as given. The second condition is a law of motion for the productivity distribution, given the policy functions of individual agents.

These two equations take the form of partial differential equations, with time and productivity levels as the two independent variables. We motivate these two equations in the next section. Then we focus on a particular solution of these equations, a balanced growth path,

[1] Our tradeoffs at an individual level will be taken with little change from the models of on-the-job learning of Ben-Porath (1967), Heckman (1976), and Rosen (1976).

[2] Jovanovic and Rob (1989) describe a similar process of adoption or imitation of ideas. In their model, however, the knowledge distribution converges to a steady state whereas our framework features sustained growth. For more recent contributions that are similar in spirit to ours, see Bental and Peled (1996) and particularly Perla and Tonetti (2011). A different approach to a similar set of questions as in our paper is pursued by Fogli and Veldkamp (2011) who model the diffusion of knowledge among individuals in a network.

along which production grows at a constant rate and the distribution of relative productivities remains constant. In Section 3 we discuss the properties of this balanced growth path and develop an algorithm to calculate it, given parameters that describe the production and search technologies.

There is an evident external effect in this decentralized equilibrium. The private return to knowledge acquisition motivates individual decisions that generate sustained productivity growth but an individual agent does not take into account the fact that increases in his own knowledge enrich the learning environment for the people around him. The social return to search exceeds the private return, raising the possibility that taxes and subsidies can equate private and social returns and improve both growth rates and welfare.[3] In Section 4 we formulate a planning problem, in which the planner directs the time allocations of each of the continuum of individual agents in the economy. We show how this problem can be broken into individual Bellman equations where the value function for each person is his marginal *social* value under an optimal plan. We study the implied balanced growth path and compare the implied policy function and distribution of relative productivities to those implied by the decentralized problem studied in Section 3. In Section 5 we consider the implementation of the planning solution through the use of a Pigovian system of taxes and subsidies.

All of the analysis in Sections 3-5 is based on a single, specific model of the search/learning process. It turns out that the algorithm we develop for this model is quite easily adapted to the analysis of a wide variety of other learning technologies. In Section 6 we describe two of these alternative technologies and consider their implications and economic interpretations. Section 7 concludes the paper.

2 A Model Economy

There is a constant population of infinitely-lived agents. We identify each person at each date as a realization of a draw $\tilde{z}$ from a cost distribution, described by its cdf

$$F(z,t) = \Pr\{\tilde{z} \leq z \text{ at date } t\},$$

or equivalently by its density function $f(z,t)$. This function $f(\cdot,t)$ fully describes the state of the economy at t. A person with cost draw $\tilde{z}$ can produce $\tilde{a} = \tilde{z}^{-\theta}$ units of a single consumption good, where $\theta \in (0,1)$.

[3]In this respect the model is a direct descendant of Arrow (1962), Romer (1986), and Grossman and Helpman (1991).

We will formulate the equilibrium and planning problems of this economy in terms of this cost distribution but of course we could instead do this in terms of the productivity distribution $G(a,t)$:

$$G(a,t) = \Pr\{\tilde{z}^{-\theta} \leq a\} = \Pr\{\tilde{z} \geq a^{-1/\theta}\} = 1 - F(a^{-1/\theta}, t).$$

For some purposes the economic interpretations seem clearer in this form, but algorithmically the cost formulation is more convenient. We will find it useful to use both of them on occasion. Here we continue with the cost formulation.

Every person has one unit of labor per year. He allocates his time between a fraction $1 - s(z,t)$ devoted to goods production and $s(z,t)$ devoted to improving his production-related knowledge. His goods production is

$$[1 - s(z,t)]\, z^{-\theta}. \tag{1}$$

Total production in the economy is

$$Y(t) = \int_0^\infty [1 - s(z,t)]\, z^{-\theta} f(z,t) dz.$$

Individual preferences are

$$V(z,t) = \mathbb{E}_t \left\{ \int_t^\infty e^{-\rho(\tau - t)} [1 - s(\tilde{z}(\tau), \tau)]\, \tilde{z}(\tau)^{-\theta} d\tau \middle| z(t) = z \right\}. \tag{2}$$

We model the evolution of the distribution $f(z,t)$ as a process of individuals meeting others from the same economy, comparing ideas, improving their own productivity. The details of this meeting and learning process are as follows.[4] A person z allocating the fraction $s(z,t)$ to learning observes the cost z' of one other person with probability $\alpha\,[s(z,t)]\,\Delta$ over an interval $(t, t+\Delta)$, where α is a given function. He compares his own cost level z with the cost z' of the person he meets, and leaves the meeting with the best of the two costs, $\min(z, z')$. (These meetings are not assumed to be symmetric: z learns from and perhaps imitates z' but z' does not learn from z and in fact he may not be searching himself at all.)

We assume that everyone in the economy behaves in this way, though the search effort $s(z,t)$ varies over time and across individuals at a point in time. Thinking of $F(z,t)$ as the fraction of people with cost below z at date t, this behavior results in a law of motion for F as

[4] The process assumed here is an adaptation of ideas in Kortum (1997), Eaton and Kortum (1999), Alvarez, Buera and Lucas (2008), and Lucas (2009).

follows:

$$\begin{aligned}
1 - F(z, t+\Delta) &= \Pr\{\text{cost above } z \text{ at } t \text{ and no lower cost found in } [t, t+\Delta)\} \\
&= \int_z^\infty f(y,t) \Pr\{\text{no lower cost found in } [t, t+\Delta)\} dy \\
&= \int_z^\infty f(y,t) \left[1 - \alpha(s(y,t))\Delta + \alpha(s(y,t))\Delta \left(1 - F(z,t)\right)\right] dy \\
&= 1 - F(z,t) - F(z,t) \int_z^\infty \alpha(s(y,t)) f(y,t) \Delta dy.
\end{aligned}$$

Then

$$\frac{F(z, t+\Delta) - F(z,t)}{\Delta} = F(z,t) \int_z^\infty \alpha(s(y,t)) f(y,t) dy$$

and letting $\Delta \to 0$ gives

$$\frac{\partial F(z,t)}{\partial t} = F(z,t) \int_z^\infty \alpha(s(y,t)) f(y,t) dy.$$

Differentiating with respect to z we obtain

$$\frac{\partial f(z,t)}{\partial t} = -\alpha(s(z,t)) f(z,t) \int_0^z f(y,t) dy + f(z,t) \int_z^\infty \alpha(s(y,t)) f(y,t) dy. \tag{3}$$

Equation (3) can also be motivated by considering the evolution of the density at z directly, as follows. Some agents who have cost z will adopt a lower cost $y \leq z$ and so there will be an outflow of these agents. Other agents who have cost $y \geq z$ will adopt cost z and there will be an inflow of these agents. Hence we can write

$$\frac{\partial f(z,t)}{\partial t} = \left.\frac{\partial f(z,t)}{\partial t}\right|_{\text{out}} + \left.\frac{\partial f(z,t)}{\partial t}\right|_{\text{in}}.$$

Consider first the outflow. The $f(z,t)$ agents at z have meetings at the rate $\alpha(s(z,t)) f(z,t)$. A fraction $F(z,t)$ of these draws satisfy $y < z$ and these agents leave z. Hence

$$\left.\frac{\partial f(z,t)}{\partial t}\right|_{\text{out}} = -\alpha(s(z,t)) F(z,t) f(z,t).$$

Next, consider the inflow. Agents with cost $y \geq z$ have meetings at the rate $\alpha(s(y,t)) f(y,t)$. Each of these meetings yields a draw z with probability $f(z,t)$. Hence

$$\left.\frac{\partial f(z,t)}{\partial t}\right|_{\text{in}} = f(z,t) \int_z^\infty \alpha(s(y,t)) f(y,t) dy.$$

Combining, we obtain (3). This type of equation is known in physics as a Boltzmann equation.

Now consider the behavior of a single agent with current cost z, acting in an environment characterized by a given density path $f(z,t)$, all $z, t \geq 0$. The agent wants to choose a policy $s(z,t)$ so as to maximize the discounted, expected value of his earnings stream, expression (2). The Bellman equation for this problem is[5]

$$\rho V(z,t) = \max_{s \in [0,1]} \left\{ (1-s)z^{-\theta} + \frac{\partial V(z,t)}{\partial t} + \alpha(s) \int_0^z [V(y,t) - V(z,t)] f(y,t) dy \right\}. \tag{4}$$

The system (3) and (4) is an instance of what Lasry and Lions (2007) have called a "mean-field game." We summarize our discussion of the economy in the

Definition: An *equilibrium,* given the initial distribution $f(z,0)$, is a triple (f, s, V) of functions on $\mathbf{R}_+^2$ such that (i) given s, f satisfies (3) for all (z,t), (ii) given f, V satisfies (4), and (iii) $s(z,t)$ attains the maximum for all (z,t).

A complete analysis of this economy would require the ability to calculate solutions for all initial distributions. This would be an economically useful project to carry out, but we limit ourselves in this paper to the analysis of a set of particular solutions on which the growth rate and the distribution of relative costs are both constant over time.

Definition: A *balanced growth path (BGP)* is a number γ and a triple of functions (ϕ, σ, v) on $\mathbf{R}_+$ such that

$$f(z,t) = e^{\gamma t} \phi(z e^{\gamma t}), \tag{5}$$

$$V(z,t) = e^{\theta \gamma t} v(z e^{\gamma t}), \tag{6}$$

and

$$s(z,t) = \sigma(z e^{\gamma t}) \tag{7}$$

for all (z,t), and (f, s, V) is an equilibrium with the initial condition $f(z,0) = \phi(z)$.

Intuitively, a BGP is simply a path for the distribution function along which all cost quantiles shrink at the same rate γ (and hence all quantiles of productivity, $z^{-\theta}$, grow at rate $\theta\gamma$). That is, on a BGP the cost cdf satisfies $F(z,t) = \Phi(z e^{\gamma t})$ and therefore the qth quantile, $z_q(t)$, satisfies $\Phi(z_q(t) e^{\gamma t}) = q$ or

$$z_q(t) = e^{-\gamma t} \Phi^{-1}(q).$$

That the value and policy functions take the forms in (6) and (7) is then immediately implied.

The analysis of balanced growth is facilitated by restating (3) and (4) in terms of relative

[5] See Appendix A for a derivation of this continuous time Bellman equation as a limit of the corresponding discrete time version.

costs $x = ze^{\gamma t}$. From (5), we have

$$\frac{\partial f(z,t)}{\partial t} = \gamma e^{\gamma t}\phi(ze^{\gamma t}) + e^{\gamma t}\phi'(ze^{\gamma t})\gamma ze^{\gamma t}$$

which from (3) and (7) implies

$$\phi(x)\gamma + \phi'(x)\gamma x = \phi(x)\int_x^\infty \alpha(\sigma(y))\phi(y)dy - \alpha(\sigma(x))\phi(x)\int_0^x \phi(y)dy. \tag{8}$$

Evaluating at $x = 0$, we have

$$\phi(0)\gamma = \phi(0)\int_0^\infty \alpha(\sigma(y))\phi(y)dy. \tag{9}$$

The Bellman equation (4) becomes

$$(\rho - \theta\gamma)\, v(x) - v'(x)\gamma x = \max_{\sigma\in[0,1]}\left\{(1-\sigma)x^{-\theta} + \alpha(\sigma)\int_0^x [v(y) - v(x)]\phi\,(y)\,dy\right\}. \tag{10}$$

Total production on a balanced growth path is

$$Y(t) = e^{\theta\gamma t}\int_0^\infty [1 - \sigma(x)]x^{-\theta}\phi(x)dx, \tag{11}$$

provided the integral converges. Hence total production grows at the rate $\theta\gamma$.

If all agents in this economy had the same cost level $\bar{z}$, say, then no one would have any motive to search and everyone would simply produce $\bar{z}^{-\theta}$ forever. Such a trivial equilibrium could be called a BGP with $\gamma = 0$, but our interest is in BGPs with $\gamma > 0$. To ensure that this is a possibility we will need to add more structure. For this purpose, we add the assumption that $f(0,0) \equiv \lim_{z\to 0} f(z,0) > 0$, implying that on a BGP $\phi(0) > 0$. This condition is sufficient to ensure that sustained growth at some rate $\gamma > 0$ is possible. Its interpretation is that the stock of good ideas waiting to be discovered is inexhaustible. The next result shows that this restriction is equivalent to the assumption that the initial distribution of productivity has a Pareto tail with tail parameter $1/\theta$.

Lemma 1: The initial cdf of productivity, $G(a,0)$ say, has a Pareto tail,

$$\lim_{a\to\infty}\frac{1 - G(a,0)}{a^{-1/\theta}} = \lambda \qquad \text{for some } \lambda > 0,$$

if and only if $z = a^{-1/\theta}$ satisfies $f(0,0) = \lambda > 0$.

Proof: We have that $G(a,0) = 1 - F(a^{-1/\theta}, 0)$ and therefore

$$\lim_{a\to\infty} \frac{1 - G(a,0)}{a^{-1/\theta}} = \lim_{a\to\infty} \frac{F(a^{-1/\theta}, 0)}{a^{-1/\theta}} = \lim_{z\to 0} \frac{F(z,0)}{z} = f(0,0) = \lambda.\square$$

Thus requiring that $f(0,0) > 0$ is the same thing as assuming a fat tailed initial productivity distribution and the parameter θ has the interpretation as the inverse of the tail parameter.

