From Inequality to Growth: The Role of Knowledge Creation

Kunal Dasgupta *
University of Toronto

November 2009

Abstract

In this paper, I revisit the classic question: how does inequality affect growth? I propose a new channel - knowledge creation - through which the relationship between growth and inequality could be affected. In the model, knowledge is created within firms through the interaction of agents with different abilities. An increase in inequality of the underlying distribution of abilities affects knowledge creation by changing how agents match within firms. I uncover two effects that determine how aggregate knowledge creation and consequently, the growth rate changes. A change in inequality increases the average ability of agents creating knowledge - the distribution effect. At the same time, there are fewer agents engaged in knowledge creation - the allocation effect. I examine the conditions under which one effect dominates the other. Using the model, I provide one possible explanation for Barro's (1993) finding that inequality reduces growth rate for less developed countries but increases the rate for developed countries.

KEYWORDS : Inequality, growth, knowledge creation, matching, team formation.

* *E-mail: kunal.dasgupta@utoronto.ca* I would like to thank Esteban Rossi-Hansberg and Gene Grossman for helpful comments.

1 Introduction

"Your best opportunity to come up with ideas is not to lock yourself alone in a room but to be out in the world, seeing what others are doing and use that in different ways."

Andrew Hargadon [1]

History credits James Watt with the invention of the steam engine, Thomas Alva Edison with the invention of the electric light bulb and Henry Ford with the invention of the assembly line. But these great inventors did not work alone; they had very able partners or teams assisting them in their endeavors. Watt had Matthew Boulton, Edison had Charles W. Batchelor while Ford had C. Harold Wills among others. Whether Watt or Edison or Ford would have been equally successful if they had worked on their own is hard to conjecture. It is true, however, that the above mentioned partners and many other individuals made significant contributions to the various innovations that have made Watt et al. household names.

In this paper, I develop a model to understand how the matching of agents within firms affects growth through its impact on knowledge creation. Agents in the model live for two periods. In the first period, they acquire ability which depends partly on government policy or institutions. In the second period, agents choose a sector to work in. The economy has two sectors - final goods and intermediate goods - that vary in three important respects. First, the intermediate goods sector is more ability-intensive. Second, production in the intermediate goods sector takes place in teams and how agents match matters. And finally, firms producing intermediate goods also create new knowledge every period; this knowledge spills over to the final goods sector in the following period, thereby raising the total factor productivity in final good production. Furthermore, agents choosing to work in the intermediate goods sector also choose their occupation. They could either be innovators, developing blueprints for new products, or managers who use the blueprints to create output. The final goods sector only requires workers.

An equilibrium of the model is characterized by two ability thresholds, where the agents with ability below the lower threshold sort into the final goods sector, while those above the higher threshold become innovators. The ones in between choose to be managers. The ability distribution not only determines the thresholds, but by affecting how agents match, also determines how much knowledge is created by each firm. I go on to study how an increase in inequality in the ability distribution affects the growth rate. A change in the distribution, by changing how agents match, affects knowledge creation within each firm. The effect on growth, however, depends on aggregate knowledge in the economy.

I define a very general aggregate knowledge function, which nests some of the more popular knowledge spillover functions used in the literature. In the case of the canonical growth model, which I briefly review in Section 2, an increase in inequality is accompanied by a decline in the growth rate. If the growth rate is determined by the ability of the best agents as in Murphy et al. (1991), an increase in inequality could lead to a higher growth rate. I show that whether an increase in inequality raises or lowers the growth rate

[1] "Andrew Hargadon talks about green initiatives during keynote address", Design News, August 8, 2008.

depends crucially on weights that capture the extent of overlap between knowledge created by individual firms/innovators. A more unequal ability distribution has fewer innovators which means less knowledge creation (*allocation effect*). But the innovators are, on average, more able which means more knowledge creation (*distribution effect*). If the weights are a sufficiently convex function of ability, the latter effect dominates, leading to a rise in the growth rate.

The mechanism proposed in the paper provides a justification for the finding in Barro (2000) that the effect of inequality on economic growth depends on the level of economic development. Barro finds that the effect of inequality on growth is negative for countries with low per capita GDP and positive for the rest. The model generates the above prediction if the weights evolve systematically as a country develops. In particular, if the knowledge created by the more able innovators becomes relatively more important as a country grows, we should expect to see the relation between growth and inequality varying across countries at different stages of development.[2]

The model is built on the presumption that knowledge generated by a firm depends on the ability of the firm's employees. In a related paper, Lucas (2009) develops a model where the aggregate knowledge of an economy is simply a list of the knowledge of its members. Similar to my model, a fundamental assumption of Lucas (2009) is that knowledge is embodied. Lucas does not, however, consider team formation or the problem of allocating agents across sectors, as I do here. In my model, both allocation, as well as, the composition of firms matters for growth. In that respect, this paper is similar to Murphy et al. (1991), who look at how allocating individuals with different levels of ability across different activities can affect growth. Murphy et al. consider two main activities, production and rent-seeking, and consider the implications of allocating the most able agent in either of these two activities.[3] They do not consider team formation, however. Since growth in their model is driven by the ability of the best production worker, inequality has no effect on growth as long as the most able agent in the economy is in the production sector. I show that if knowledge creation depends on team formation, then a change in inequality can affect the growth rate.

Economists have been studying the relation between inequality and growth for a long time. The classical view was that inequality is good for growth. This could be either because investment projects involve large indivisibilities or, as pointed out by Kaldor (1955-56), because the rich have a higher propensity to save compared to the poor (See Aghion et al., 1999, for a discussion of these and other causes). As evidence suggesting a negative relation between inequality and growth was unearthed in the last two decades,

[2] The empirical literature has failed to reach a consensus on the relationship between inequality and growth. While studies based on ordinary least squares analyses of cross-country data have typically found a negative relation between inequality and growth (See, for example, Alesina and Rodrik, 1994; Perotti, 1996; Persson and Tabellini, 1994), studies based on fixed-effects estimation yield a positive relation (See, for example, Benhabib and Spiegel, 1997; Forbes, 2000). As pointed out by Banerjee and Duflo (2003), however, the above mentioned studies are not comparable; not only are the specifications quite different across these studies, they also use different data sets. Banerjee and Duflo also point out out that these two opposite findings are not necessarily inconsistent with each other.

[3] Murphy et al. (1991) do not have an explicit dynamic model where the production and rent-seeking sectors evolve endogenously. Ehrlich and Lui (1999) develop an endogenous growth model where growth results from the interaction of two activities - accumulating physical capital which promotes growth, and accumulating political capital, which promotes rent-seeking.

theories have been forwarded that predict a negative causal link between inequality and growth. These theories either emphasize the absence of well-functioning credit-markets (Banerjee and Newman, 1991; Galor and Zeira, 1993; Aghion and Bolton, 1997) or introduce a political process through which agents choose the level of re-distribution in the economy (Alesina and Rodrik, 1994; Persson and Tabellini, 1994).[4] None of these papers consider the effect of inequality on knowledge creation.

The remainder of the paper is organized as follows. In Section 2, I briefly review the canonical growth model and its treatment of inequality. In Section 3, I introduce heterogeneity and matching. The equilibrium is characterized in Section 4. I examine the effect of inequality on growth in Section 5. Section 6 concludes.

2 A Canonical Growth Model

Let us consider a simple version of the endogenous growth model of Romer (1990). The analysis is based on Jones (2005). Time is discreet and labor is the only factor of production. At time t, the economy produces a final good Y_t, let us call it GDP, according to the following technology :

$$Y_t = K_t^{\sigma} L_{Y,t}$$

where K_t is the stock of knowledge in the economy while $L_{Y,t}$ is the amount of labor employed in production at time t. Labor can also be employed in research, leading to the creation of new knowledge. Knowledge is produced according to the following linear production function :

$$K_{t+1} = e^{\delta L_{K,t}} K_t$$

where $L_{K,t}$ is the amount of labor employed in research at time t. In the absence of population growth, the resource constraint of the economy is

$$L_{Y,t} + L_{K,t} = L$$

where L is the size of the work force. Suppose a constant fraction χ of the labor force is allocated towards research in each period, (such a choice can be derived from a micro-founded model where χ would depend on preference and institutional parameters) i.e.,

$$L_{K,t} = \chi L \;\text{ and }\; L_{Y,t} = (1-\chi)L$$

Define the growth rate of per capita income $y_t (= Y_t/L)$ as g. Then it follows that,

[4] See Galor (2011) for a summary of these two main lines of work.

$$\begin{aligned} g &= log(y_{t+1}) - log(y_t) \\ &= \sigma(log(K_{t+1}) - log(K_t)) \\ &= \sigma\delta\chi L \end{aligned}$$

The growth rate of per capita output is simply the growth rate of knowledge, augmented by σ which measures the sensitivity of output to knowledge; how a country grows essentially depends on how knowledge is created in the country. As shown above, g depends on the the total number of workers in the economy. This is the standard scale effect in these class of models.[5] A bigger population means more researchers, which translates into greater knowledge creation and faster growth of output. A direct implication is that other things being the same, two countries with the same size of the labor force should grow at the same rate.

It is clear that in the above model, inequality can affect growth rate only through χ; all the other parameters are exogenous. *A priori* we do not have any idea of how, following an increase in inequality, χ would change. We do know, however, that an increase in inequality can have a positive impact on the growth rate only if there is an increase in the number of researchers; if an increase in inequality leads to a decline in χ, growth rate of the economy would go down because of fewer researchers in the economy.[6] Since in these models researchers are homogeneous in terms of their research productivity, only the mass of researchers matters for growth. In the following section, I present a model in which knowledge creation depends not only on the number of researchers in an economy, but also on how they match within firms.

3 A model with heterogeneous agents

In this section, I lay out the basic economic environment of the model with heterogeneous agents.

3.1 Preferences and Ability

Consider a two period overlapping generations model populated by a continuum of agents. Every agent has a parent and a child; thus, population is constant. An agent also owns one unit of labor which he supplies inelastically in the second period. Since the measure of agents does not play any role in my analysis, I normalize it to 1.

For simplicity, I assume that agents consume only in the second period. They also leave bequests for their children. The utility function is given by

[5] Apart from Romer (1990), other papers that predict this *strong* scale effect include Grossman and Helpman (1991) and Aghion and Howitt (1992) among others. Failing to find evidence for strong scale effects in 20th century U.S., Jones (1999) proposes a model with *weak* scale effects, where the growth rate of the economy depends on the growth rate of the population and not its level.

[6] If there is a difference between the workforce L and the population and the participation of agents in economic activity depends on the distribution of wealth, then a change in inequality might affect the growth rate by changing L.

$$U = c^{\rho} b^{1-\rho} \tag{1}$$

where c is consumption, b is bequest and $0 < \rho < 1$ captures the preference of agents for their own, rather than their children's, well-being. The bequest is the only wealth which a newborn agent inherits.

In the first period, an agent acquires education that determines his ability in period two, a, where $a \in [\underline{a}, \overline{a}]$ with $\underline{a} > 0$ and $\overline{a} \leq 1$. This ability depends on (i) his inherited wealth, b, and (ii) a parameter Γ that captures policy, institutions or market frictions. For example, a change in Γ could be interpreted as an expansion of primary education at the cost of tertiary education. Γ could also be the state of credit market which, given the distribution of inherited wealth, determines how agents acquire education and hence their ability. I assume that the determination of ability involves a stochastic component which allows me to write the conditional distribution of abilities as $G(a|b, \Gamma)$. I make the following assumptions about G :

(A1) $G(a|b, \Gamma)$ has full support for all values of b.

(A2) For all $b' > b$, $G(a|b', \Gamma)$ dominates $G(a|b, \Gamma)$ in the sense of first-order stochastic dominance.

Assumption A1 implies that no matter how large (small) a wealth an agent inherits, he could end up with very low (high) ability in period two. But at the same time, higher is the wealth, more likely it is that the agent will have higher ability (Assumption A2). Denoting an agent's income in the second period as $y(a)$, the preferences given by (1) imply that $b = (1 - \rho)y(a)$. This allows me to write,

$$\begin{aligned} G(a_{t+1}|b_{t+1}, \Gamma) &= G(a_{t+1}|(1-\rho)y_t(a_t), \Gamma) \\ &= \tilde{G}_t(a_{t+1}|a_t, \Gamma) \end{aligned} \tag{2}$$

where the dependence of income $y(a)$ on t is due to general equilibrium effects - an agent's income depends not only on his own ability but also on the ability of others. This will become clear in the next section. Let the distribution of abilities at time t be denoted by $\Phi_t(a)$. Using (2), I can write

$$\Phi_{t+1}(a) = \int_{\underline{a}}^{\overline{a}} \int_{\underline{a}}^{a} \tilde{G}_t(s|a_t, \Gamma) ds d\Phi_t(a) \tag{3}$$

The above equation represents the law of transition of the ability distribution. The distribution of abilities at t determines production in the economy and the income of every agent. Then, through the bequests, the distribution of abilities at $t + 1$ is determined. We turn to production next.

3.2 Production and Occupation

The economy consists of two sectors - a final goods (F) sector and an intermediated goods (M) sector. The F sector produces Y_t while the M sector produces differentiated intermediate inputs that are used in the production of the final good.[7]

3.2.1 Final goods sector

The final good Y_t is produced competitively using the following technology :

$$Y_t = K_t L_{Y,t}^{1-\alpha} \int_{\Omega_t} x_t(\omega)^\alpha d\omega, \tag{4}$$

where K_t, as before, is the stock of knowledge at time t, $L_{Y,t}$ is the labor employed in sector F, $x_t(\omega)$ is the amount of the differentiated input ω and Ω_t is the set of inputs that are produced at time t. The set of inputs that can be potentially produced is un-bounded. The inputs could be goods that get used up in production or machines that depreciate fully. (See Acemoglu, 2008). Equation (4) can, alternatively, be written as

$$Y_t = K_t L_{Y,t}^{1-\alpha} X_t^\alpha,$$

where

$$X_t = (\int_{\Omega_t} x_t(\omega)^\alpha d\omega)^{\frac{1}{\alpha}}.$$

The alternative form shows that the production technology in sector F exhibits constant-returns-to-scale in labor and X_t, which is a CES aggregator of the differentiated inputs $x_t(\omega)$. Workers in sector F perform routine tasks; what matters for production is the mass of workers rather than their ability. As a result, all workers in this sector earn an uniform wage s_t.

Each differentiated goods firm is a monopolist and sets a price of $p_t(\omega)$ for its inputs. Every period, a typical sector F firm hires labor and purchases inputs in the spot market, carries out production and makes payment to the factors which exhausts total output. Since the firm does not make any inter temporal allocation decisions, it simply maximizes instantaneous profits. The problem of a F firm is

$$\max_{L_{Y,t}, x_t(\omega)} K_t L_{Y,t}^{1-\alpha} \int_{\Omega_t} x_t(\omega)^\alpha d\omega - s_t L_{Y,t} - \int_{\Omega_t} p_t(\omega) x_t(\omega) d\omega. \tag{5}$$

The first-order condition with respect to labor gives rise to the following expression for s_t :

$$s_t = (1-\alpha) K_t L_{Y,t}^{-\alpha} (\int_{\Omega_t} x_t(\omega)^\alpha d\omega). \tag{6}$$

More inputs raises the marginal productivity of the sector F workers and thereby increases their wage. The first-order condition with respect to input $x_t(\omega)$ yields the following iso-elastic demand curves :

[7] Hence, we are following Romer (1990) in making the intermediate input a differentiated good. The alternative, following Ethier (1982) and Grossman and Helpman (1991), is to assume that agents have a love-for-variety utility function and consequently, the final good is differentiated.

$$x_t(\omega) = (\alpha K_t / p_t(\omega))^{\frac{1}{1-\alpha}} L_{Y,t}^{1-\alpha}.$$

The demand curve can be inverted to obtain an expression for the price of the intermediate inputs :

$$p_t(\omega) = \alpha K_t L_{Y,t}^{1-\alpha} x_t(\omega)^{\alpha-1}. \tag{7}$$

Note that the only dependence of input price $p_t(a)$ on the aggregate knowledge distribution is through the allocation of labor across the two sectors. In particular, it does not depend on how much of the other inputs are being produced.

3.2.2 Intermediated goods sector

Each M firm consists of a manager and an innovator, each of whom performs a separate task. The innovator develops a blueprint. The manager uses the knowledge embodied in a blueprint to build equipments/machines that produce the intermediate good ω. Since the set of inputs is un-bounded, each firm chooses to produce a different input. The innovator pays a wage to the manager and is the residual claimant of output.

Creation of new knowledge (development of a blueprint) is not deterministic. The probability that an innovator is successful depends on his ability. Let us assume that this probability is simply a_I, the ability of the innovator in question. Upon successful development, the manager tries to follow the instructions in the blueprint to create a machine.[8] The ability of the manager determines how well he can implement the complex instructions contained in the blueprint. I assume that conditional on there being a blueprint, a manager with ability a_M builds a machine with productivity a_M^β, where $\beta < 1$. Therefore, I assume that innovation is more sensitive to ability relative to implementation. Success in innovation is almost all about ability or talent whereas implementation requires both ability as well as effort. Although effort is partly endogenous, assuming that innovation is more *ability-intensive* seems reasonable.[9]

Given the above structure for knowledge creation and implementation, the expected output of a M firm is given by

$$x_t(a_I, a_M) = a_I a_M^\beta$$

The innovator is successful with probability a_I, whereby a_E^β units of intermediate input is produced. When the innovator fails to develop a blueprint, there is no production. Given x_t and the price of intermediate inputs given by equation (7), the innovator chooses a manager to maximize his profits,[10]

[8] In a recent paper, Meisenzahl and Mokyr (2011) argue that the industrial revolution in Britain was the outcome of effort by not only great innovators, but also workmen who carried out the instructions in the blueprints.