Under the restriction $f(0,0) > 0$, $\phi(0) > 0$ and (9) imply that γ will be an average of the search intensities α at different cost levels x:

$$\gamma = \int_0^\infty \alpha(\sigma(x))\phi(x)dx. \tag{12}$$

We also assume that the learning technology function $\alpha : [0,1] \to \mathbf{R}_+$ satisfies

$$\alpha(s) \geq 0,\ \alpha'(s) > 0,\ \alpha''(s) < 0, \quad \text{all } s,$$

and

$$\alpha(1) > 0,\ \alpha'(1) > 0,\ \lim_{s\to 0} \alpha'(s) = \infty. \tag{13}$$

The discount rate ρ satisfies

$$\rho \geq \theta\alpha(1). \tag{14}$$

This will ensure that the preferences in (2) are well-defined.

3 Calculation and Analysis of Balanced Growth Paths

In this section we describe the algorithm we use to calculate BGPs, given a specified function α, values for the parameters ρ and θ, and a value $\lambda = \phi(0)$ for the density at $x = 0$.

We begin an iteration with initial guesses (ϕ_0, γ_0) for (ϕ, γ). Then for $n = 0, 1, 2, ...$ we follow

Step 1. Given (ϕ_n, γ_n), use (10) to calculate v_n and σ_n.

Step 2. Given σ_n, solve (8) and (12) jointly to generate a new guess $(\phi_{n+1}, \gamma_{n+1})$.

When these steps are completed, $(\phi_{n+1}, \gamma_{n+1})$ and (v_n, σ_n) have been calculated. When $(\phi_{n+1}, \gamma_{n+1})$ is close enough to (ϕ_n, γ_n), we call $(\phi_n, \gamma_n, v_n, \sigma_n)$ a BGP equilibrium. Steps 1 and 2 themselves involve iterative procedures which we describe in turn.

For step 1, consider the Bellman equation (10). Define the function

$$S(x) = \int_0^x [v(y) - v(x)]\phi\,(y)\,dy.$$

Then the first order condition for σ is

$$S(x)\alpha'(\sigma) \geq x^{-\theta} \quad \text{with equality if } \ \sigma < 1. \tag{15}$$

Under our assumptions on α, this condition can be solved for a unique $\sigma(x) \in (0, 1]$, that satisfies $\sigma'(x) > 0$ as long as $\sigma(x) < 1$. There will be a unique value $\hat{x}$ that satisfies

$$\alpha'(1) = \frac{\hat{x}^{-\theta}}{S(\hat{x})}.$$

Agents with relative costs x below $\hat{x}$ will divide their time between producing and searching; agents at or above $\hat{x}$ will be searching full time. For $x \geq \hat{x}$, $v(x)$ is constant at $v\,(\hat{x})$ and thus $S(x)$ is constant at $S(\hat{x})$. The value function v will satisfy $v(x) > 0$, $v'(x) \leq 0$, $\lim_{x\to 0} v(x) = \infty$ and

$$\lim_{x\to\infty} v'(x) = 0. \tag{16}$$

The last condition motivates a boundary condition for the integro-differential equation (10). All these conclusions hold for any density ϕ and $\gamma > 0$.

The computation of (v_n, σ_n) given (ϕ_n, γ_n), follows itself an iterative procedure. We begin an iteration with an initial guess v_n^0 for v_n.[6] Then for $j = 0, 1, 2, ...$ we follow

Step 1a. Given $v_n^j(x)$, compute $S_n^j(x)$ from (3) and $\sigma_n^j(x)$ from (15).

Step 1b. Given $\sigma_n^j(x)$, solve (10) together with the boundary condition (16) for $v_n^{j+1}(x)$. To carry out these calculations, we applied a finite difference method on a grid $(x_1, x_2, ..., x_I)$ of I values. Details are provided in Appendix C.1.

When v_n^{j+1} and v_n^j are sufficiently close, we set $(v_n, \sigma_n) = (v_n^j, \sigma_n^j)$. This completes step 1.

For step 2, we express (8) as

$$\phi(x)\gamma + \phi'(x)\gamma x = \phi(x)\psi(x) - \alpha(\sigma(x))\phi(x)\Phi(x) \tag{17}$$

[6] We use $v_n^0(x) = x^{-\theta}/(\rho - \theta\gamma_n)$.

where ψ and Φ are defined by

$$\psi(x) = \int_x^\infty \alpha(\sigma(y))\phi(y)dy \quad \text{and} \quad \Phi(x) = \int_0^x \phi(y)dy \ .$$

Then

$$\psi'(x) = -\alpha(\sigma(x))\phi(x) \tag{18}$$

$$\Phi'(x) = \phi(x) \tag{19}$$

We further have $\phi(0) = \lambda$, $\Phi(0) = 0$. Finally, equation (12) can be written as $\gamma = \psi(0)$. The computation of $(\phi_{n+1}, \gamma_{n+1})$ given (v_n, σ_n) again follows an iterative procedure. We begin an iteration with an initial guess γ_{n+1}^0 for γ_{n+1}. Then for $j = 0, 1, 2, ...$ we follow

Step 2a. Given γ_{n+1}^j and σ_n, solve for functions $\phi_{n+1}^j(x), \Phi_{n+1}^j(x), \psi_{n+1}^j(x)$ by solving the system of ODEs (17) to (19) with boundary conditions

$$\phi_{n+1}^j(0) = \lambda, \quad \Phi_{n+1}^j(0) = 0, \quad \psi_{n+1}^j(0) = \gamma_{n+1}^j.$$

We again use a finite difference method with details provided in Appendix C.3.

Step 2b. Given ϕ_{n+1}^j, γ_{n+1}^j and σ_n, update

$$\gamma_{n+1}^{j+1} = \xi \int_0^\infty \alpha(\sigma_n(x))\phi_{n+1}^j(x)dx + (1-\xi)\gamma_{n+1}^j$$

where $\xi \in (0, 1]$ is a relaxation parameter.

When γ_{n+1}^{j+1} and γ_{n+1}^j are sufficiently close, we set $(\phi_n, \gamma_n) = (\phi_n^j, \gamma_n^j)$. This completes step 2. For the initial guess we use an exponential with parameter λ, $\phi_0(x) = \lambda \exp(-\lambda x)$, and a growth rate $\gamma_0 = \alpha(1)$. For the function α we used

$$\alpha(s) = ks^\eta, \quad \eta \in (0, 1).$$

The computational procedure is outlined in more detail in the Appendix C.

The mathematics of each of the steps just described, the solution to a Bellman equation, the solution to an ordinary differential equation with given boundary questions, and the solution to a fixed point problem in the growth parameter γ, are all well understood. We have not been able to establish the existence or uniqueness of a BGP with $\gamma > 0$, but the algorithm we have described calculates solutions to a high degree of accuracy for the exponential initial cost density that we use as an initial guess and a variety of reasonable parameter values.

Figures 1-4 report the results of one simulation of this model, and provide some information on the sensitivity of the policy function to changes in parameters. Figure 5 provides some typical sample paths, to illustrate the kind of changes over time an individual's choices and earnings will exhibit along the BGP we have computed. The figures are intended to illustrate the qualitative properties of the model, and the calibration of parameters will depend on the application and available data. But there is a good deal of closely related research that uses time series on aggregate growth rates and cross-section data on individual agents to estimate parameters related to our parameters θ and η and it will be useful to describe how the numbers we use are related to this evidence.

The growth rate of per capita GDP in the United States and other OECD countries has fluctuated around two percent at least since World War II. This fact supports the application of models that have a BGP equilibrium and suggest the value .02 for the product $\theta\gamma$. The parameter θ has interpretations both as a log variance parameter or as a tail parameter. Thinking of agents in the model as individual workers as we have done, suggests using the variance of log earnings to estimate θ. Lucas (2009), using a model with constant search effort, finds $\theta = 0.5$ to be consistent with U.S. census earnings data. Gabaix (2009), Luttmer (forthcoming), and others who identify agents (in our sense) with firms estimates $\theta = 1$ (Zipf's Law)as a good tail parameter based on the size distribution of firms. Eaton and Kortum (2002) associate costs of any specific good with an entire country, and obtain estimates of θ less than one, using international relative prices. Here we use the value $\theta = 0.5$; results for $\theta = 0.7$ are also shown in Figure 1. Then given a choice of θ and a value for the parameter η, we can choose the constant k so that $\gamma = (.02)/\theta$.

None of the studies cited above provides evidence on η, which measures the elasticity of search intensity with respect to the time spent searching. To obtain information on η we need evidence on the technology of on-the-job human capital accumulation, such as that used by Ben-Porath (1967), Rosen (1976), Heckman (1976) and Hause (1980).[7] Rosen (1976) used a parameter similar to our η. He assigned the value $\eta = 0.5$, in part to get a functional form that was easy to work with. We used $\eta = 0.3$. Perla and Tonetti (2011) use a model similar to ours in which $\alpha(s)$ is linear in s, so that workers work full time above a productivity threshold and search full time otherwise. Our model approaches this situation as $\eta = 1$, although the Perla and Tonetti model is not a special case of ours. See Figure 4 for experiments at η values 0.3, 0.6, and 0.9.

[7] Ben-Porath and Rosen suggested that any particular human capital path could be interpreted as a property of an occupation, in which case one could view a person's time allocation choices as implied by an initial, one-time occupational choice. This appealing interpretation is open to us as well, as long as the path is interpreted as a productivity-contingent stochastic process.

Figure 1 plots the equilibrium time allocation function, $\sigma(x)$, against relative productivity levels, $x^{-\theta}$ for the two θ values 0.5 and 0.7. The units on the productivity axis are arbitrary. We normalized productivity by dividing by median productivity for each value of θ. A higher θ value (higher variance, fatter tail) induces a higher return to search. At either θ value the least productive people search full time; the most productive work almost full time.

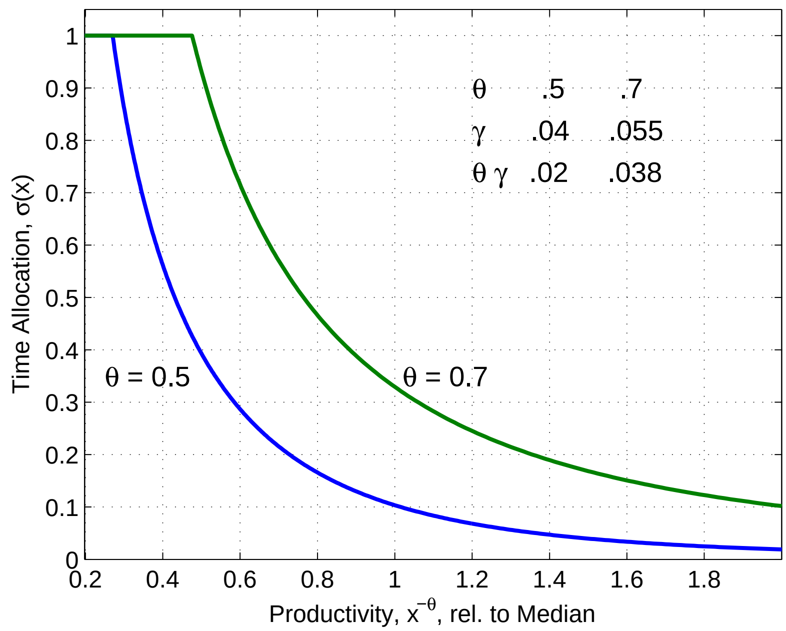

Figure 1: Optimal Time Allocation, $\sigma(x)$, for $\theta = 0.5$ and $\theta = 0.7$

Figure 2 plots the productivity density for $\theta = 0.5$, superimposed on a plot of a Pareto density with tail parameter $1/\theta = 2$. The two curves coincide for large productivity levels. Again, units are relative to the median value under the equilibrium density.

Figure 3 plots two equilibrium Lorenz curves for the same case $\theta = 0.5$. The curve furthest from the diagonal (the one with the most inequality) plots the fraction of current production $(1 - \sigma(x))x^{-\theta}$ attributed to workers with productivity less than $x^{-\theta}$. This is the standard income flow Lorenz curve. The other curve, the one with less inequality, plots the fraction of total discounted expected earnings $v(x)$ accounted for by people with current productivity less than $x^{-\theta}$. Here $v(x)$ is the value function calculated in our algorithm. This value Lorenz curve takes into account the effects of mobility along with the effect of current productivity. In dynamic problems such as the one we study, it will be more informative to examine present value rather than flow Lorenz curves.

Figure 4 plots the time allocation functions for three η values with θ set at 0.5. The $\eta = 0.3$ curve coincides with the $\theta = 0.5$ curve in Figure 1.