[9] Under these assumptions, output is more sensitive to the ability of the innovator. The property rights theory of Grossman and Hart (1986) and Hart and Moore (1990) then suggests that residual rights to output should lie with the innovator.

[10] Since every firm chooses to produce a different intermediate input, we can represent x_t as a function of a_I and a_M, rather than ω.

$$\begin{aligned}\pi_t(a_I) &= p_t(x_t)x_t(a_I, a_M) - w_t(a_M) \\ &= \alpha K_t L_{Y,t}^{1-\alpha}[x_t(a_I, a_M)]^\alpha - w_t(a_M)\end{aligned} \tag{8}$$

where $w_t(a)$ is the wage schedule for managers and we have replaced $p_t(x_t)$ with its expression from equation (7). The first-order condition for profit-maximization with respect to a_M yields

$$w_t'(a_M) = \alpha^2 K_t L_{Y,t}^{1-\alpha}[x_t(a_I, a_M)]^{\alpha-1}\frac{\partial x_t}{\partial a_M}. \tag{9}$$

At the margin, the change in revenue coming from hiring a slightly more able manager should be equal to the change in cost due to the higher wage that has to be paid to the manager.[11] The envelope condition yields

$$\pi_t'(a_I) = \alpha^2 K_t L_{Y,t}^{1-\alpha}[x_t(a_I, a_M)]^{\alpha-1}\frac{\partial x_t}{\partial a_I}. \tag{10}$$

The above two equations suggest that the slope of the wage and the profit functions depend on the sign of $\frac{\partial x_t}{\partial a_I}$ and $\frac{\partial x_t}{\partial a_M}$, both of which, it can be easily checked, are positive. Hence, within an occupation, the earning of an agent increases with his ability. Furthermore, $\frac{\partial^2 x_t}{\partial a_I \partial a_M} > 0$ - the production function for intermediate inputs exhibits complementarity with respect to a_M and a_I. A direct implication of complementarity is that equilibrium is characterized by positive assortative matching or PAM as stated formally in the following lemma.

Lemma 1. *The equilibrium matching function in sector M, $m_t(a_M)$, is monotonic increasing in a_M.*

The above lemma allows us to write $x_t(a_I, a_M) = x_t(a_I, m_t^{-1}(a_I)) = x_t(a_I)$. In equilibrium, production of intermediate inputs is simply a function of the ability of the innovator producing that input. More able innovators hire more able managers and also earn higher profits relative to less able innovators. Note that this result arises out of the optimization behavior of innovators and does not, in any way, depend on the shape of $\Phi_t(a)$, the ability distribution at time t.

3.2.3 Occupational choice

In each period, the agents face the following occupational choices : (1) Being a worker in sector F, (2) being a manager in sector M, and (3) being an innovator in sector M. An agent chooses the occupation which generates the highest income, i.e., the problem facing an agent with knowledge a is $max\{s_t, w_t(a), \pi_t(a)\}$. Labor market clearing in each period requires that

[11] Note that this condition is slightly different from the standard profit maximization condition that equates wage with marginal product of labor. In this model, since every firm hires only one manager, marginal product of labor has no meaning; what matters is the marginal returns to ability.

$$L_{Y,t} + \int\limits_{Ma_t} d\Phi_t(a) + \int\limits_{In_t} d\Phi_t(a) = 1. \quad (11)$$

where Ma_t and In_t denotes the set of managers and innovators respectively.

3.3 Evolution of Knowledge

Every firm in sector M creates new knowledge. I make two assumptions about knowledge creation. First, knowledge created within a firm in period t spills over to the rest of the economy in period $t+1$. A firm developing a blueprint can use patents to exclude others from using it. I consider two scenarios. In the first scenario, firms use patents but they are effective for only one period; in the period following the development of a blueprint, the blueprint is publicly available. In the second scenario, firms do not use patents for fear of disclosing their knowledge. In this case, the product of the firm "partially" reveals the knowledge created by the firm. The non-rival nature of knowledge implies that this knowledge can be costlessly used by sector F firms, thereby increasing TFP in final good production.

Second, new knowledge adds to the stock of existing knowledge; I rule out obsolescence of knowledge. I continue to assume that the stock of aggregate knowledge evolves linearly with the difference being that $L_{K,t}$ in equation (2) is now replaced by a function that depends on the entire knowledge distribution through the matching of innovators with managers. As the distribution changes, the matches change too, thereby affecting the amount of knowledge that is created and eventually spills over. This, in turn, affects growth. I assume that K_t has the following law of motion

$$K_{t+1} = e^{\lambda_t} K_t, \quad (12)$$

where $\lambda_t : \Phi_t(k) \to \Re_+$ is an aggregation of new knowledge that spills over. At this point, I do not make any further assumptions about λ_t and defer my discussion of λ_t till Section 5.

4 Equilibrium

Now that the basic structure of the model has been laid out, an equilibrium can be defined.

Definition 1. *An equilibrium in this economy consists of the following objects :*
(i) allocation of agents to different occupations,
(ii) time paths of sector F wages and employment, $\{s_t, L_{Y,t}\}_{t=0}^{\infty}$,
(iii) time paths of quantities and prices of differentiated inputs, $\{x_t(a), p_t(a)\}_{t=0}^{\infty}\big|_{k\in In_t}$,
(iv) time paths of sector M wage schedule, $\{w_t(a)\}_{t=0}^{\infty}\big|_{a\in Ma_t}$ *and profit schedule,* $\{\pi_t(a)\}_{t=0}^{\infty}\big|_{a\in In_t}$,
(v) time path of the matching function, $\{m_t(a)\}_{t=0}^{\infty}\big|_{a\in Ma_t}$,
(vi) law of motion for knowledge $\{K_t\}_{t=0}^{\infty}$,

such that (a) agents maximize utility, (b) firms maximize profit, (c) labor market clears and (d) knowledge evolves according to equation (12).

4.1 Characterization of equilibrium

In the previous section, I had established that the manufacturing sector exhibits PAM. The implication is that if there is an interval $[a_1, a_2]$ of managers, then the corresponding innovators will lie in the interval $[m_t(a_1), m_t(a_2)]$, where $m_t(a_1) > a_2$. There are numerous allocations that are consistent with PAM. Before proceeding any further, we impose the following restriction on the parameter space:

(A3) $\underline{a} > \beta^{\frac{1}{\alpha}}\overline{a}$.

Given assumption (A3), not all allocations are consistent with equilibrium, as shown in the following lemmas.

Lemma 2. *In equilibrium, the sets of workers in sector F are connected and is of the form $[\underline{a}, a_t^*]$, where a_t^* is endogenous.*

Lemma 3. *In equilibrium, the sets of managers and innovators in sector M are connected.*

Lemma 2 says that the least able agents select into F sector. The agents with low ability have a comparative advantage in sector F where ability is not rewarded. The intuition behind Lemma 3, on the other hand, is a bit more subtle. There are two opposing forces that determine matching in sector M - complementarity, which encourages agents with similar abilities to match together and asymmetry, which goes in the opposite direction (Kremer and Maskin, 1996). Given the functional forms and assumption (A3), the latter effect dominates, whereby managers and workers with different abilities match.

Lemmas 2 and 3 together imply that the equilibrium allocation is characterized by two thresholds, a_t^* and a_t^{**}, such that agents with ability $\in [\underline{a}, a_t^*]$ are in sector F, while those with ability $\in [a_t^*, a_t^{**}]$ and ability $\in [a_t^{**}, \overline{a}]$ are managers and innovators in sector M respectively. Figure 1 shows a possible equilibrium allocation.

The agent with knowledge a_t^* must be indifferent between being a worker in sectors F and M, i.e., a_t^* must solve

$$s_t = w_t(a_t^*). \tag{13}$$

On the other hand, one-to-one matching in sector M implies that a_t^{**} must solve the following :

$$\frac{1}{1-\Phi_t(a_t^*)}\int_{a_t^*}^{a_t^{**}} \phi_t(a)dk = \frac{1}{2},$$

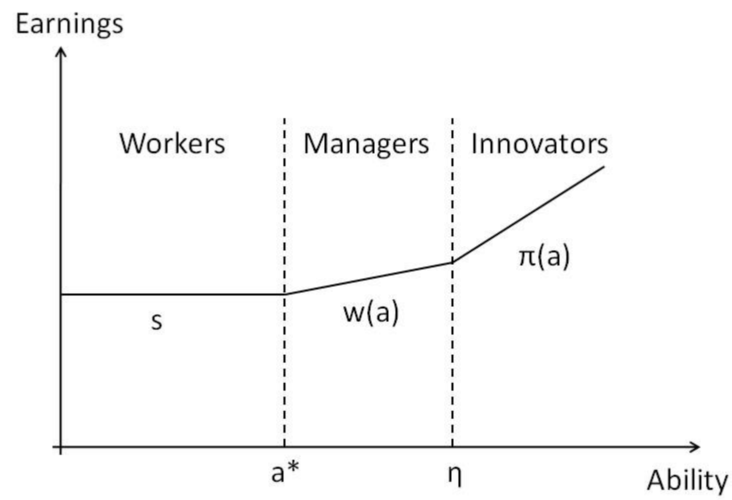

Figure 1: *An equilibrium allocation of agents*

i.e., a_t^{**} is the median of the truncated distribution lying between a_t^* and $\overline{a}$. Denote $a_t^{**} = \eta(a_t^*)$. Clearly, $\frac{\partial \eta(a_t^*)}{\partial a_t^*} > 0$. Henceforth, we drop the argument and write $\eta(a_t^*)$ simply as η_t. Conditional on a_t^*, the labor market clearing condition in sector M is given by

$$\int_{a_t^*}^{a} \phi_t(s)ds = \int_{\eta_t}^{m_t(a)} \phi_t(s)ds \ \ \forall a_t^* \leq a \leq \eta_t. \tag{14}$$