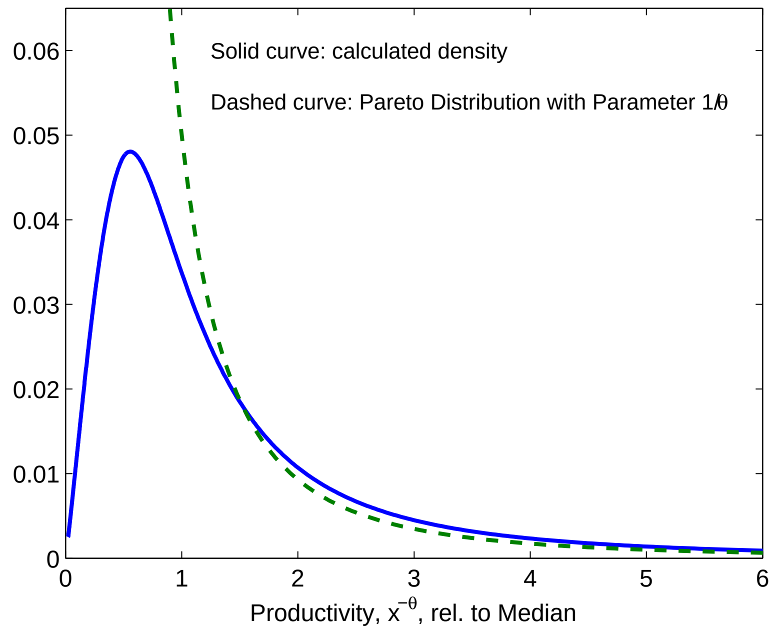

Figure 2: Productivity Density for $\theta = 0.5$

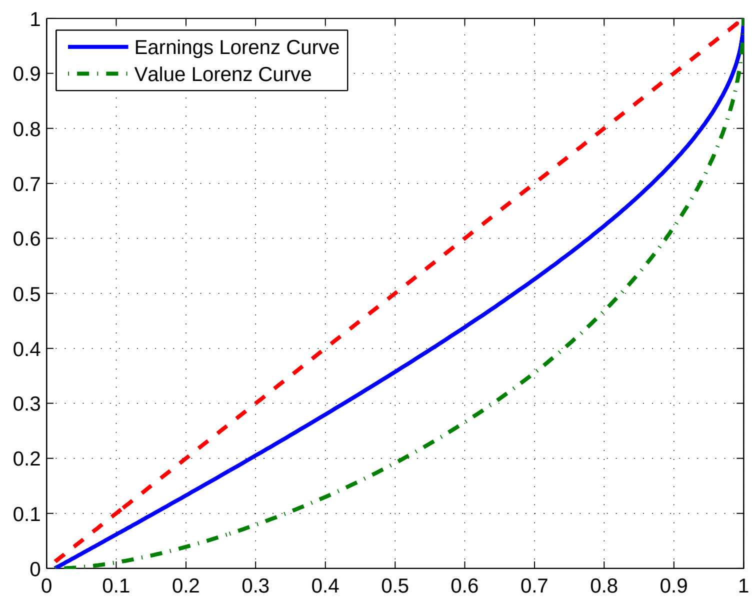

Figure 3: Earnings and Value Lorenz Curves for $\theta = 0.5$

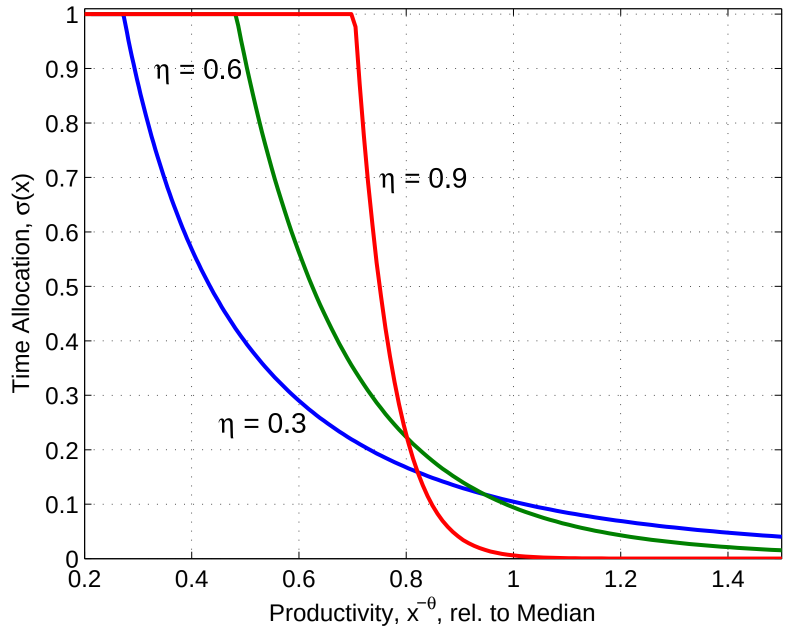

Figure 4: Optimal Time Allocation, for various η values.

Figure 5 shows various aspects of two randomly generated sample paths, which we call stochastic careers. Agents in our model are infinitely-lived. A particular productivity sample path will never decrease—knowledge in our model is never lost—but on a BGP relative productivities $x^{-\theta}$ will wander forever with long run averages described by the cdf $\Phi(x)$. This means, for example, that every sample path will be in the interval $[\hat{x}, \infty)$ for a fraction $1-\Phi(\hat{x})$ of his career, where $\hat{x}^{-\theta}$ is the productivity level (defined in Section 3) below which it is not worthwhile to work. He will return to $[\hat{x}, \infty)$ infinitely often. (We can make the same statement about *any* x value but $\hat{x}$ is chosen here for a reason.) We can get a good sense of an individual career by thinking of each return to $[\hat{x}, \infty)$ as a death or retirement, where the departing worker is replaced by a new potential worker who begins with some productivity $x_0^{-\theta} \leq \hat{x}^{-\theta}$. Like a school child, this entrant starts with some work-relevant knowledge and can begin to acquire more right away, but it may be some time before his knowledge level has a market value. In the same way, some older workers, even those with successful careers in their past, will find that the market value of their accumulated knowledge has fallen to zero, not because they forget what they once knew but because the number of others who know more has grown.

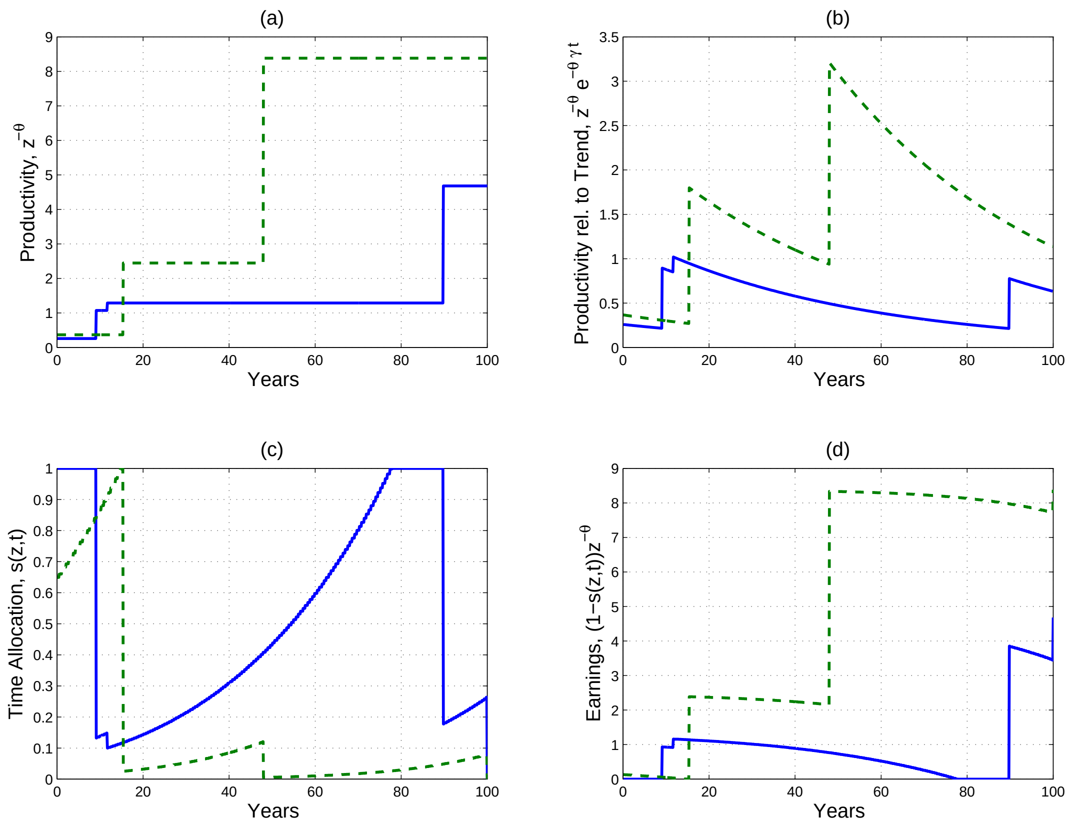

Figure 5: Two Stochastic Careers.

4 An Optimally Planned Economy

Neither the equilibrium conditions (3) and (4) for the decentralized economy nor their BGP counterparts describe an economically efficient allocation. Each agent allocates his time to maximize his own present value, but assigns no value to the benefits that increasing his knowledge will have for others. Yet we are studying an economy where learning from others is the sole engine of technological change.

In this section, we ask how a hypothetical, benificent planner would allocate resources. In our model economy, such a planner's instruments are the time allocations of agents at different cost levels and his objective is to maximize the expected value, discounted at ρ, of total production. The state variable for this problem is the density $f(z,t)$: a point in an infinite dimensional space. We denote the value function, which maps a space of densities into $\mathbb{R}_+$, by W. The problem is then to choose a function $s : \mathbb{R}_+^2 \to [0,1]$ to solve

$$W[f(z,t)] = \max_{s(\cdot,\cdot)} \int_t^\infty e^{-\rho(\tau-t)} \int_0^\infty [1 - s(z,\tau)]\, z^{-\theta} f(z,\tau) dz d\tau$$

subject to the law of motion for f :

$$\frac{\partial f(z,\tau)}{\partial \tau} = -\alpha(s(z,\tau))f(z,\tau)\int_0^z f(y,\tau)dy + f(z,\tau)\int_z^\infty \alpha(s(y,\tau))f(y,\tau)dy. \qquad (20)$$

and with $f(z,t)$ given.

Instead of looking at the planner's Bellman equation directly, it turns out to be more convenient to work with the *marginal value to the planner of one type z individual*, which we denote by $w(z,t)$. This marginal value is more formally defined in Appendix B but the idea is as follows. First, define by $\tilde{w}(z,f)$ the marginal value of one type z individual if the distribution is any function f:

$$\tilde{w}(z,f) \equiv \frac{\delta W(f)}{\delta f(z)}.$$

Here $\delta/\delta f(z)$ is the "functional derivative" of the planner's objective with respect to f at point z, the analogue of the partial derivative $\partial W(\mathbf{f})/\partial f_i$ for the case where z is discrete and hence the distribution $\mathbf{f}$ takes values in $\mathbb{R}^n$. See Appendix B.1 for a rigorous definition of such a derivative. Note that the function $\tilde{w}(z,f)$ is defined over the entire state space, the space of all possible density functions f.

Now we define $w(z,t)$ as the marginal value *along the optimal trajectory* of the distribution, $f(z,t)$:

$$w(z,t) \equiv \tilde{w}(z, f(z,t)),$$

thereby reducing the planner's problem from an infinite-dimensional to a two dimensional problem. The logic is the familiar variational argument: If a plan is optimal, it cannot be improved by telling any individual at any time to deviate from it.

Proposition 1 *The marginal value to the planner of one type z individual, $w(z,t)$, satisfies the Bellman equation*

$$\begin{aligned}\rho w(z,t) = \max_{s\in[0,1]} & \left\{(1-s)z^{-\theta} + \frac{\partial w(z,t)}{\partial t} + \alpha(s)\int_0^z [w(y,t)-w(z,t)]f(y,t)dy\right\} \\ & - \int_z^\infty \alpha(s(y,t))\,[w(y,t)-w(z,t)]\,f(y,t)dy.\end{aligned} \tag{21}$$

This result is intuitive. It states that the flow value $\rho w(z,t)$ contributed by one type z individual is a sum of three terms. The first term is simply the output produced by this individual. The second term is the expected value from improvements in type z's future cost to some $y < z$. We refer to this term as the "internal benefit from search": It takes the same form here as in the problem of an individual stated in (4), with private continuation values replaced by the planner's "social" values. Finally, the third term is the expected value from improvements in the cost of other types $y > z$ to z in case they should meet z. It is only in this term, which we refer to as the "external benefit from search," that the planning problem differs from the individual optimization problem in the decentralized equilibrium. That is, individuals internalize the benefit from search to themselves, but not the benefit to others.

The planner's optimal choice of search intensity satisfies

$$z^{-\theta} = \alpha'(s(z,t))\int_0^z [w(y,t)-w(z,t)]f(y,t)dy. \tag{22}$$

The planner trades off costs and benefits from changing individual search intensities, $s(z,t)$. Increasing $s(z,t)$ has three effects. First, production decreases by $z^{-\theta}$. Second, the outflow of people at z increases by $\alpha'(s(z,t))$, corresponding to a loss

$$-\alpha'(s(z,t))w(z,t)\int_0^z f(y,t)dy.$$

Third, the inflow of people into $y < z$ increases by $\alpha'(s(z,t))$. This corresponds to a gain

$$\alpha'(s(z,t))\int_0^z w(y,t)f(y,t)dy.$$

Note that the integral on the right hand-side of (22) is only taken over $y \leq z$. This is because from (20) changing $s(z,t)$ has no *direct* effect on the distribution at $y > z$ which only depends on the search intensities, $s(y,t)$, of those individuals with costs $y > z$.

As in the decentralized allocation, the Bellman equation here for the marginal value $w(z,t)$ (21) and the law of motion for the distribution (20) constitute a system of two partial differential equations that completely summarize the necessary conditions for a solution to the planning problem.

A balanced growth path for the planning problem is defined in the same way as in the decentralized equilibrium:

$$f(z,t) = e^{\gamma t}\phi(ze^{\gamma t}), \quad w(z,t) = e^{\theta\gamma t}\omega(ze^{\gamma t}).$$

Again, restating (20) and (21) in terms of relative productivities $x = ze^{\gamma t}$, we obtain a BGP Bellman equation

$$\begin{aligned}(\rho - \theta\gamma)\,\omega(x) - \omega'(x)\gamma x = \max_{\sigma\in[0,1]} &\left\{(1-\sigma)x^{-\theta} + \alpha(\sigma)\int_0^x [\omega(y) - \omega(x)]\phi(y)dy\right\} \\ &- \int_x^\infty \alpha[\varsigma(y)][\omega(y) - \omega(x)]\phi(y)dy\end{aligned} \tag{23}$$

and an equation for the BGP distribution, (8). It is important to note that while the equation for the distribution is the same as in the decentralized equilibrium, the planner will generally choose a different time allocation, $\varsigma(x)$, and hence different arrival rates, $\alpha(\varsigma(x))$, implying a different BGP distribution. Here and below we use the notation $\varsigma(x)$ for the planner's policy function, to distinguish it from the policy function $\sigma(x)$ chosen by individual agents. Finally, the parameter γ is given by (12) evaluated using the planner's time allocation, $\varsigma(x)$.

Figure 6 compares the time allocation, $\varsigma(x)$, chosen by the planner with the outcome of the decentralized equilibrium. Not surprisingly, the planner assigns a higher fraction of time spent searching to all individuals so as to internalize the "external benefit from search" discussed above. This implies a higher growth rate $\theta\gamma$ in the planning problem vis-à-vis the decentralized economy. The larger amount of time allocated towards search is also reflected in a lower initial level of total production, $Y(0)$.

Figures 7 and 8 compare the Lorenz curves for flow income and the present value of future income in the decentralized equilibrium and planning problem. An immediate implication of more time allocated towards search is a higher degree of income inequality in the planning problem. This effect is, however, much more muted if we instead measure inequality by the value Lorenz curve, which takes into account mobility in the productivity distribution.

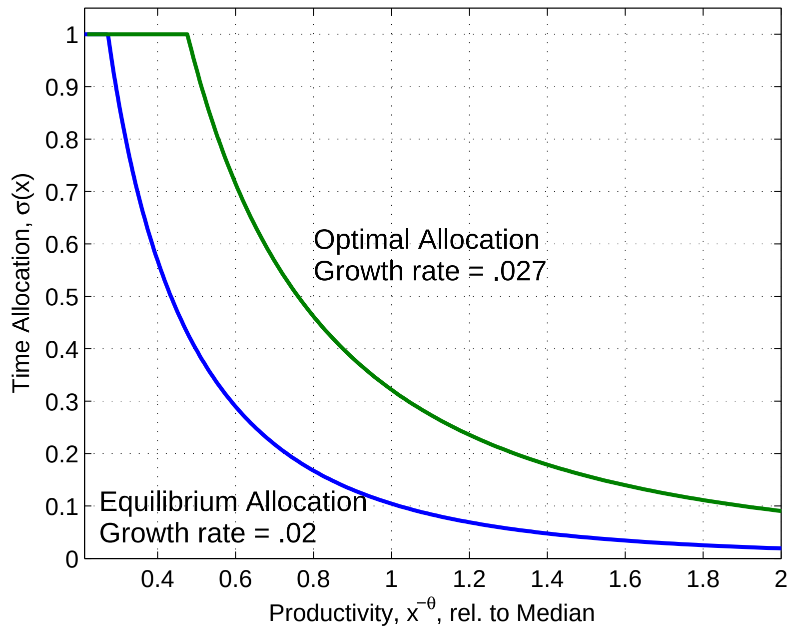

Figure 6: Optimal time allocation, $\sigma(x)$, in decentralized equilibrium and $\varsigma(x)$ in the planning problem

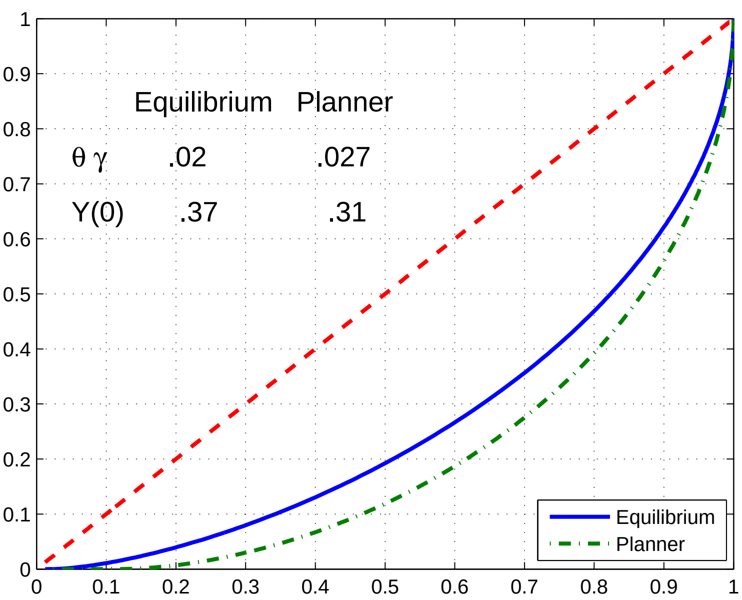

Figure 7: Income Lorenz curves and growth rate, $\theta\gamma$, in decentralized equilibrium and planning problem.

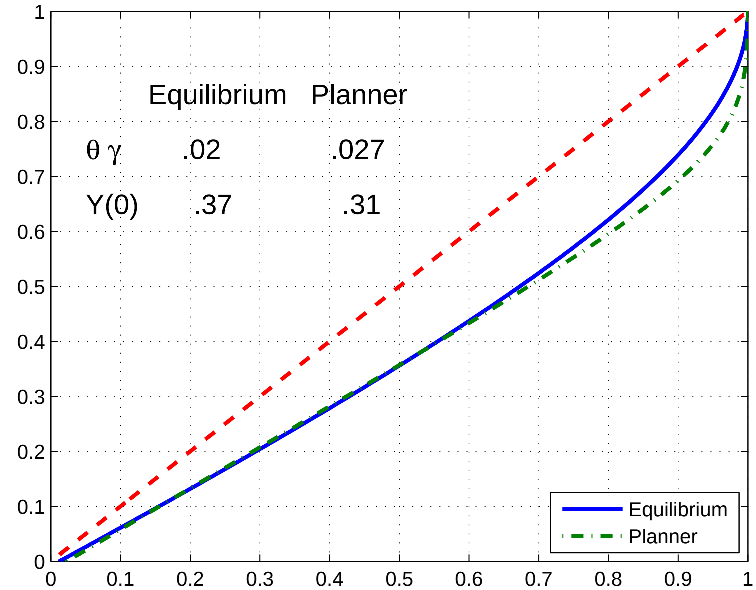

Figure 8: Present value Lorenz curves and growth rate, $\theta\gamma$, in decentralized equilibrium and planning problem.

5 Tax Implementation of the Optimal Allocation

In this section we propose and illustrate a Pigovian tax structure that implements the optimal allocation by aligning the private and social returns to search. In this model a flat tax on income will be neutral: it will have identical effects on both sides of the first order condition (15). We use such a tax to finance a productivity related subsidy $\tau(z,t)$ to offset the opportunity cost $z^{-\theta}s$ of search time s. The flat tax τ_0 satisfies the government budget constraint

$$\int_0^\infty \tau(z,t)s(z,t)z^{-\theta}f(z,t)dz = \tau_0 \int_0^\infty (1-s(z,t))z^{-\theta}f(z,t)dz.$$

Under this tax structure, the individual Bellman equation becomes

$$\rho V(z,t) = \max_{s\in[0,1]} \left\{ (1-\tau_0)[(1-s)z^{-\theta} + \tau(z,t)z^{-\theta}s] + \frac{\partial V(z,t)}{\partial t} + \alpha(s)\int_0^z [V(y,t)-V(z,t)]f(y,t)dy \right\}.$$

The law of motion for the distribution (3) and the expression for aggregate output (11) are unchanged.

Let $v_n(x)$ (n for "net") be the present value of an individual's earnings, net of subsidies and taxes, and replace the equation defining the value function on a BGP (6) by

$$V(z,t) = (1-\tau_0)e^{\theta\gamma t}v_n(ze^{\gamma t}).$$

In addition $\tau(z,t) = \tau(ze^{\gamma t})$. This function $v_n(x)$ satisfies

$$(\rho - \theta\gamma)\, v_n(x) - v_n'(x)\gamma x = \max_{\sigma\in[0,1]} \left\{ (1-\sigma)x^{-\theta} + \tau(x)x^{-\theta}\sigma + \alpha(\sigma)\int_0^x [v_n(y) - v_n(x)]\phi\,(y)\,dy \right\},$$

where both the density ϕ and the growth rate γ are taken from the planning problem.

As before, we let

$$S_n(x) = \int_0^x [v_n(y) - v_n(x)]\phi\,(y)\,dy.$$

The first order condition is

$$(1-\tau(x))\,x^{-\theta} \leq \alpha'(\sigma)S_n(x) \;\text{ with equality if } \sigma < 1.$$

The agent takes $\tau(x)$ as given and chooses $\sigma(x)$.

The planner wants to choose the subsidy rate $\tau(x)$ so that individuals choose $\sigma(x) = \varsigma(x)$, the time allocation that the planner has already decided on. This choice is then

$$(1-\tau(x))x^{-\theta} = \alpha'(\varsigma(x))S_n(x) \tag{24}$$

provided that $\varsigma(x) < 1$. At the smallest value $\bar{x}$ at which $\varsigma(x) = 1$, $\tau(\bar{x})$ is the rate at which the agent is indifferent between working a small amount and not working at all. For $x > \bar{x}$, equality in (24) gives the subsidy rate that maintains indifference as cost increases from $\bar{x}$. Of course, any higher subsidy rate in this range would have the same effect.

The Bellman equation under the tax policy just described is

$$(\rho - \theta\gamma)\, v_n(x) - v_n'(x)\gamma x = x^{-\theta} - \varsigma(x)\alpha'(\varsigma(x))S(x) + \alpha(\varsigma(x))\int_0^x [v_n(y) - v_n(x)]\phi\,(y)\,dy.$$

With the function $\alpha(\sigma) = k\sigma^\eta$ that we use, $\alpha(\sigma) - \sigma\alpha'(\sigma) = k\sigma^\eta - \sigma\eta k\sigma^{\eta-1} = (1-\eta)\,\alpha(\sigma)$ and so

$$(\rho - \theta\gamma)\, v_n(x) - v_n'(x)\gamma x = x^{-\theta} + (1-\eta)\,\alpha(\varsigma(x))\int_0^x [v_n(y) - v_n(x)]\phi\,(y)\,dy \tag{25}$$

on $(0,\bar{x})$. On $[\bar{x},\infty)$, $v_n\,(x) = v_n\,(\bar{x})$.

Given $\varsigma(x)$ and γ from the planning problem, (25) can be solved for $v_n(x)$ and $S_n(x)$, applying the algorithm used earlier. The tax rate $\tau(x)$ can then be computed using (24). Figure

9 plots the two policy functions $\sigma(x)$ and $\varsigma(x)$ and the subsidy rate $\tau(x)$. On the interval A on the figure, agents choose $\sigma = 1$ in both the decentralized and planned cases, so no tax is needed to encourage more search. On the interval B, the planner wants everyone to search full time so $\tau(x)$ is chosen to induce agents to prefer this to doing any production. The agents with the lowest productivity on the interval B choose to work in the decentralized economy but the planned allocation implemented by the tax improves their return from search enough that no additional tax incentive is needed. On the interval C, the planner wants to increase everyone's search: compare $\sigma(x)$ to $\varsigma(x)$. The opportunity cost of search increases without limit as $x^{-\theta} \to \infty$. This requires that $\tau(x)$ be an increasing function on C.

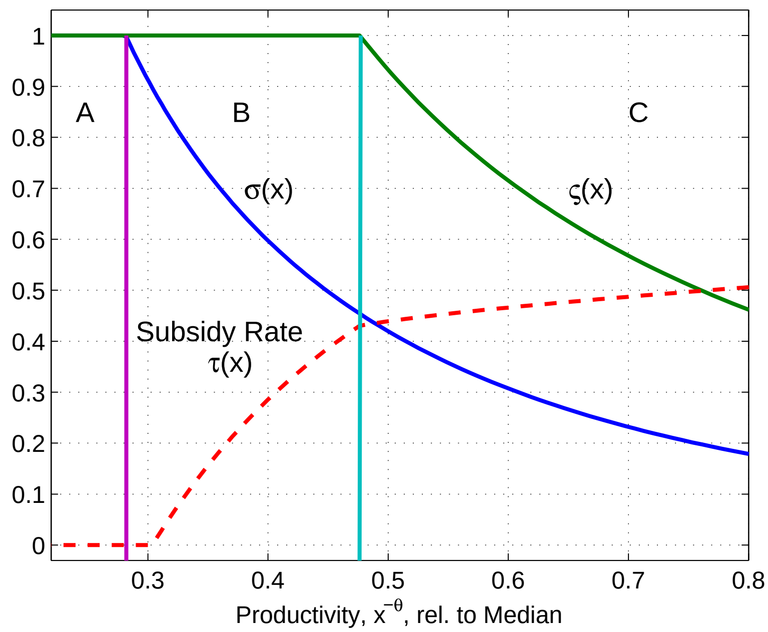

Figure 9: Pigovian Implementation of the Optimal Allocation

In the example shown in Figure 9, the only agents with positive earnings are those on the interval C. All of them pay the flat tax τ_0 on earnings and receive offsetting subsidies designed to encourage search. These subsidy rates increase faster than earnings, making the tax system as a whole regressive. It is worth emphasizing that this is a feature of a tax system which has the single purpose of encouraging productivity innovation. Considerations of distorting taxes and distribution, central to much of tax analysis, have simply been set aside.