The left-hand side of the above equation denotes the supply of managers in the interval $[a_t^*, a]$ for some a. The right-hand side denotes the demand for managers coming from innovators in the corresponding interval $[\eta_t, m_t(a)]$. PAM implies that these two measures must be equal for *every* a. It is not enough for the above equation to be satisfied only for $a = \eta_t$, as would be the case in a standard Lucas span-of-control model without matching. Differentiating the above equation with respect to a, we have

$$m_t'(a) = \frac{\phi(a)}{\phi(m_t(a))}. \tag{15}$$

The above differential equation is intuitive. The numerator measures the density of the managers with knowledge a, while the denominator measures the density of the corresponding innovators. Loosely speaking, if the numerator is small relative to the denominator, the managers are more dispersed relative to the innovators in terms of abilities. Then measure consistency would require that managers with very different ability be matched with innovators with very similar ability. Consequently, $m_t(a)$ would increase slowly, i.e., the slope of the matching function would be small.

For a given a_t^*, the terminal conditions (i) $m_t(a_t^*) = \eta_t$ and (ii) $m_t(\eta_t) = \overline{a}$ can be used to solve equation (15) and obtain the matching function $m_t(a)$. Replacing the equilibrium matching function $m_t(a)$

in equation (9) and integrating, we have

$$w_t(\eta_t) - w_t(a_t^*) = \alpha^2 K_t L_{Y,t}^{1-\alpha} \int_{a_t^*}^{\eta_t} [x_t(m_t(a))]^{\alpha-1} \frac{\partial x_t}{\partial a} da. \tag{16}$$

Furthermore, the agent with knowledge η_t must be indifferent between being a worker or a manager in the sector M, i.e., $w_t(\eta_t) = \pi_t(\eta_t)$. Replacing $\pi_t(\eta_t)$ from equation (8) yields

$$w_t(\eta_t) = \alpha K_t L_{Y,t}^{1-\alpha} [x_t(\eta_t)]^\alpha - w_t(a_t^*). \tag{17}$$

Combining equations (16) and (17), we have

$$w_t(a_t^*) = \frac{\alpha K_t}{2} L_{Y,t}^{1-\alpha} ([x_t(\eta_t)]^\alpha - \alpha \int_{a_t^*}^{\eta_t} [x_t(m_t(a))]^{\alpha-1} \frac{\partial x_t}{\partial a} da). \tag{18}$$

Solving for the equilibrium requires solving a fixed-point problem. For a particular value of K_t, the equilibrium can be solved by carrying out the following iterations :

1. Guess a value for a_t^*, which in turn determines $L_{Y,t}$ (since $L_{Y,t} = \Phi_t(a_t^*)$).
2. Solve for $m_t(a)$. Sector M has a block recursive structure which allows me to solve for the matching function without any information on prices and earnings. (See Dasgupta, 2010, for more on this).
3. Use equation (6) to solve for s_t.
4. Use equation (18) to solve for $w_t(a_t^*)$.

If s_t and $w_t(a_t^*)$ thus computed are equal, then we are done. Otherwise, we just guess a different value for a_t^* and repeat steps 2-4. The following proposition shows that such iterations necessarily converge.

Proposition 1. *An equilibrium exists with a threshold a_t^* such that $s_t = w_t(a_t^*)$.*

Finally, the growth rate of the economy is given by

$$g = ln(Y_{t+1}) - ln(Y_t). \tag{19}$$

This completes the characterization of the equilibrium.

4.2 Balanced Growth Path

Along a Balanced Growth Path (BGP), the growth rate of income is constant. As discussed in the previous section, given K_t, a_t^* is a sufficient statistic for the equilibrium. But a_t^* solves equation (13). Notice that K_t appears in the same form on both sides of this equation. Consequently, at any point in time, a_t^* is *independent* of the level of knowledge in the economy. Intuitively, a higher K_t not only raises sector F wages, but by raising the demand for intermediate inputs, also raises sector M wages. Since K_t enters the production function of Y_t multiplicatively, the increase in the two wages cancel each other out, keeping the threshold unchanged. The threshold, a_t^*, depends only on the distribution of abilities at time t, $\Phi_t(a)$.

Proposition 2. *There exists a unique stationary distribution of abilities $\Phi(a)$ such that any initial distribution $\Phi_t(a)$ converges to it.*

Once the stationary distribution is reached, the threshold a_t^* becomes equal to a constant a^*. Hence all variables except the stock of knowledge K_t are also constant. In particular, λ_t, which captures the knowledge spillover in equation (12) is also a constant λ. The growth rate is then given by equation (19),

$$\begin{aligned} g &= ln(A_{t+1}) - ln(A_t) \\ &= \lambda \end{aligned}$$

Once the distribution of abilities converges to the stationary distribution, the economy move along a BGP, growing at the rate λ. For the rest of the paper, we focus attention on the stationary distribution. A change in the distribution changes λ. In the next section, we explore how a change in inequality affects the growth rate.

5 Inequality and Growth

One of the fundamental parameters in this model is Γ which could capture government policy or the quality of institutions. From equation (3) it follows that the distribution of abilities at time t and accordingly, the stationary distribution depends on Γ. Without analyzing in detail how a change in Γ might affect $\Phi(a)$, I simply assume that the government can change policy in such a way that $\Phi(a)$ displays greater inequality.

How does higher inequality affect growth? A change in $\Phi(a)$ affects the creation of knowledge by changing how agents match. As knowledge spills over to the final goods sector, the incentives of the agents change too. This, in turn, affects the matches. To analyze the effect of inequality on growth, I need to precisely define a change in inequality.

Consider two distributions $\Phi_A(a)$ and $\Phi_B(a)$, which differ only in terms of the degree of inequality. I shall assume that $\Phi_B(a)$ is more unequal than $\Phi_A(a)$ in the sense of second-order stochastic dominance. Since I want to compare countries at the same level of development, a mean-preserving spread seems like a more reasonable assumption rather than some arbitrary difference in the two distributions. The only restriction I impose on the shape of $\Phi_A(a)$ is that the distribution is downward-sloping, at least in the upper tail. For illustrative purpose, I consider the simplest form of a mean-preserving spread. In particular, I choose t_1, t_2, s_1, s_2 as shown in the following diagram.

As indicated in Figure (2), the knowledge distribution in country B, $\Phi_B(k)$, is created by making the following changes to $\Phi_A(k)$:

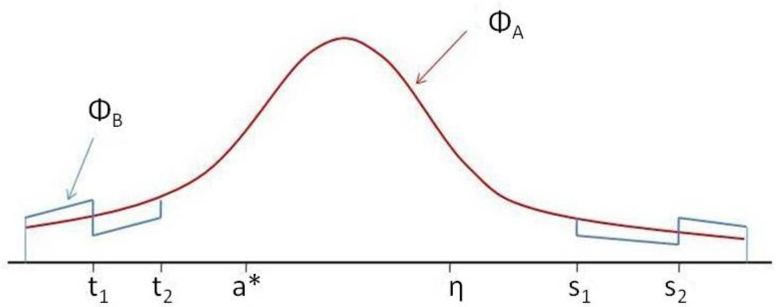

Figure 2: *Second-order stochastic dominance*

$$\phi_B(a) = \begin{cases} \phi_A(a) + \epsilon & \text{for } a \in [\underline{a}, t_1] \\ \phi_A(a) - \epsilon & \text{for } a \in [t_1, t_2] \\ \phi_A(a) - \epsilon & \text{for } a \in [s_1, s_2] \\ \phi_A(a) + \epsilon & \text{for } a \in [s_2, \overline{a}] \end{cases} \tag{20}$$

The effect of higher inequality is asymmetric across the two sectors. Since abilities do not matter for production in sector F, there is no direct effect of a rise in inequality. Production in sector M, on the other hand, depends on the abilities of the agents. A higher inequality, as defined above, implies that there are relatively fewer low ability innovators under B. The relative scarcity of low ability innovators under B implies that their earnings are high relative to the more able innovators and managers. When inequality goes up, the managers are made relatively worse-off, forcing the least able among them to switch to sector F. The following proposition, which lies at the heart of my analysis, shows the final effect of a rise in inequality on labor allocation across the different occupations.