6 Alternative Learning Technologies

All of the analysis so far has been carried out under the learning technology described in Section 2. Even under the limits of a one-dimensional model of knowledge, however, there are many other models of learning that might be considered. It turns out that the algorithm we describe in Section 3 is not difficult to adapt to some alternatives. Ultimately, which of these and other alternatives are substantively interesting will depend on the evidence we are trying to understand. In this section, we simply illustrate some theoretical possibilities with two examples.

6.1 Limits to Learning

In the theory we have considered so far, a person's current productivity level determines his ability to produce goods but has no effect on his ability to acquire new knowledge. The outcome of a search by agent z who meets an agent $y < z$ is y, regardless of the value of his own cost z. But it is easy to think of potential knowledge transfers that cannot be carried out if the "recipient's" knowledge level is too different from that of the "donor." To explore this possibility, we make use of an appropriate "kernel" to modify the law of motion for the distribution (3). Assume for example that if an agent at z meets another agent at y, he can adopt y with probability $k(y, z, t)$; with probability $1 - k(y, z, t)$ he cannot do this and retains his previous cost z. Then the law of motion for the distribution becomes

$$\frac{\partial f(z,t)}{\partial t} = f(z,t)\int_z^\infty \alpha(s(y,t))f(y,t)k(z,y,t)dy - \alpha(s(z,t))f(z,t)\int_0^z f(y,t)k(y,z,t)dy.$$

A natural assumption on the kernel k is that the probability of z learning from y is unchanged over time if z and y are at the same quantiles of the cost distribution

$$k(y,z,t) = k(y',z',t'), \quad \text{for } F(y,t) = F(y',t') \text{ and } F(z,t) = F(z',t'). \tag{26}$$

(Think of our cohort interpretation of the stochastic careers in Figure 5. A "newborn" beginning at date $t + \tau$ immediately benefits from the fact that productivities in general are $e^{\theta\gamma\tau}$ larger than they were for his "parent" who arrived at t.) Along a balanced growth path for the distribution as defined in (5), all cost quantiles shrink at a common rate γ. Hence (26) can be written as[8]

$$k(y,z,t) = k(ye^{\gamma t}, ze^{\gamma t}, 0).$$

[8]To see this, use that along a BGP $F(z,t) = \Phi(ze^{\gamma t})$ and let $t' = 0$ in (26).

We find it convenient to work with the functional form

$$k(y, z, 0) = e^{-\kappa|y-z|}, \tag{27}$$

where $\kappa > 0$ is the rate at which learning probabilities fall off as knowledge differences increase. We can think of this kernel as reflecting an ordering in the learning process or some limits to intellectual range.[9] An equivalent interpretation of this kernel is that *meeting* probabilities depend on the distance between different productivity types, so that each person has a higher chance of meeting those with a knowledge level close to his own. In this interpretation, the parameter κ captures the degree of socioeconomic segregation or stratification in a society.

With the functional form in (27), we can derive the following expressions for the law of motion for the distribution along a BGP

$$\phi(x)\gamma + \phi'(x)\gamma x = \phi(x)\int_x^\infty \alpha(\sigma(y))\phi(y)e^{-\kappa(y-x)}dy - \alpha(\sigma(x))\phi(x)\int_0^x \phi(y)e^{-\kappa(x-y)}dy$$

Evaluating at $x = 0$, the growth rate of the economy is

$$\gamma = \int_0^\infty \alpha(\sigma(y))\phi(y)e^{-\kappa y}dy. \tag{28}$$

Analogously, the corresponding Bellman equation is

$$(\rho - \theta\gamma)\, v(x) - v'(x)\gamma x = \max_{\sigma\in[0,1]} \left\{ (1-\sigma)x^{-\theta} + \alpha(\sigma)\int_0^x [v(y) - v(x)]\phi(y)e^{-\kappa(x-y)}dy \right\}.$$

Figure 10 plots the optimal time allocation, $\sigma(x)$, for various values of the parameter measuring the limits to learning, κ. Going from $\kappa = 0$ to $\kappa = 0.01$ changes people's search behavior dramatically. High productivity types allocate a roughly equal amount of time towards knowledge acquisition regardless of κ. But low productivity types are discouraged from search, resulting in search intensity being a hump-shaped function of current productivity. The reason for this is that the benefit from search

$$S(x) = \int_0^x [v(y) - v(x)]\phi(y)e^{-\kappa(x-y)}dy$$

is no longer very high for low productivity (high cost) types. Because low productivity types

[9] Jovanovic and Nyarko (1996) suggested the following rationale for such limits to learning: different productivity types, $z^{-\theta}$, correspond to different *activities* and human capital is partially specific to a given activity. When an agent switches to a new activity, he loses some of this human capital, and more so the more different is the new activity.

also have a low probability of benefiting from a meeting with a high productivity type, their expected payoff from search is low and their search effort is discouraged. Increases in κ have little effect on income inequality, as Figure 11 shows, but a large effect on present value inequality, especially for low productivity types, as Figure 12 shows. Another way of putting this is that social mobility decreases dramatically as κ increases.

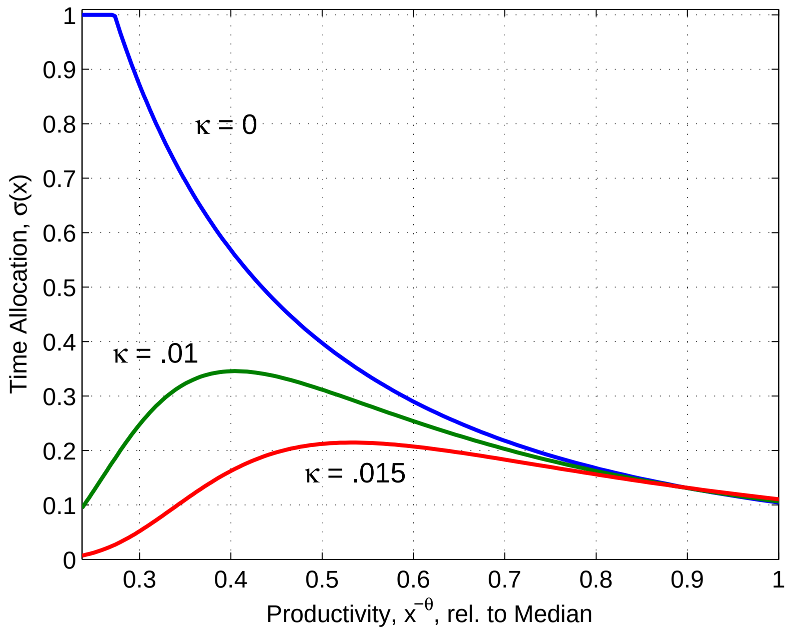

Figure 10: Optimal Time Allocation, $\sigma(x)$, for various κ values.

As can also be seen on these figures, the growth rate of the economy, $\theta\gamma$, declines as the limits to learning, κ, increase. This is due to two effects. First and as just discussed, some low productivity types allocate less time towards search. Because the growth rate of the economy is an average of individual search intensities, this depresses growth. Second, there is a direct negative effect of κ on the growth rate as can be seen in the formula (28). This direct effect comes from the fact that the number of agents whose productivity exceeds any value $1/\varepsilon^{\theta}$ at any point in time (the inflow into the cost interval $(0, \varepsilon)$) is lower because low productivity agents face a lower probability of meeting high productivity ones.

Finally, Figure 13 also reports the time allocation chosen by an idealized social planner who is constrained by the limits to learning that characterize the decentralized equilibrium. Compared with the time allocation in the decentralized equilibrium reported in Figure 10, the planner chooses a higher search intensity for each κ value. Note that even this planner gives up on low productivity types because the social value of their attempts at knowledge acquisition is diminished by the fact that they are unlikely to learn from those with high productivity.

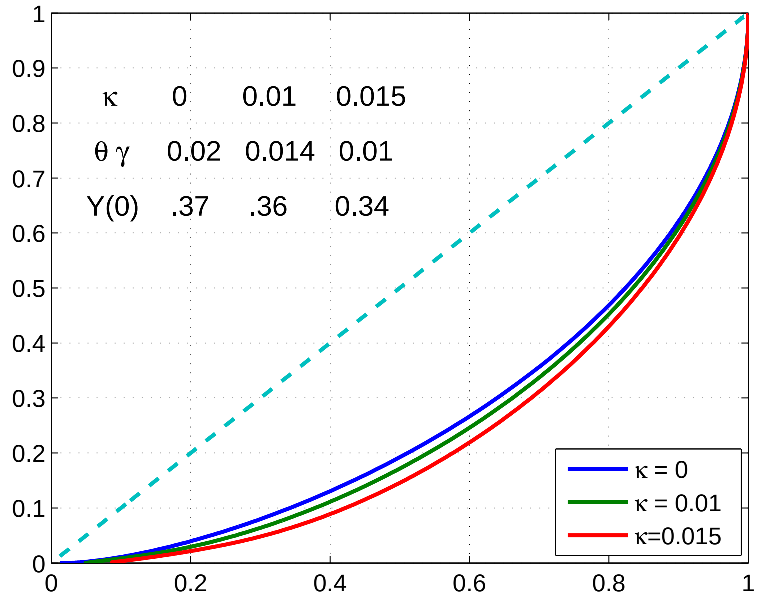

Figure 11: Income Lorenz curves for various κ values.

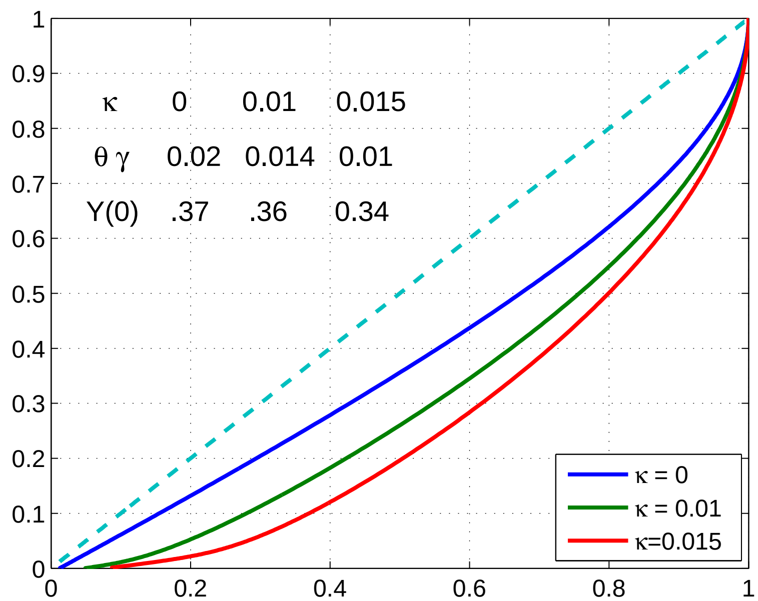

Figure 12: Present Value Lorenz curves for various κ values.

To obtain information on our parameter κ, our theory suggests studying the degree of social

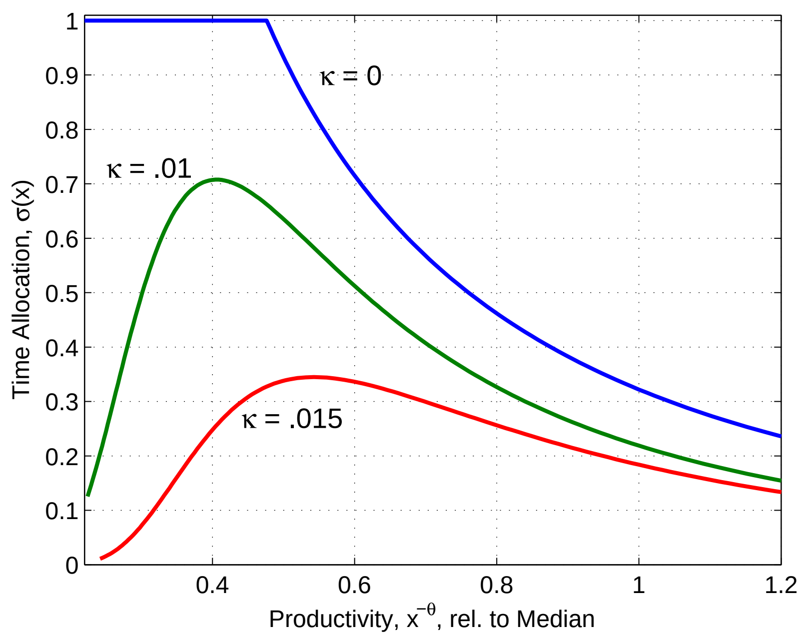

Figure 13: Planner's Time Allocation, $\sigma(x)$, for various κ values.

mobility in a society. Comparing income and present value Lorenz curves in Figures 11 and 12, social mobility decreases sharply with κ. Both evidence on intra- and inter-generational mobility is informative, even though our theory does not distinguish between the two. In section 3, we have already cited some studies on on-the-job human capital accumulation and the slope of earnings profiles. There are also many studies examining the correlation in lifetime income between parents and children (e.g. Solon, 1992) or intergenerational transition probabilities between different income quantiles (e.g. Zimmerman, 1992).[10]

6.2 Symmetric Meetings

Another feature of the learning technology applied in Sections 1-5 is the fact that meetings between two agents z and y are completely asymmetric. Agents could only upgrade their knowledge through active search while the other party to the meeting gains nothing and can well be unaware that he is being met.