Proposition 3. *A rise in inequality increases a^*, the threshold between sectors F and M.*

A corollary of the above proposition is that η rises as inequality increases. An increase in inequality raises the proportion of more able innovators. PAM implies an increase in demand for the more able managers resulting in their wages going up. Therefore, the opportunity cost of the least able innovators goes up, with some of them switching occupation.

Corollary 1. *A rise in inequality increases the ability of the best sector M manager, η.*

An increase in inequality not only changes the occupation of agents but by changing the underlying ability distribution, also changes how agents match in sector M. This change in matches affects how much knowledge is produced in each firm, and how much of this knowledge spills over. Recall from Section 4, that the steady-state growth rate is given by λ, which is a measure of aggregate knowledge. For distribution $i = \{A, B\}$, I now define λ_i as follows :

$$\lambda_i = \int_{\eta_i}^{\overline{a}} \omega_i(a) k(a, m_i^{-1}(a)) \phi_i(a) da. \tag{21}$$

where $k(a, m_i^{-1}(a))$ is the knowledge created by each firm, with $\frac{\partial k}{\partial a} > 0, \frac{\partial k}{\partial m_i^{-1}(a)} \geq 0$. $\omega_i(a) \geq 0$ is a weight attached to knowledge created by each firm. $\omega_i(a)$ captures the potential "overlap" between knowledge created by different firms as I elaborate below. Before analyzing the general case, I consider two special cases that will help us better understand equation (21).

Example 1 (Canonical growth model) : Setting $\omega_i(a) = \frac{1}{k(a, m_i^{-1}(a))}$ in equation (21), the growth rate of the economy becomes,

$$\begin{aligned} \lambda_i &= \int_{\eta_i}^{\overline{a}} \phi_i(a) da \\ &= (1 - \Phi_i(\eta_i)) \end{aligned}$$

The growth rate depends only on the fraction of innovator in the economy (or the mass of innovators since we have set population to 1). According to equation (20), $\Phi_A(\eta_A) = \Phi_B(\eta_A)$. Corollary (1) then implies that $\Phi_A(\eta_A) < \Phi_B(\eta_B)$. Therefore, the more unequal distribution has fewer innovators and consequently, lower growth. This result is similar to the one obtained in Section 2, where growth rate was a function of the size of the population.

How do we interpret the assumption that $\omega_i(a) = \frac{1}{k(a, m_i^{-1}(a))}$? This assumption says that the knowledge generated by firms owned by more able innovators do not contribute more to aggregate knowledge. This could due to the "fishing in the pond" effect. If the set of knowledge is bounded and invention or innovation amounts to discovery of such knowledge, then knowledge creation independently by many firms would involve lots of duplications. In this particular case, a firm that creates a (expected) knowledge of $\frac{1}{2}$ has more duplications than a firm that creates knowledge of $\frac{1}{3}$ in a way that both firms end up creating the same amount of *unique* knowledge.[12] This result might seem surprising, given that the average ability of the innovator in the unequal distribution is higher. The ability of the innovators, however, does not really matter for aggregate knowledge. Although a more unequal knowledge distribution has innovators who are, on average, more able, the more able among them actually drive out the less able through general equilibrium effects (in this model, the latter become managers) with the result that there are fewer agents innovating in equilibrium.

Example 2 (Murphy, Shleifer, Vishny (1990) : Consider the following functional form for $\omega_i(a)$:

[12] Note that an additional unit of an intermediate good has a positive marginal product, even if this unit is identical to other units. The same is not true for knowledge.

$$\omega_i(a) = \begin{cases} 0 & \forall a < \overline{a} \\ \frac{1}{\phi(a)} & \text{if } a = \overline{a} \end{cases}$$

Given these weights, the growth rate of the economy is given by

$$\lambda_i = k(\overline{a}, \eta_i).$$

Although every sector M firm creates knowledge, the knowledge created by only the best firm matters for growth. This would be true if knowledge had only one dimension. For example, the knowledge embodied in every blueprint could be used to lower the production cost of the final good. Then the knowledge that leads to the greatest reduction in cost is adopted by sector F; the *best practice* in period t becomes the standard in period $t+1$ and the knowledge created by the most able innovator becomes critical for growth.

This is similar to Murphy et al. (1991), where the growth rate is determined by the knowledge of the most able entrepreneur in the economy. In their model, if the parameters are such that the best agents become entrepreneurs, the growth rate would correspond to the knowledge of the most able agent in the economy. In this model, as in Murphy et al., an increase in inequality does not change the identity of the best innovator in the economy, because comparative advantage implies that the most able agents always select into being innovators. But, unlike Murphy et al., an increase in inequality has an additional effect. The most able innovator is matched with a better manager in the more unequal distribution ($\eta_B > \eta_A$). Hence, given our assumptions about k, the best firm creates weakly more knowledge under $\Phi_B(a)$.

The above examples suggest that the definition of λ is quite general and nests some of the more popular endogenous growth models in the literature. The examples also suggest that the potential overlap of ideas, captured by $\omega_i(a)$, is critical for growth. These weights could determine whether an increase in inequality causes the growth rate to fall (as in Example 1), or rise weakly (as in Example 2). The examples presented above, however, make some extreme assumptions. Let us now proceed with the general case. We know from Corollary (1) that $\eta_B > \eta_A$. This allows me to write,

$$\begin{aligned} \lambda_A &= \int_{\eta_A}^{\overline{a}} \omega_A(a) k(a, m_A^{-1}(a)) \phi_A(a) da \\ &= \underbrace{\int_{\eta_A}^{\eta_B} \omega_A(a) k(a, m_A^{-1}(a)) \phi_A(a) da}_{P} + \underbrace{\int_{\eta_B}^{\overline{a}} \omega_A(a) k(a, m_A^{-1}(a)) \phi_A(a) da}_{Q} \end{aligned} \tag{22}$$

and

$$\lambda_B = \underbrace{\int_{\eta_B}^{\overline{a}} \omega_B(a) k(a, m_B^{-1}(a)) \phi_B(a) da}_{R} \tag{23}$$

The terms Q and R defined above capture the contribution of innovators with ability in $[\eta_B, \overline{a}]$. These are the innovators who operate under both distributions. λ_A has an additional component, denoted by P in equation (22), which captures the contribution of innovators who operate under Φ_A but switch occupation under Φ_B. An increase in inequality affects growth through two effects - (i) a *distribution effect*, captured by $Q - R$ and (ii) an *allocation effect*, captured by P. A more unequal distribution has fewer innovators which has a dampening effect on the knowledge created. At the same time, the average ability of innovators under the more unequal distribution is higher, which has a positive effect on knowledge creation. Naturally, the final impact depends on which effect dominates. This would depend on the weights, the density and the function $k(a, m_A^{-1}(a))$.

In this paper, I consider two different scenarios for knowledge creation. In the first scenario, innovators protect their blueprints using patents. Since patents are effective for a single period, the blueprint is freely available in the following period. In this case, new knowledge is completely embodied in the blueprint created by each firm. Hence, $k(a, m_A^{-1}(a)) = a$. In the second scenario, innovators do not disclose their blueprints. In this case, new knowledge is embodied in the products created by each firm. Hence, $k(a, m_A^{-1}(a)) = x(a, m_A^{-1}(a)) = a[m_A^{-1}(a)]^\beta$. I consider both cases in turn. Henceforth, I also assume that $\omega_i(a) = \omega(a)$ with $\omega' \geq 0$.

5.1 Case I: Knowledge is created only by innovator

When $k(a, m_A^{-1}(a)) = a$, a change in the ability distribution has no impact on knowledge created by an innovator of ability a. The densities, on the other hand, change. In particular, conditional on being truncated at η_B, Φ_B first-order stochastic dominates Φ_A. Therefore, with the weights being the same, $Q - R < 0$; the distribution effect of an increase in inequality is positive.[13] At the same time, since a rise in inequality leads the least-able managers to switch occupations, some knowledge is lost. Therefore, the allocation effect is always negative, provided that the weights attached to the knowledge generated by the less able innovators are positive.

With $\omega(a)$ being non-decreasing in knowledge, the effect of $\omega(a)$ and a are not separately identifiable. Then we can simply define $\hat{k}(a) = \omega(a)a$. Observe that the marginal increase in $R - Q$ will be greater, the more convex $\hat{k}(a)$ is for $a \in [\eta_B, \overline{a}]$. At the same time, P will be small if $\hat{k}(a)$ is small over the relevant range. Combining the two observations, we can infer that if $\hat{a}$ is sufficiently convex, the distribution effect dominates and the growth rate rises following an increase in inequality. An extreme form of this convexity is illustrated in Example 2. The above discussion can be summarized in the following proposition :

Proposition 4. *When knowledge creation depends only on the ability of the innovator, an increase in inequality leads to higher growth if the weights are increasing and sufficiently convex in ability.*

A convex $\omega(a)$ function implies that knowledge created by more able innovators are more effective in the sense that they have less overlaps. Therefore, although an increase in inequality reduces the mass of

[13]This follows from the definition of first-order stochastic dominance

innovators, the corresponding loss of knowledge is small compared to the increase in knowledge coming from innovators who are, on average, more able.