Depending on the specific application, this may not be the best assumption. For example, Arrow (1969) argues that "the diffusion of an innovation [is] a process formally akin to the

[10]See Becker and Tomes (1979), Benabou (2002) and Benhabib, Bisin and Zhu (2011) for alternative theories of the relationship between inequality and the degree of intragenerational mobility.

spread of an infectious disease." This description of meetings has a symmetric component: a person can get "infected" even when he didn't actively search for the "disease". The model can easily be extended to encompass the case where meetings are symmetric, as we now show. To capture symmetric meetings, we assume that even if y initiated the meeting, z can learn from y with probability β. Therefore, β parameterizes how strong passive spillovers are: $\beta = 0$ corresponds to our benchmark model; $\beta = 1$ is the case of perfectly symmetric meetings. Under this assumption, we obtain the new law of motion

$$\begin{aligned}\frac{\partial f(z,t)}{\partial t} =& - f(z,t)\int_0^z [\alpha(s(z,t)) + \beta\alpha(s(y,t))]f(y,t)dy \\ &+ f(z,t)\int_z^\infty [\alpha(s(y,t)) + \beta\alpha(s(z,t))]f(y,t)dy.\end{aligned}$$

The main difference from the asymmetric law of motion (3) is that here the search intensities $s(z,t)$ and $s(y,t)$ enter in a symmetric fashion. Agents at z now have opportunities to upgrade their productivities even if another agent y initiated the meeting. These opportunities arrive at rate $\alpha(s(z,t)) + \beta\alpha(s(y,t))$ rather than just $\alpha(s(z,t))$. The Bellman equation now becomes

$$\rho V(z,t) = \max_{s\in[0,1]} \left\{(1-s)z^{-\theta} + \frac{\partial V(z,t)}{\partial t} + \int_0^z [\alpha(s) + \beta\alpha(s(y,t))][V(y,t) - V(z,t)]f(y,t)dy\right\}.$$

The corresponding equations along a BGP are found as above. Figures 14 and 15 report the optimal time allocation and productivity density for various values of the parameter measuring the amount of passive spillovers, β. The more knowledge that can be acquired without actively searching, the lower is agents' incentive to search. Since the economy-wide growth rate is still an average of individual search intensities, this "free-riding" implies that the growth rate is actually *lower* the higher are spillovers, β. At the same time, a higher β implies that the BGP distribution places more mass on high productivity types (Figure 15). Figures 16 to 18 compare the decentralized equilibrium just described to the allocation chosen by a social planner when meetings are symmetric. The time allocation chosen by the social planner now differs dramatically from that in the decentralized equilibrium. The planner makes the most productive agents search full time, the high opportunity cost notwithstanding. He views them as even more valuable as "teachers," reaching out to meet less productive agents, increasing the probability that less productive agents will learn from them. After such an unproductive agent becomes productive, he searches full time for a while, but as his relative productivity declines (as in panel (b) of Figure 5) he resumes working. While period-by-period income is more unequally distributed under the planner's time (Figure 17), this is no longer true for the

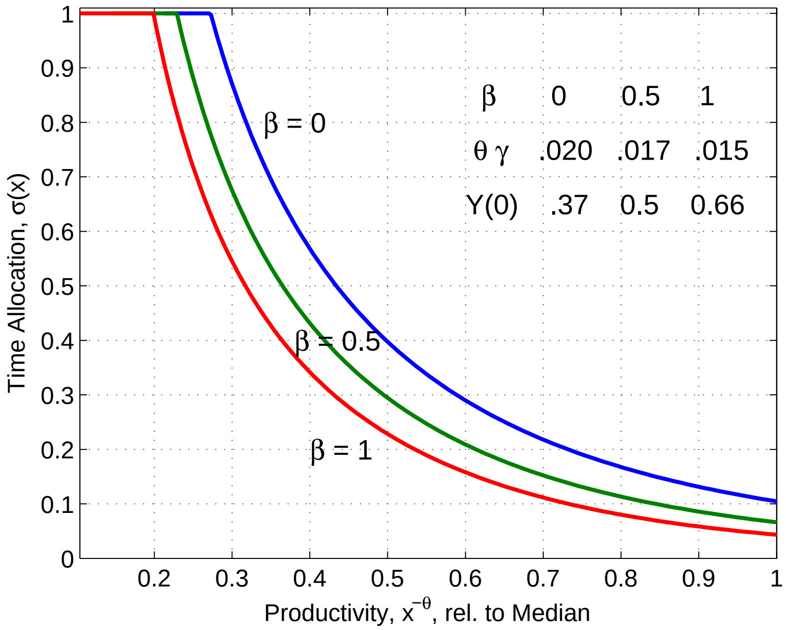

Figure 14: Optimal Time Allocation, $\sigma(x)$, for various β values

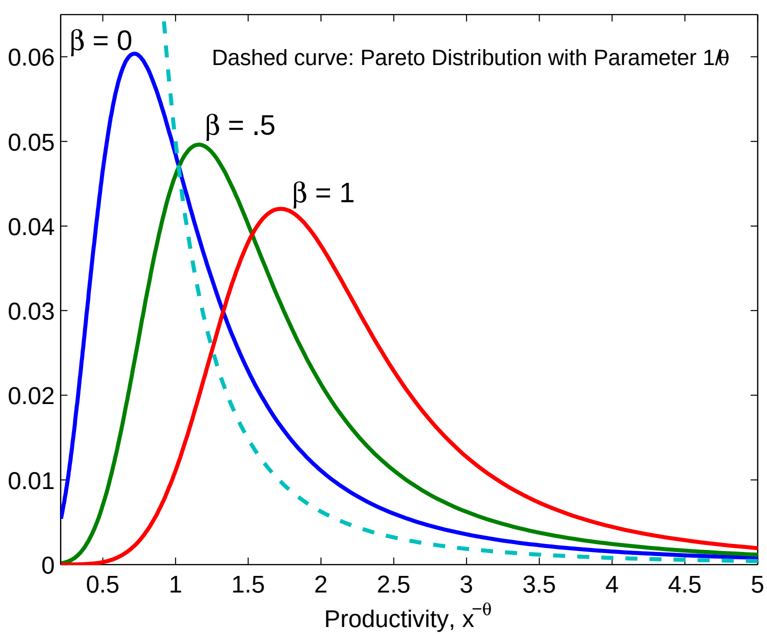

Figure 15: Productivity Densities for various β values

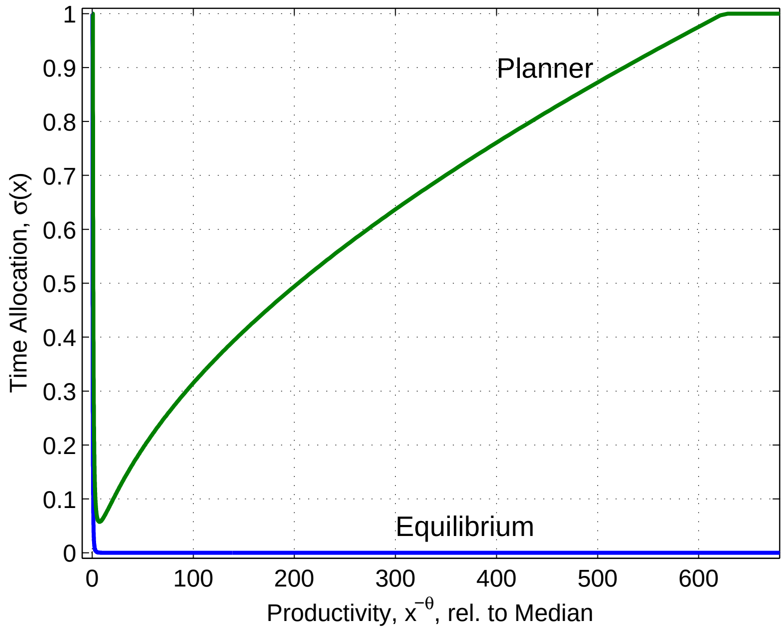

Figure 16: Optimal Time Allocation with Symmetric Meetings

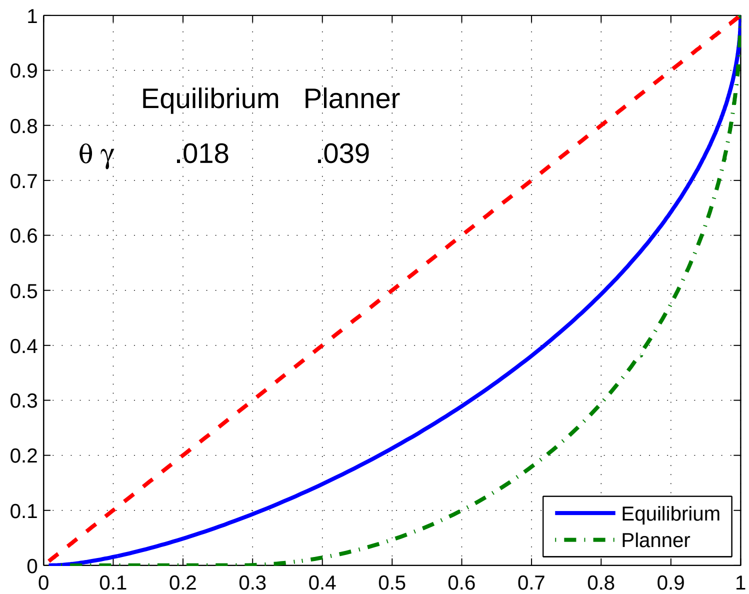

Figure 17: Income Lorenz Curves, Symmetric Meetings

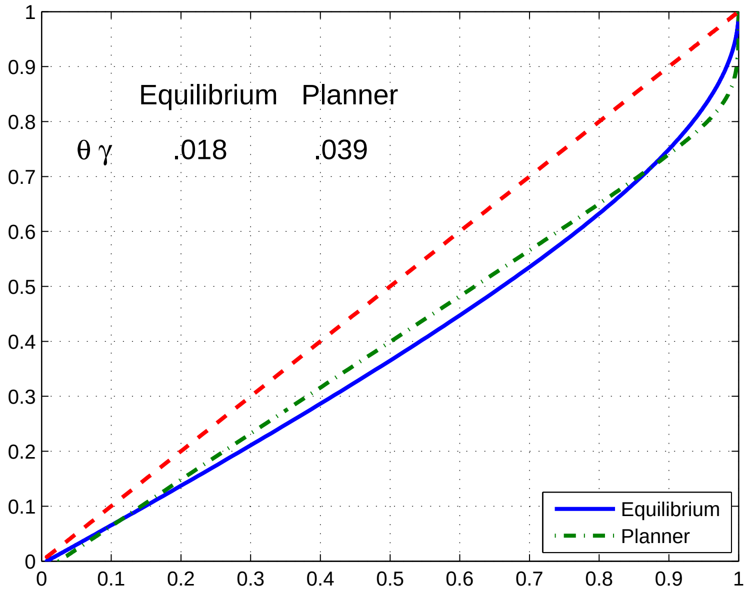

Figure 18: Value Lorenz Curves, Symmetric Meetings

present value of income. The Lorenz curves in Figure 18 cross, meaning that in parts of the distribution the decentralized equilibrium features too little mobility relative to the planning problem.

7 Conclusion

We have proposed and studied a new model of economic growth in which individuals differ only in their current productivity, and the state of the economy is fully described by the probability distribution of productivities. The necessary conditions for equilibrium in the model take the form of a Bellman equation describing individual decisions on the way to allocate time between producing and searching for new ideas and a law-of-motion for the economy-wide productivity distribution. With the right kind of initial conditions these forces can interact to generate sustained growth. We show that among these possibilities is a balanced growth path, characterized by a constant growth rate and a stable Lorenz curve describing relative incomes. We provide an algorithm for calculating solutions along this path.

This solution is the outcome of a decentralized system in which each agent acts in his own interest. But the new knowledge obtained by any one agent benefits others by enriching their intellectual environment and raising the return to their own search activities. We then formulate the problem of a hypothetical planner who can allocate people's time so as to internalize this external effect. We show how the decentralized algorithm can be adapted to compute the planning solution as well, and compare it to the decentralized solution. We then consider tax structures that implements an optimal solution. Finally we provide two examples of alternative learning technologies, as concrete illustrations of other directions that might be pursued.