5.2 Case II: Knowledge creation exhibits complementarity

In this case, a change in the distribution not only affects $\phi_i(a)$ but also $m_i^{-1}(a)$. In order to analyze the effect of a mean-preserving spread on matching, I need the following lemma.

Lemma 4. *With an increase in inequality, the match of the manager with knowledge η_B worsens, i.e., $k_B^* \leq m_A^{-1}(\eta_B)$.*

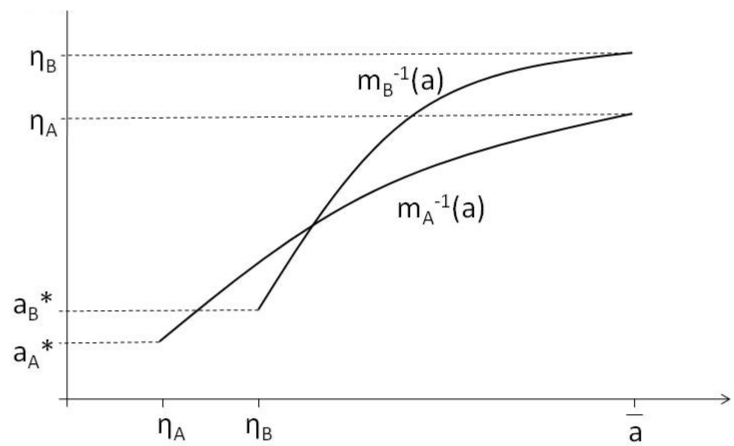

Figure 3: *The $m^{-1}(k)$ function under the two distributions*

We know that $m_A^{-1}(\overline{a}) < m_B^{-1}(\overline{a})$. With an increase in inequality, the inverse matching function becomes steeper for lower values of a but eventually flattens out. Combined with Lemma (4), this implies that there is a set $E \subset [\eta_B, \overline{k}]$, such that $m_A^{-1}(a) > m_B^{-1}(a)$ for $a \in E$ and $m_A^{-1}(a) < m_B^{-1}(a)$ for $a \in E^c$. This is illustrated in Figure 3. The figure suggests that the distribution effect defined earlier is no longer unambiguously positive. Following a rise in inequality, some of the incumbent innovators have better matches, but there are some innovators who are matched with less able managers than before. Of course, how this affects knowledge creation in each firm and consequently in the aggregate depends on β, the elasticity of knowledge (and also output) with respect to the manager's ability.

For a given increase in inequality, the sign of the distribution effect depends on how β compares with the curvature of $\omega(a)$. Holding everything constant, a fall in β reduces the distribution effect. A lower β, which makes $[m_A^{-1}(a)]^\beta$ more concave, makes knowledge creation less sensitive to the manager's ability for high values of $m_A^{-1}(a)$, but more sensitive to the manager's ability for low values of $m_A^{-1}(a)$. Since under Φ_B, the matches improve for those innovators who were matched with high ability managers but worsen for those matched with low ability managers, a low β reduces aggregate knowledge. As far as

$\omega(a)$ is concerned, the intuition is very similar to the one offered in the previous subsection. A sufficiently convex $\omega(a)$ re-enforces the effect of a shift in the mass towards more able innovators. The increase in knowledge coming from the improvement in matches of the most able innovators carries much more weight compared to the decrease in knowledge arising from the worsening of some matches. This allows us to state the following proposition.

Proposition 5. *When knowledge creation is complementary, an increase in inequality will raise the growth rate if, ceteris paribus, (i) β is sufficiently high or (ii) the weights are increasing and sufficiently convex in ability.*

5.3 Discussion

Propositions 4 and 5 underline the significance of the weights $\omega(a)$ in determining the relation between inequality and growth. We have already seen that different functional forms for the weights generate very different outcomes. But is it possible to discipline the weights in some ways so that we could make more systematic statements about the inequality-growth relation?

The interpretation of $\omega(a)$ as capturing the potential overlap of knowledge is useful in this respect. One could claim that the "uniqueness" of knowledge generated by each firm is a function of the level of development of an economy. When a less-developed economy starts to innovate, each innovator discovers something unique; there is a huge mass of knowledge to be discovered, and the probability of multiple innovators discovering the same knowledge is small. But as an economy accumulates more and more knowledge, i.e., as K_t increases, it becomes increasingly difficult to discover something unique. In such a situation, it is the more able innovators who are more likely to create unique knowledge. Accordingly, knowledge created by the more able innovators will carry more weight in aggregate knowledge. In terms of our analysis, as an economy develops, the $\omega(a)$ function becomes more convex. The implication is the following: for less-developed economies, the relation between inequality and growth could be negative, while for developed economies, this relation could be positive. This is consistent with the findings of Barro in his study of the relation between inequality and growth in a panel of countries. Barro finds that the effect of inequality on growth is negative for GDP per capita below $2070 and positive for higher values, where GDP is measured in 1985 U.S. dollars. One possible explanation for this finding, as pointed out by Barro, is the reduction in credit market imperfections as a country develops. When the credit constraints are severe, a rise in inequality reduces the growth rates. But with well-functioning credit markets, the growth-enhancing effects of inequality dominate. This paper provides an alternative explanation of this phenomenon.

It should be noted that the mechanism (through which inequality affects growth) that I have highlighted in this paper has a fundamental difference with what has been proposed in the existing literature on growth and inequality. In the existing literature, inequality affects growth primarily through its impact on the lower tail of the income/wealth distribution. In contrast, in this model, growth depends primarily on what is happening at the upper tail of the distribution. This is because growth is driven by knowledge which is created in the ability-intensive sector of the economy. The less able agents (usually the poor) in the economy sort into the final goods sector which benefits from the spillover but does not generate any. A change in the distribution of agents working in the sector F has no direct impact on growth, because it does not affect the rate of knowledge creation.[14]

6 Conclusion

In this paper, although an increase in inequality keeps the mean of the ability distribution unchanged, the mean ability of the innovators goes up. If the weights increase sufficiently quickly with ability, fatter tails would lead to more knowledge creation by the surviving innovators. This is not only because more innovators create more (expected) knowledge, but also because the better innovators are matched with better managers which has a positive effect on knowledge creation, if the latter exhibits complementarity. For an unchanged ability support for the innovators, this results in more knowledge on the aggregate. But a change in inequality affects the allocation of agents across occupations, which in turn means fewer innovators. The effect on the growth rate depends on whether having relatively more able innovators can compensate for having fewer innovators.

Neo-Classical growth theory has come a long way since the seminal work of Robert Solow (Solow, 1956). Economists have studied the relation between the rate of growth of a country and a myriad phenomena, including innovation, migration, corruption and pollution, to name just a few. Although it has received considerable attention, the interaction between inequality and growth is far from being well-understood (See Helpman, 2004, Chapter 6 for a summary of the relevant issues). For one, the direction of causality is un-clear. In this paper, I take a position on causality and proceed to analyze the nature of this causality, being fully aware of the fact that there could be reverse causality. I take the fundamental determinants of the distribution of ability as given and go on to examine the implications of a mean-preserving spread of this distribution on the the growth rate of the economy. Since the ability distribution of a country maps into a wealth distribution, this model generates an indirect link between wealth distribution and the growth rate of a country. Of course, the exogenous parameter that shapes the ability distribution, Γ, is itself determined by the wealth distribution of a country. My focus on the ability distribution, rather than the wealth distribution, stems from the belief that the pool of knowledge in an economy ultimately depends on the ability of the agents engaged in the creation and implementation of knowledge.

[14]This prediction is consistent with Barro (2000)'s finding that growth is positively related to the average years of attainment at the secondary and higher levels of education, but insignificantly related to primary attainment at the start of each period.

References

Acemoglu, D. (2008), *Introduction to Modern Economic Growth*, Princeton University Press.

Aghion, P. and Bolton, P. (1997), 'A theory of trickle-down growth and development', *The Review of Economic Studies* **64**(2), 151–172.

Aghion, P., Caroli, E. and García-Peñalosa, C. (1999), 'Inequality and economic growth: The perspective of the new growth theories', *Journal of Economic Literature* **37**(4), 1615–1660.

Aghion, P. and Howitt, P. (1992), 'A model of growth through creative destruction', *Econometrica* **60**(2), 323–51.

Alesina, A. and Rodrik, D. (1994), 'Distributive politics and economic growth', *The Quarterly Journal of Economics* **109**(2), 465–490.

Banerjee, A. V. and Duflo, E. (2003), 'Inequality and growth: What can the data say?', *Journal of Economic Growth* **8**(3), 267–99.

Banerjee, A. V. and Newman, A. F. (1991), 'Risk-bearing and the theory of income distribution', *The Review of Economic Studies* **58**(2), 211–235.

Barro, R. J. (2000), 'Inequality and growth in a panel of countries', *Journal of Economic Growth* **5**(1), 5–32.

Benhabib, J. and Spiegel, M. (1997), Cross-country growth regressions, Technical report.

Dasgupta, K. (2010), Learning and knowledge diffusion in a global economy, Working paper, University of Toronto.

Ehrlich, I. and Lui, F. T. (1999), 'Bureaucratic corruption and endogenous economic growth', *Journal of Political Economy* **107**(S6), S270–S293.

Ethier, W. J. (1982), 'National and international returns to scale in the modern theory of international trade', *The American Economic Review* **72**(3), 389–405.

Forbes, K. J. (2000), 'A reassessment of the relationship between inequality and growth', *The American Economic Review* **90**(4), pp. 869–887.

Galor, O. (2011), Inequality, human capital formation and the process of development, Working Paper 17058, NBER.

Galor, O. and Zeira, J. (1993), 'Income distribution and macroeconomics', *The Review of Economic Studies* **60**(1), 35–52.

Grossman, G. M. and Helpman, E. (1991), 'Quality ladders in the theory of growth', *Review of Economic Studies* **58**(1), 43–61.

Grossman, S. J. and Hart, O. D. (1986), 'The costs and benefits of ownership: A theory of vertical and lateral integration', *The Journal of Political Economy* **94**(4), pp. 691–719.

Hart, O. and Moore, J. (1990), 'Property rights and the nature of the firm', *The Journal of Political Economy* **98**(6), pp. 1119–1158.

Helpman, E. (2004), *The Mystery of Economic Growth*, Belknap Harvard.

Jones, C. I. (1999), 'Growth: With or without scale effects?', *The American Economic Review* **89**(2), 139–44.

Jones, C. I. (2005), Growth and ideas, *in* P. Aghion and S. Durlauf, eds, 'Handbook of Economic Growth, Volume 1B', Elsevier, pp. 1063–1111.

Kremer, M. and Maskin, E. (1996), Wage inequality and segregation by skill, Working paper, Massachusetts Institute of Technology (MIT), Department of Economics.

Lucas, R. E. (2009), 'Ideas and growth', *Economica* **76**(301), 1–19.

Meisenzahl, R. and Mokyr, J. (2011), The rate and direction of invention in the british industrial revolution: Incentives and institutions, Working Paper 16993, NBER.

Murphy, K. M., Shleifer, A. and Vishny, R. W. (1991), 'The allocation of talent: Implications for growth', *The Quarterly Journal of Economics* **106**(2), 503–30.

Perotti, R. (1996), 'Growth, income distribution, and democracy: What the data say', *Journal of Economic Growth* **1**(2), 149–87.

Persson, T. and Tabellini, G. (1994), 'Is inequality harmful for growth?', *The American Economic Review* **84**(3), 600–621.

Romer, P. M. (1990), 'Endogenous technological change', *Journal of Political Economy* **98**(5), S71–102.

Solow, R. M. (1956), 'A contribution to the theory of economic growth', *The Quarterly Journal of Economics* **70**(1), 65–94.

Appendix

Proof of Lemma 1 : Replacing the value of $x_t(a_I, a_M)$ in equation (9) and totally differentiating, we have

$$w_t''(a_M)da_M = \alpha^2 K_t L_{Y,t}^{1-\alpha}[(\alpha\beta - 1)a_I^\alpha a_M^{\alpha\beta-2}da_M + \alpha a_I^{\alpha-1}a_M^{\alpha\beta-1}da_I].$$

This can be re-written as,

$$[w_t''(a_M)da_M - \alpha^2 K_t L_{Y,t}^{1-\alpha}(\alpha\beta - 1)a_I^\alpha a_M^{\alpha\beta-2}] = \alpha^2 K_t L_{Y,t}^{1-\alpha}\alpha a_I^{\alpha-1}a_M^{\alpha\beta-1}\frac{da_I}{da_M}.$$

If there is positive assortative matching, $\frac{da_I}{da_M} > 0$, which implies that the right-hand side of the above equation is positive. Then the left-hand side must be positive too. It can be easily checked that the left-hand side being positive corresponds to the second-order condition for profit maximization being satisfied.

Proof of Lemma 2 : Suppose that, on the contrary, we have the following allocation: $[a_1, a_2] \in F$, $[a_2, a_3] \in M$, $[a_3, a_4] \in F$. Then both a_2 and a_3 must be earning s_t, since both of them are indifferent between being in sector F and M. If a_2 chooses to be a worker, then $w_t(a_2) = s_t$. Otherwise, $\pi_t(a_2) = s_t$. Either way, since $w_t'(a) > 0$ and $\pi_t'(a) > 0$, a_3 must be earning more than a_2 and consequently, more than s_t. A contradiction.

To show that the least able agents are in sector F, suppose that, on the contrary, $[\underline{a}, a_1] \in M$, $[a_1, a_2] \in F$. Without loss of generality, suppose that a_1 is a manager. Consider the agent with ability $a_1 + \epsilon$, where $\epsilon \to 0$. Since he is in sector F, he must be earning s_t. However, if he moves to sector M as a manager, he can earn $w_t(a_1 + \epsilon) > w_t(a_1) = s_t$, since $w_t'(a_1) > 0$. Therefore, this allocation cannot be an equilibrium.

Proof of Lemma 3 : See Dasgupta (2010).

Proof of Proposition 1 : Define $\zeta(a_t^*) = w_t(a_t^*) - s_t$, where $w_t(a_t^*)$ is given by equation (18) and s_t is given by equation (6). Differentiating $w_t(a_t^*)$ with respect to a_t^*,

$$\begin{aligned}\frac{\partial w_t(a_t^*)}{\partial a_t^*} &= \frac{\alpha^2 K_t}{2}L_{Y,t}^{1-\alpha}([x(\eta_t)]^{\alpha-1}\eta_t' + [x(\eta_t)]^{\alpha-1}\frac{\partial x(m_t(a))}{\partial a}|_{a=a_t^*} - [x(\overline{a})]^{\alpha-1}\eta_t'\frac{\partial x(m_t(a))}{\partial a}|_{a=\eta_t}) \\ &+ \frac{\alpha A_t}{2}\frac{\partial L_{Y,t}^{1-\alpha}}{\partial a_t^*}([x_t(\eta_t)]^\alpha - \alpha\int_{a_t^*}^{\eta_t}[x_t(m_t(a))]^{\alpha-1}\frac{\partial x_t}{\partial a}da)\end{aligned}$$

Now, $L_{Y,t} = \Phi_t(k_t^*)$. So, $\frac{\partial L_{Y,t}^{1-\alpha}}{\partial a_t^*} > 0$ and hence the second term on the right-hand side is positive. Therefore, the above expression will be positive if the first term is positive too. A sufficient condition for that is

$$[x(\eta_t)]^{\alpha-1} - [x(\overline{a})]^{\alpha-1}\frac{\partial x(m_t(a))}{\partial a}|_{a=\eta_t} > 0$$

Using the definition of x_t, $\frac{\partial x(m_t(a))}{\partial a}|_{a=\eta_t} = \beta\overline{a}\eta^{\beta-1}$. Replacing in the above expression,

$$\begin{aligned}[x(\eta_t)]^{\alpha-1} - [x(\overline{a})]^{\alpha-1}\frac{\partial x(m_t(a))}{\partial a}|_{a=\eta_t} &= (\eta a_t^{*\beta})^{\alpha-1} - (\overline{a}\eta^\beta)^{\alpha-1}(\beta\overline{a}\eta^{\beta-1}) \\ &= (\eta a_t^{*\beta})^{\alpha-1} - \beta\overline{a}^\alpha\eta^{\alpha\beta-1}\end{aligned}$$

Hence, for the left-hand side to be positive, we must have $\eta^{\alpha(1-\beta)}a_t^{*\beta(\alpha-1)} > \beta\overline{a}^\alpha$. Now, $\eta, a_t^* > \underline{a}$. Therefore, a sufficient condition is $\underline{a}^{\alpha(1-\beta)}\underline{a}^{*\beta(\alpha-1)} > \beta\overline{a}^\alpha$, i.e., $\underline{a}^{\alpha-\beta} > \beta\overline{a}^\alpha$. Since $\underline{a} < 1$, $\underline{a}^{\alpha-\beta} > \underline{a}^\alpha$. Therefore the required condition is $\underline{a}^\alpha > \beta\overline{a}^\alpha$ which is satisfied by Assumption1. It is also straightforward to show that

$\frac{\partial s_t}{\partial a_t^*} < 0$. Therefore, $\frac{\partial \zeta(a_t^*)}{\partial a_t^*} > 0$. Furthermore, when $a_t^* = 1$, $L_{Y,t} = 0$ and $\zeta(a_t^*)|_{a_t^*=1} = -\infty$. And when $a_t^* = \overline{a}$, $\zeta(a_t^*)|_{a_t^*=\overline{a}} = \frac{\alpha A_t}{2}[x(\overline{a})]^\alpha > 0$. Applying the intermediate value theorem, there exists a unique a_t^* such that $\zeta(a_t^*) = 0$.