All of this is carried out in a starkly simple context in order to reveal the economic forces involved and the nature of their interactions, and to build up our experience with a novel and potentially useful mathematical structure. But we also believe that the external effects we study here are centrally important to the understanding of economic growth and would like to view our analysis as a step toward a realistically quantitative picture of the dynamics of production and distribution.[11]

Appendix

A Derivation of Bellman Equation

Let time be indexed by $t, t+\Delta, ...$ Denote the discount factor between two periods by $1-\Delta\rho$. The Bellman equation is

$$\begin{aligned} V(z,t) = \max_{s\in[0,1]} &\Delta(1-s)z^{-\theta} + (1-\Delta\rho) \\ &\times \left\{ \Delta\alpha(s) \int_0^\infty \max\{V(y,t+\Delta), V(z,t+\Delta)\} f(y,t+\Delta)dy + (1-\Delta\alpha(s))V(z,t+\Delta) \right\} \end{aligned}$$

Rewrite as

$$\begin{aligned} V(z,t) = \max_{s\in[0,1]} &\Delta(1-s)z^{-\theta} + (1-\Delta\rho) \\ &\times \left\{ \Delta\alpha(s) \int_0^z [V(y,t+\Delta) - V(z,t+\Delta)] f(y,t+\Delta)dy + V(z,t+\Delta) \right\}. \end{aligned}$$

Subtract $(1-\Delta\rho)V(z,t)$ from both sides

$$\begin{aligned} \Delta\rho V(z,t) = \max_{s\in[0,1]} &\Delta(1-s)z^{-\theta} + (1-\Delta\rho) \\ &\times \left\{ \Delta\alpha(s) \int_0^z [V(y,t+\Delta) - V(z,t+\Delta)] f(y,t+\Delta)dy + V(z,t+\Delta) - V(z,t) \right\}. \end{aligned}$$

Dividing by Δ and taking the limit as $\Delta \to 0$ yields equation (4).

[11] In this regard, see also Choi (forthcoming).

B Proof of Proposition 1

B.1 Mathematical Preliminaries

The value function of the planner $W[f(z,t)]$ is a *functional*, that is a map from a space of functions to the real numbers, or informally a "function of a function". The planner chooses a function $f(z,t)$ to maximize this functional, which is the prototypical problem in the calculus of variations. The concept of a *functional derivative* is helpful in solving this problem.

Definition: The *functional derivative* of W with respect to f at point y is

$$\frac{\delta W[f(z)]}{\delta f(y)} \equiv \lim_{\varepsilon \to 0} \frac{W[f(z)+\varepsilon\delta(z-y)] - W[f(z)]}{\varepsilon} = \frac{d}{d\varepsilon} W[f(z)+\varepsilon\delta(z-y)]\bigg|_{\varepsilon=0} \tag{29}$$

where $\delta(\cdot)$ is the Dirac delta function.

The functional derivative is the natural generalization of the partial derivative. Thus, consider the case where z is discrete and takes on n possible values, $z \in \{z_1, ..., z_n\}$. The corresponding distribution function is then simply a vector $\mathbf{f} \in \mathbb{R}^n$ and the planner's value function is an ordinary function of n variables, $W : \mathbb{R}^n \to \mathbb{R}$. The partial derivative in this case is defined as

$$\frac{\partial W(\mathbf{f})}{\partial f_i} \equiv \lim_{\varepsilon \to 0} \frac{W(f_1, ..., f_i + \varepsilon, ..., f_n) - W(f_1, ..., f_i, ..., f_n)}{\varepsilon} \tag{30}$$

If we denote by $\boldsymbol{\delta}(i) \in \mathbb{R}^n$ the vector that has elements $\delta_i(i) = 1$ and $\delta_i(j) = 0$ for all $i \neq j$, then (30) cab be written as

$$\frac{\partial W(\mathbf{f})}{\partial f_i} \equiv \lim_{\varepsilon \to 0} \frac{W(\mathbf{f} + \varepsilon\boldsymbol{\delta}(i)) - W(\mathbf{f})}{\varepsilon} = \frac{d}{d\varepsilon} W(\mathbf{f} + \varepsilon\boldsymbol{\delta}(i))\bigg|_{\varepsilon=0}$$

It can be seen that the functional derivative in (29) is defined in the exact same way.

Another fact that will be useful below, is that the integral of the Dirac delta function can be expressed as the Heaviside step function

$$\int_{-\infty}^{z} \delta(\zeta) d\zeta = \begin{cases} 1, & z \geq 0 \\ 0, & z < 0 \end{cases} \equiv H(z)$$

Similarly, integrals of the Dirac delta function centered at y are

$$\int_{-\infty}^{z} \delta(\zeta - y) d\zeta = H(z-y), \quad \int_{z}^{\infty} \delta(\zeta - y) d\zeta = H(y-z)$$

B.2 Bellman Equation

With this mathematical apparatus in hand, the planner's problem can be written in recursive form as

$$\rho W[f] = \max_s \int_0^\infty [1 - s(z)] z^{-\theta} f(z) dz + \int_0^\infty \frac{\delta W[f]}{\delta f(z)} \hat{f}(z; s, f) dz \tag{31}$$

where

$$\hat{f}(z; s, f) = -\alpha(s(z)) f(z) \int_0^z f(y) dy + f(z) \int_z^\infty \alpha(s(y)) f(y) dy. \tag{32}$$

Lemma 2: A solution $s(\cdot)$ to the planning problem must satisfy

$$z^{-\theta} = \alpha'(s(z)) \int_0^z \left[\frac{\delta W[f]}{\delta f(y)} - \frac{\delta W[f]}{\delta f(z)} \right] f(y) dy \quad \text{for all } z. \tag{33}$$

Proof: The planner's first order condition is

$$0 = \frac{\delta}{\delta s(z)} \left[\int_0^\infty [1 - s(y)] y^{-\theta} f(y) dy + \int_0^\infty \frac{\delta W[f]}{\delta f(y)} \hat{f}(y; s, f) dy \right]$$

Using the definition of a functional derivative

$$\begin{aligned} &\frac{\delta}{\delta s(z)} \int_0^\infty [1 - s(y)] y^{-\theta} f(y) dy = \frac{d}{d\varepsilon} \int_0^\infty [1 - (s(y) + \varepsilon \delta(y - z))] y^{-\theta} f(y) dy \bigg|_{\varepsilon = 0} \\ &= -\frac{d}{d\varepsilon} \int_0^\infty \delta(y - z) y^{-\theta} f(y) dy = -z^{-\theta} f(z) \end{aligned} \tag{34}$$

Using similar manipulations for the second term:

$$z^{-\theta} f(z) = \int_0^\infty \frac{\delta W[f]}{\delta f(y)} \frac{\delta \hat{f}(y; s, f)}{\delta s(z)} dy \tag{35}$$

From (32), we have

$$\begin{aligned} \frac{\delta \hat{f}(y; s, f)}{\delta s(z)} &= -\alpha'(s(y)) \delta(y - z) f(y) \int_0^y f(\zeta) d\zeta + f(y) \int_y^\infty \alpha'(s(\zeta)) \delta(\zeta - z) f(\zeta) d\zeta \\ &= -\alpha'(s(y)) \delta(y - z) f(y) \int_0^y f(\zeta) d\zeta + f(y) \alpha'(s(z)) f(z) H(z - y) \end{aligned}$$

where we use the relation between Dirac delta and Heaviside step function.

Therefore (35) can be written as

$$
\begin{aligned}
z^{-\theta}f(z) &= \int_0^\infty \frac{\delta W[f]}{\delta f(y)} f(y)\alpha'(s(z))f(z)H(z-y)dy - \int_0^\infty \frac{\delta W[f]}{\delta f(y)}\alpha(s(y))\delta(y-z)f(y)\int_0^y f(\zeta)d\zeta dy \\
&= \int_0^z \frac{\delta W[f]}{\delta f(y)}\alpha'(s(z))f(z)f(y)dy - \frac{\delta W[f]}{\delta f(z)}\alpha'(s(z))f(z)\int_0^z f(\zeta)d\zeta
\end{aligned}
$$

Collecting terms yields (33).□

B.3 Proof of Proposition 1

Differentiating (31) with respect to $f(z)$, we obtain

$$
\rho\tilde{w}(z,f) = (1-s(z))z^{-\theta} + \frac{\delta}{\delta f(z)}\int_0^\infty \frac{\delta W(f)}{\delta f(y)}\hat{f}(y;s,f)dy \tag{36}
$$

We have

$$
\frac{\delta}{\delta f(z)}\int_0^\infty \frac{\delta W(f)}{\delta f(y)}\dot{f}(y;s,f)dy = \int_0^\infty \frac{\delta \tilde{w}(y,f)}{\delta f(y)}\hat{f}(y;s,f)dy + \int_0^\infty \tilde{w}(y,f)\frac{\delta \hat{f}(y;s,f)}{\delta f(z)}dy \tag{37}
$$

From (32) and using the relation between Dirac delta and Heaviside step function,

$$
\begin{aligned}
\frac{\delta \hat{f}(y;s,f)}{\delta f(z)} =& -\alpha(s(y))\delta(y-z)\int_0^y f(\zeta)d\zeta - \alpha(s(y))f(y)H(y-z) \\
&+ \delta(y-z)\int_y^\infty \alpha(s(\zeta))f(\zeta)d\zeta + f(y)\alpha(s(z))H(z-y)
\end{aligned}
$$

and similarly for the last term. Hence

$$
\begin{aligned}
\int_0^\infty \tilde{w}(y,f)\frac{\delta \hat{f}(y;s,f)}{\delta f(z)}dy =& -\tilde{w}(z,f)\alpha(s(z))\int_0^z f(\zeta)d\zeta - \int_z^\infty \tilde{w}(y,f)\alpha(s(y))f(y)dy \\
&+ \tilde{w}(z,f)\int_z^\infty \alpha(s(\zeta))f(\zeta)d\zeta + \int_0^z \tilde{w}(y,f)f(y)\alpha(s(z))dy \\
=& \alpha(s(z))\int_0^z [\tilde{w}(y,f)-\tilde{w}(z,f)]f(y)dy \\
&- \int_z^\infty \alpha(s(y))\left[\tilde{w}(y,f)-\tilde{w}(z,f)\right]f(y)dy
\end{aligned}
$$

Combining with (36) and (37), we have

$$
\begin{aligned}
\rho\tilde{w}(z,f) =& (1-s(z))z^{-\theta} + \int_0^\infty \frac{\delta \tilde{w}(y,f)}{\delta f(y)}\hat{f}(y;s,f)dy \\
&+ \alpha(s(z))\int_0^z [\tilde{w}(y,f)-\tilde{w}(z,f)]f(y)dy - \int_z^\infty \alpha(s(y))\left[\tilde{w}(y,f)-\tilde{w}(z,f)\right]f(y)dy
\end{aligned}
$$

Define

$$w(z,t) \equiv \tilde{w}(z, f(z,t)) \tag{38}$$

Then

$$\frac{\partial w(z,t)}{\partial t} = \int_0^\infty \frac{\delta \tilde{w}(y, f(y,t))}{\delta f(y,t)} \hat{f}(y; s(y,t), f(y,t)) dy$$

and hence

$$\begin{aligned}\rho w(z,t) =& (1 - s(z,t)) z^{-\theta} + \frac{\partial w(z,t)}{\partial t} + \alpha(s(z,t)) \int_0^z [w(y,t) - w(z,t)] f(y,t) dy \\ & - \int_z^\infty \alpha(s(y,t)) \left[w(y,t) - w(z,t)\right] f(y,t) dy\end{aligned} \tag{39}$$

Further, using (38), the FOC (33) can be written as

$$z^{-\theta} = \alpha'(s(z,t)) \int_0^z \left[w(z,t) - w(y,t)\right] f(y) dy \tag{40}$$

(39) and (40) can be summarized as (21).□

C Computation

C.1 Step 1: Solution to Bellman Equation – Decentralized Equilibrium

The BGP Bellman equation (10) can be rewritten as

$$(\rho - \theta\gamma) v(x) = x^{-\theta}[1 - \sigma(x)] + \gamma x v'(x) + \alpha[\sigma(x)] S(x)$$

where $S(x)$ is defined as

$$S(x) \equiv \int_0^x [v(y) - v(x)] \phi(y) dy = \int_0^x v(y)\phi(y) y - v(x)\Phi(x)$$

and $\Phi(x) = \int_0^x \phi(y)dy$, that is the cdf corresponding to ϕ. The optimal choice $\sigma(x)$ is defined implicitly by the first order condition (15). We further have a boundary condition (16).

We solve these equations using a finite difference method which approximates the function $v(x)$ on a finite grid, $x \in \{x_1, ..., x_I\}$. We use the notation $v_i = v(x_i), i = 1, ..., I$.[12] We approximate the derivative of v using a forward difference

$$v'(x_i) \approx \frac{v_{i+1} - v_i}{h_i}$$

[12] A useful reference is Candler (1999).

where h_i is the distance between grid points x_i and x_{i+1}. The boundary condition (16) then implies

$$0 \approx v'(x_I) = \frac{v_{I+1} - v_I}{h_I} \quad \Rightarrow \quad v_{I+1} = v_I. \tag{41}$$

Similarly, we approximate $S(x)$ by

$$S_i = S(x_i) \approx \sum_{l=1}^{i} v_l \phi_l h_l - v_i \Phi_i \tag{42}$$

Further, denote by $\sigma_i = \sigma(x_i)$ and $\alpha_i = \alpha[\sigma(x_i)]$ the optimal time allocation and search intensity.

We proceed in an iterative fashion: we guess v_i^0 and then for $j = 0, 1, 2...$ form v_i^{j+1} as follows. Form S_i^j as in (42), and obtain σ_i^j and α_i^j from the first order condition (44). Write the Bellman equation as

$$(\rho - \theta\gamma) v_i^{j+1} = (1 - \sigma_i^j) x_i^{-\theta} + \gamma x_i \frac{v_{i+1}^{j+1} - v_i^{j+1}}{h_i} + \alpha_i^j \left[\sum_{l=1}^{i} v_l^{j+1} \phi_l h_l - v_i^{j+1} \Phi_i \right], \quad i = 1, ..., I \tag{43}$$

Given v^j and hence σ^j and α^j, and using the boundary condition $v_{I+1}^{j+1} = v_I^{j+1}$, (43) is a system of I equations in I unknowns, $(v_1^{j+1}, ..., v_I^{j+1})$, that can easily be solved for the updated value function, v^{j+1}. Using matrix notation

$$\mathbf{A}^j v^{j+1} = b^j, \quad b_i^j = (1 - \sigma_i) x_i^{-\theta}, \quad \mathbf{A}^j = \mathbf{B}^j - \mathbf{C}^j$$

where[13]

$$\mathbf{B}^j = \begin{bmatrix} \rho-\theta\gamma+\alpha_1^j\Phi_1+\frac{\gamma x_1}{h_1} & -\frac{\gamma x_1}{h_1} & 0 & \dots & 0 \\ 0 & \rho-\theta\gamma+\alpha_2^j\Phi_2+\frac{\gamma x_2}{h_2} & -\frac{\gamma x_2}{h_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & \rho-\theta\gamma+\alpha_I^j\Phi_I \end{bmatrix}$$

$$\mathbf{C}^j = \begin{bmatrix} \alpha_1\phi_1 h_1 & 0 & \dots & \dots & 0 \\ \alpha_2\phi_1 h_1 & \alpha_2\phi_2 h_2 & 0 & \dots & 0 \\ \alpha_3\phi_1 h_1 & \alpha_3\phi_2 h_2 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \alpha_I\phi_1 h_1 & \alpha_I\phi_2 h_2 & \dots & \alpha_I\phi_{I-1}h_{I-1} & \alpha_I\phi_I h_I \end{bmatrix}$$

Solve the system of equations and iterate until v^{j+1} is close to v^j.

C.2 Step 1: Solution to Bellman Equation – Planning Problem

The Bellman equation for the planning problem (23), can be written as

$$(\rho-\theta\gamma)\,\omega(x) - \omega'(x)\gamma x = [1-\sigma(x)]x^{-\theta} + \alpha[\sigma(x)]S(x) + Q(x)$$

where

$$S(x) \equiv \int_0^x [\omega(y)-\omega(x)]\phi(y)dy = \int_0^x \omega(y)\phi(y)dy - \omega(x)\Phi(x)$$
$$Q(x) \equiv -\int_x^\infty \alpha[\sigma(y)][\omega(y)-\omega(x)]\phi(y)dy = -\int_x^\infty \alpha[\sigma(y)]\omega(y)\phi(y)dy + \omega(x)\psi(x)$$

$$\psi(x) \equiv \int_x^\infty \alpha(\sigma(y))\phi(y)dy$$

and the optimal choice $\sigma(x)$ is defined implicitly by the first-order condition

$$x^{-\theta} \geq \alpha'[\sigma(x)]S(x). \tag{44}$$

We use the same finite difference approximation as above, that is approximate $\omega(x)$ on a finite grid $x \in \{x_1, ..., x_I\}$. We again approximate the functions $S(x), \sigma(x)$ and $\alpha(\sigma(x))$ as in (42), and the

[13] This follows from rearranging the Bellman equation as

$$\left[\rho-\theta\gamma+\alpha_i^j\Phi_i+\frac{\gamma x_i}{h_i}\right]v_i^{j+1} - \frac{\gamma x_i}{h_i}v_{i+1}^{j+1} - \alpha_i^j\sum_{l=1}^{i} h_l\phi_l v_l^{j+1} = (1-\sigma_i^j)x_i^{-\theta}$$

and then rewriting it in matrix notation.

functions $Q(x)$ and $\psi(x)$ as

$$Q_i = Q(x_i) \approx -\sum_{l=i}^{N} \alpha_l \omega_l \phi_l h_l + \omega_i \psi_i, \quad \psi_i = \psi(x_i) \approx \sum_{l=i}^{N} \alpha_l \phi_l h_l \tag{45}$$

We again impose the boundary condition

$$0 \approx \omega'(x_I) = \frac{\omega_{I+1} - \omega_I}{h} \quad \Rightarrow \quad \omega_{I+1} = \omega_I.$$

We again proceed in an iterative fashion: we guess ω_i^0 and then for $j = 0, 1, 2...$ form ω_i^{j+1} as follows. Form S_i^j and Q_i^j as in (42) and (45), and obtain s_i^j and α_i^j from the first order condition (44). Write the Bellman equation as

$$\begin{aligned}(\rho - \theta\gamma)\omega_i^{j+1} &= (1 - \sigma_i^j)x_i^{-\theta} + \gamma x_i \frac{\omega_{i+1}^{j+1} - \omega_i^{j+1}}{h_i} \\ &+ \alpha_i^j \left[\sum_{l=1}^{i} \omega_l^{j+1} \phi_l h_l - \omega_i^{j+1} \Phi_i \right] - \sum_{l=i}^{N} \alpha_l^j \omega_l^{j+1} \phi_l h_l + \omega_i^{j+1} \psi_i^j \end{aligned} \tag{46}$$

Given ω^j and hence α^j and σ^j, this is again a system of I equations in I unknowns $(\omega_1^{j+1}, ..., \omega_I^{j+1})$ that we can solve for the value function at the next iteration ω^{j+1}. We again write (46) in matrix notation as

$$\mathbf{A}^j \omega^{j+1} = b^j, \quad b_i^j = (1 - \sigma_i)x_i^{-\theta}, \quad \mathbf{A}^j = \mathbf{B}^j - \mathbf{C}^j + \mathbf{D}^j$$

where[14]

$$
\mathbf{B}^j = \begin{bmatrix} \rho - \theta\gamma + \alpha_1^j \Phi_1 - \psi_1^j + \frac{\gamma x_1}{h_1} & -\frac{\gamma x_1}{h_1} & 0 & \dots & 0 \\ 0 & \rho - \theta\gamma + \alpha_2^j \Phi_2 - \psi_2^j + \frac{\gamma x_2}{h_2} & -\frac{\gamma x_2}{h_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & \rho - \theta\gamma + \alpha_I^j \Phi_I - \psi_I^j \end{bmatrix}
$$

$$
\mathbf{C}^j = \begin{bmatrix} \alpha_1\phi_1 h_1 & 0 & \dots & \dots & 0 \\ \alpha_2\phi_1 h_1 & \alpha_2\phi_2 h_2 & 0 & \dots & 0 \\ \alpha_3\phi_1 h_1 & \alpha_3\phi_2 h_2 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \alpha_I\phi_1 h_1 & \alpha_I\phi_2 h_2 & \dots & \alpha_I\phi_{I-1}h_{I-1} & \alpha_I\phi_I h_I \end{bmatrix}
$$

$$
\mathbf{D}^j = \begin{bmatrix} \alpha_1\phi_1 h_1 & \alpha_2\phi_2 h_2 & \alpha_3\phi_3 h_3 & \dots & \alpha_I\phi_I h_I \\ 0 & \alpha_2\phi_2 h_2 & \alpha_3\phi_3 h_3 & \dots & \alpha_I\phi_I h_I \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \alpha_{I-1}\phi_{I-1}h_{I-1} \\ 0 & 0 & \dots & 0 & \alpha_I\phi_I h_I \end{bmatrix}
$$

Solve the system of equations and iterate until ω^{j+1} is close to ω^j.

C.3 Step 2: Distribution Function

This section briefly describes the finite difference method used to compute the functions $\phi_{n+1}^j(x)$, $\Phi_{n+1}^j(x)$, $\psi_{n+1}^j(x)$ in Step 2a of the algorithm described in section 3. For notational simplicity, we suppress the dependence of these functions on n (the main iteration). We approximate these functions on a finite grid $(x_1, ..., x_I)$ of I values. We approximate the derivatives in (17) to (19) by

$$
(\phi^j)'(x_i) \approx \frac{\phi_{i+1}^j - \phi_i^j}{h_i}, \quad (\Phi^j)'(x_i) \approx \frac{\Phi_{i+1}^j - \Phi_i^j}{h_i}, \quad (\psi^j)'(x_i) \approx \frac{\psi_{i+1}^j - \psi_i^j}{h_i}
$$

[14]This follows from rearranging the Bellman equation as

$$
\left[\rho - \theta\gamma + \alpha_i^j \Phi_i - \psi_i^j + \frac{\gamma x_i}{h_i}\right] \omega_i^{j+1} - \frac{\gamma x_i}{h_i}\omega_{i+1}^{j+1} - \alpha_i^j \sum_{l=1}^{i} \phi_l \omega_l^{j+1} h_l + \sum_{l=i}^{I} \alpha_l^j \omega_l^{j+1} \phi_l h_l = (1 - \sigma_i^j) x_i^{-\theta}
$$

and then rewriting it in matrix notation.

so the finite difference approximation to (17) to (18) is

$$\phi_i^j \gamma + \gamma \frac{\phi_{i+1}^j - \phi_i^j}{h_i}, x_i = \phi_i^j \psi_i^j - \alpha(\sigma_i)\phi_i^j \Phi_i^j$$
$$\frac{\psi_{i+1}^j - \psi_i^j}{h_i} = -\alpha(\sigma_i)\phi_i^j$$
$$\frac{\Phi_{i+1}^j - \Phi_i^j}{h_i} = \phi_i^j$$

with boundary conditions

$$\phi_1^j = \lambda, \quad \Phi_1^j = 0, \quad \psi_1^j = \gamma_n^j.$$

This is a simple initial value problem which can simply be solved by running the system forward.

References

Alvarez, Fernando E., Francisco J. Buera, and Robert E. Jr. Lucas. 2008. "Models of Idea Flows." National Bureau of Economic Research, Inc NBER Working Papers 14135.

Arrow, Kenneth. 1962. "The Economic Implications of Learning by Doing." *Review of Economic Studies*, 29: 155–173.

Arrow, Kenneth J. 1969. "Classificatory Notes on the Production and Transmission of Technological Knowledge." *American Economic Review*, 59(2): 29–35.

Becker, Gary S, and Nigel Tomes. 1979. "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." *Journal of Political Economy*, 87(6): 1153–89.

Benabou, Roland. 2002. "Tax and Education Policy in a Heterogenous Agent Economy: What Levels of Redistribution Maximize Growth and Effiency?" *Econometrica*, 70(2): 481–517.

Benhabib, Jess, Alberto Bisin, and Shenghao Zhu. 2011. "The Distribution of Wealth and Fiscal Policy in Economies With Finitely Lived Agents." *Econometrica*, 79(1): 123–157.

Ben-Porath, Yoram. 1967. "The Production of Human Capital and the Life Cycle of Earnings." *Journal of Political Economy*, 75: 352.

Bental, Benjamin, and Dan Peled. 1996. "The Accumulation of Wealth and the Cyclical Generation of New Technologies: A Search Theoretic Approach." *International Economic Review*, 37(3): 687–718.

Candler, Graham V. 1999. "Finite-Difference Methods for Dynamic Programming Problems." In *Computational Methods for the Study of Dynamic Economies.* , ed. Ramon Marimon and Andrew Scott. Cambridge, England:Cambridge University Press.

Choi, Seung Mo. forthcoming. "How Large are Learning Externalities." *International Economic Review.*

Eaton, Jonathan, and Samuel Kortum. 1999. "International Technology Diffusion: Theory and Measurement." *International Economic Review*, 40(3): 537–70.

Eaton, Jonathan, and Samuel Kortum. 2002. "Technology, Geography, and Trade." *Econometrica*, 70(5): 1741–1779.

Fogli, Alessandra, and Laura Veldkamp. 2011. "Germs, Social Networks and Growth." NYU mimeo.

Gabaix, Xavier. 2009. "Power Laws in Economics and Finance." *Annual Review of Economics*, 1(1): 255–293.

Grossman, Gene M, and Elhanan Helpman. 1991. "Quality Ladders in the Theory of Growth." *Review of Economic Studies*, 58(1): 43–61.

Hause, John C. 1980. "The Fine Structure of Earnings and the On-the-Job Training Hypothesis." *Econometrica*, 48(4): 1013–29.

Heckman, James J. 1976. "A Life-Cycle Model of Earnings, Learning, and Consumption." *Journal of Political Economy*, 84(4): S11–44.

Jovanovic, Boyan, and Rafael Rob. 1989. "The Growth and Diffusion of Knowledge." *Review of Economic Studies*, 56(4): 569–82.

Jovanovic, Boyan, and Yaw Nyarko. 1996. "Learning by Doing and the Choice of Technology." *Econometrica*, 64(6): 1299–1310.

Kortum, Samuel S. 1997. "Research, Patenting, and Technological Change." *Econometrica*, 65(6): 1389–1420.

Lasry, Jean-Michel, and Pierre-Louis Lions. 2007. "Mean field games." *Japanese Journal of Mathematics*, 2: 229–260.

Lucas, Robert E. 2009. "Ideas and Growth." *Economica*, 76(301): 1–19.

Luttmer, Erzo G.J. forthcoming. "Technology Diffusion and Growth." *Journal of Economic Theory*.

Perla, Jesse, and Christopher Tonetti. 2011. "Endogenous Risk and Growth." NYU mimeo.

Romer, Paul M. 1986. "Increasing Returns and Long-run Growth." *Journal of Political Economy*, 94(5): 1002–37.

Rosen, Sherwin. 1976. "A Theory of Life Earnings." *Journal of Political Economy*, 84(4): S45–67.

Solon, Gary. 1992. "Intergenerational Income Mobility in the United States." *American Economic Review*, 82(3): 393–408.

Zimmerman, David J. 1992. "Regression toward Mediocrity in Economic Stature." *American Economic Review*, 82(3): 409–29.