Proof of Proposition 2 : Let $\mathbb{B}$ be any Borel set of $[\underline{a}, \overline{a}]$. Then the transition function for the ability distribution satisfies, for every $a \in [\underline{a}, \overline{a}]$,

$$P_t(a, \mathbb{B}) = \int\limits_{\mathbb{B}} d\tilde{G}_t(s|a, \Gamma)$$

Suppose P is monotone, has the Feller property and satisfies a mixing condition. Then P has a unique, invariant probability measure $\Phi(a)$ (Stokey, Lucas with Prescott, 1989, Theorem 12.12). Let us define the operator T as

$$(Tf)(a) = \int f(a')P(a, da'), \quad \text{all } a \in [\underline{a}, \overline{a}]$$

where $f : [\underline{a}, \overline{a}] \to \mathbb{R}$ is a bounded function. If f is non-decreasing, then the first-order stochastic dominance property of $G(a|b, \Gamma)$ implies that Tf is also non-decreasing (Monotone Property). It is straight-forward to verify that if f is bounded and continuous, then the same holds for Tf, i.e., $T : C(a) \to C(a)$ (Feller Property). The mixing condition requires that $\exists c \in [\underline{a}, \overline{a}]$, $\epsilon > 0$ and $N \geq 1$ such that $P^N(\underline{a}, [c, \overline{a}]) \geq \epsilon$ and $P^N([\underline{a}, c], \overline{a}]) \geq \epsilon$. The assumption that $G(a|b, \Gamma)$ has full support guarantees this. Therefore all the conditions for the existence and uniqueness of the invariant distribution are satisfied.

Proof of Proposition 3 : Suppose, on the contrary, the threshold between sectors F and M remain unchanged following an increase in inequality, i.e., $a_A^* = a_B^* = a^*$. Then, $\Phi_A(k) = \Phi_B(k)$. Since the density between a_A^* and η_A is the same for the two distributions, we must also have $\eta_A = \eta_B = \eta$. From equation (15), we have $m'_B(a^*) = \frac{\phi_B(a^*)}{\phi_B(\eta)} = \frac{\phi_A(a^*)}{\phi_A(\eta)} = m'_A(a^*)$. In fact, this relation must hold for all $k \in [k^*, m_A^{-1}(s_1)]$, i.e.

$$m'_B(a) = m'_A(a) \quad \forall k \in [a^*, m_A^{-1}(s_1)]$$

It follows that,

$$m_B(a) = m_A(a) \quad \forall a \in [a^*, m_A^{-1}(s_1)] \tag{i}$$

We also know that, $\phi_B(s_1) < \phi_A(s_1)$. Therefore, at $m_B^{-1}(s_1) = m_A^{-1}(s_1)$,

$$m'_B(m_B^{-1}(s_1)) = \frac{\phi_B(m_B^{-1}(s_1))}{\phi_B(s_1)} > \frac{\phi_A(m_A^{-1}(s_1))}{\phi_A(s_1)} = m'_A(m_A^{-1}(s_1))$$

In fact, $\forall s$ such that $m_B(k) \in [s_1, s_2]$

$$m'_B(a) = \frac{\phi_B(a)}{\phi_B(m_B(a))} > \frac{\phi_A(a)}{\phi_A(m_B(a))} > \frac{\phi_A(a)}{\phi_A(m_A(a))} = m'_A(a)$$

where the first inequality follows from definition and the second inequality follows from the assumption that density is decreasing in knowledge. It follows that

$$m_B(k) > m_A(k) \quad \forall a \in [m_B^{-1}(s_1), m_B^{-1}(s_2)] \tag{ii}$$

By assumption, $m_A(\eta) = m_B(\eta) = \overline{a}$. Furthermore, $m'_B(\eta) = \frac{\phi_B(\eta)}{\phi_B(\overline{a})} < \frac{\phi_A(\eta)}{\phi_A(\overline{a})} = m'_A(\eta)$. Therefore, in an $\epsilon - ball$ around η, $m_B(a) > m_A(a)$. In fact, using a similar logic as above, it can be shown that

$$m_B(a) > m_A(a) \quad \forall a \in [m_B^{-1}(s_2), \eta] \tag{iii}$$

Combining equations (i), (ii) and (iii), we see that $\exists S \in [a^*, \eta]$, such that

$$\begin{array}{ll} m_B(a) > m_A(a) & \forall a \in S \\ m_B(a) = m_A(a) & \text{otherwise} \end{array}$$

Using the result that matches weakly improve with inequality in equation (18), we see that $w(a^*)$ under Φ_B is less than that under Φ_A. We also know that $s = (1-\alpha)AL_Y^{-\alpha}(\int_\eta^{\overline{a}} x(a)^\alpha d\Phi(a))$. But,

$$\int_\eta^{\overline{a}} x(a)^\alpha d\Phi(a) = \int_{a^*}^{\eta} x(m(a))^\alpha d\Phi(a)$$

i.e., output produced can be attributed either to the worker or the manager. Since the difference between Φ_A and Φ_B occurs only in the tails, $\phi_A(a) = \phi_B(a)$. Therefore,

$$\begin{aligned} \int_{a^*}^{\eta} x(m_A(a))^\alpha d\Phi_A(a) &= \int_{a^*}^{\eta} m_A(a)^\alpha a^{\alpha\beta} d\Phi_A(a) \\ &= \int_{a^*}^{\eta} m_A(a)^\alpha a^{\alpha\beta} d\Phi_B(a) \\ &= \int_{a^*}^{m_A^{-1}(s_1)} m_A(a)^\alpha a^{\alpha\beta} d\Phi_B(a) + \int_{m_A^{-1}(s_1)}^{\eta} m_A(a)^\alpha a^{\alpha\beta} d\Phi_B(a) \end{aligned}$$

Now,

$$\int_{a^*}^{m_A^{-1}(s_1)} m_A(a)^\alpha a^{\alpha\beta} d\Phi_B(a) = \int_{a^*}^{m_A^{-1}(s_1)} m_B(a)^\alpha a^{\alpha\beta} d\Phi_B(a)$$

while

$$\int_{m_A^{-1}(s_1)}^{\eta} m_A(a)^\alpha a^{\alpha\beta} d\Phi_B(a) < \int_{m_A^{-1}(s_1)}^{\eta} m_B(a)^\alpha a^{\alpha\beta} d\Phi_B(a)$$

Hence,

$$\begin{aligned} \int_{a^*}^{\eta} x(m_A(a))^\alpha d\Phi_A(a) &< \int_{a^*}^{m_A^{-1}(s_1)} m_B(a)^\alpha a^{\alpha\beta} d\Phi_B(a) + \int_{m_A^{-1}(s_1)}^{\eta} m_B(a)^\alpha a^{\alpha\beta} d\Phi_B(a) \\ &= \int_{a^*}^{\eta} x(m_B(a))^\alpha d\Phi_B(a) \end{aligned}$$

It follows that s increases as inequality goes up. Thus, as we move from Φ_A to Φ_B, if we hold the threshold a^* constant, $w(a^*)$ goes down while s goes up. Therefore, at $a = a^*$, $w(a^*) < s$. This can not be an equilibrium. Hence, the thresholds must change. In particular, under Φ_B, there is an excess demand for workers in the retail sector if the threshold is at a_A^*. Hence, workers move from sector M to F, resulting in a^* moving to the right. The conditions on the parameters ensure that $s - w(a^*)$ falls as a^* rises. In equilibrium, $a_B^* > a_A^*$.

Proof of Lemma 4 : From Proposition (3), we know that $\eta_B > \eta_A$. By definition, $\Phi_A(\eta_B) - \Phi_A(\eta_A) = \Phi_A(m_A^{-1}(\eta_B)) - \Phi_A(a_A^*)$. Also, $\Phi_A(m_A^{-1}(\eta_B)) - \Phi_A(a_A^*) = \Phi_B(m_A^{-1}(\eta_B)) - \Phi_B(a_A^*)$. Let us consider the following two scenarios. In scenario one, s_1 in Figure (2) is greater than η_B. Then, we must have $\Phi_B(\eta_B) - \Phi_B(\eta_A) = \Phi_A(\eta_B) - \Phi_A(\eta_A)$. In scenario two, s_1 is less than η_B. In this case, $\Phi_B(\eta_B) - \Phi_B(\eta_A) < \Phi_A(\eta_B) - \Phi_A(\eta_A)$. Combining the two, we have $\Phi_B(\eta_B) - \Phi_B(\eta_A) \leq \Phi_A(\eta_B) - \Phi_A(\eta_A)$. Therefore, $\Phi_B(\eta_B) - \Phi_B(\eta_A) \leq \Phi_B(m_A^{-1}(\eta_B)) - \Phi_B(a_A^*)$. It follows that $m_B^{-1}(\eta_B) = a_B^* \leq m_A^{-1}(\eta_B)$.