

On the distribution of income and worker assignment under intra-firm spillovers, with an application to ideas and networks.

Gilles Saint-Paul
Universitat Pompeu Fabra and CEPR*

October 5, 1999

Abstract

We study the earnings structure and the equilibrium assignment of workers when workers exert intra-firm spillovers on each other. We allow for arbitrary spillovers provided output depends on some aggregate index of workers' skill. Despite the possibility of increasing returns to skills, equilibrium typically exists. We show that equilibrium will typically be segregated; that the skill space can be partitioned into a set of segments and any firm hires from only one segment.

Next, we apply the model to analyze the effect of information technology on segmentation and the distribution of income. There are two types of human capital, productivity and creativity, i.e. the ability to produce ideas that may be duplicated over a network. Under plausible assumptions, inequality rises and then falls when network size increases, and the poorest workers cannot lose. We also analyze the impact of an improvement in worker quality and of an increased international mobility of ideas.

*I thank Roland Bénabou, Fernando Alvarez, Adriana Kugler, Rafael Repullo, Sherwin Rosen, Masako Ueda, participants at the NBER Summer Institute workshop on inequality, 1998, the Conference on Technology, Regulation, and Employment, Madrid, 1999, as well as seminar participants at Universitat Pompeu Fabra and Université des Sciences Sociales de Toulouse, for helpful comments and suggestions.

□

1 Introduction

Over the last ten years there has been a renewed interest in problems associated with the distribution of income. This interest has been motivated by the upward trend in inequality that has taken place in the United States and other countries since the early seventies, and has led to a spurt of empirical and theoretical research.¹ This research has studied the role of such diverse factors as the role of international trade, computers, educational achievements, and stratification in firms and neighborhoods.

The present paper is a theoretical contribution to that literature, which was originally motivated by an attempt to understand the effect of information technology on the distribution of income. It is mainly concerned with how technology and stratification interact with each other in determining the distribution of income. We study the earnings structure and the equilibrium assignment of workers to firms in a model where workers exert intra-firm spillovers on each other. We allow for such spillovers to be somewhat arbitrary, provided the number of workers within firms is fixed—an assumption that can be relaxed to some extent— and output is a function of some aggregate index of all the workers' skills in the firm. Because of the spillover there are non constant marginal returns to skills, and we allow returns to be either increasing or decreasing. Despite the possibility of increasing returns to skills, equilibrium typically exists because the fixed number of workers within each firm prevents them from increasing profits by simply replicating themselves.

We show that equilibrium will typically be segregated, in the sense that the skill space can be partitioned into a set of segments and that any firm

¹See Levy and Murnane (1992), Bound and Johnson (1992), Katz and Murphy (1992), Juhn, Murphy and Pierce (1993). Some of the theoretical literature on segmentation and inequality that this burst of empirical research has motivated is referred to below in the text.

hires from only one segment. The number of segments depends on the shape of the curve relating total output to aggregate skills, with convex portions acting as a force toward segregation and concave ones as a force toward unification. The wage schedule is always convex in skills, regardless of the shape of spillovers, implying that market forces, by allocating people of similar skills into the same firms, magnifies differences in skills and tends to skew the distribution of income to the right. Wages can be interpreted as the difference between the worker's marginal product and a shadow cost of bodies that captures the spillover exerted by the worker on other workers within the firm. This cost is positive if there are increasing returns so that firms are willing to pay a negative wage for some categories of workers.

Next, we develop an application of the model which yields some insights about how information technology affects the distribution of income. There are two extreme ways in which one can think about that problem. At one extreme is the standard view, which ignores job characteristics and treats each labor type as homogeneous and divisible. According to that view, computers are substitute for unskilled workers, so that their introduction is quite similar to a rise in the supply of unskilled workers or trade with a country abundantly endowed in that factor.² At the other extreme there is the "Superstar", or "Winner-takes-all", view, according to which information technology allows the best people to spread their talent over an ever larger share of the market, at the expense of less able workers who may find themselves displaced into low-paying jobs, despite being only marginally less talented. This trend toward a society where all the rewards go to the best has been analyzed in Frank and Cook (1995), who provide numerous examples from sports, performing arts, and top executives. The question, however, is whether such effects are pervasive throughout the labor market or limited to football players and divas. Rosen, (1996), in his review of the book, argues that these phenomena are confined to those segments of the market where a product

²See Krueger (1993) and DiNardo and Pischke (1996) for the empirical aspect of that debate.

can be duplicated at no cost (the media) and/or where being the best is the essence of the product (sport competition).

The model we develop takes an intermediate view. We think of wages as the return to two types of human capital. One is productivity, that is how many units of output one can produce; the other is "creativity", i.e. the ability to produce ideas that may be duplicated over a segment of the economy, that we call a "network". In our model such ideas increase the output of all other workers in the same network—this is the nature of the spillover. We assume that firms and networks coincide, so that the beneficial effect of ideas will increase profitability and, through competition, eventually increase wages for the most creative workers.³ There is a "winner-takes-all" dimension to creativity because only the best idea is applied. We model progress in information technology as an increase in the size of the network over which ideas can be spread. It is therefore a force that increases the return to the most creative workers. But, contrary to what happens in e.g. sports, best performers are complementary to other workers since a good idea increases the productivity of all participants in the network. This is a force that mitigates the inegalitarian effects of information technology. Furthermore, as networks get larger, the most creative workers end up competing with each other in the same networks, which reduces the return to creativity.

Under plausible assumptions, we show that inequality rises and then falls when network size increases. Our conclusions are therefore far less pessimistic regarding the impact of information technology on inequality than the extreme "superstar" approach. We also show that the poorest workers cannot lose from an improvement in information technology and analyze the impact of an improvement in worker quality and of an increased international mobility of ideas on income distribution and worker segmentation.

³Thus we abstract from the problems of imperfect appropriability of intellectual contributions, as emphasized by the literature on endogenous growth (See Aghion and Howitt (1998), Grossman and Helpman (1992), Romer (1990), Lucas (1988)).

A Coasian argument would lead to the conclusion that the boundaries of the firms will endogenously adjust so as to internalize the externality; therefore in equilibrium firms will coincide with networks.

The present paper is related to several strands of literature. One is the literature on job assignment, whose central claim is that an earnings function that relates wages to worker characteristics is insufficient for predictive purposes, because it ignores the assignment of workers to jobs (See Sattinger (1993) for a survey). In our model this critique holds too, since the distribution of income depends not only on the distribution of skills but also on how workers are endogenously grouped together in the labor market, and a change in the underlying distribution of skills will also affect the way workers are grouped. However, this literature generally assumes that workers are assigned to occupations of jobs whose differences are specified exogenously, while in our case jobs are symmetrical ex-ante, but end up being different because one's productivity depends on one's colleagues' skills, which differ across firms in equilibrium.

Another literature is the one about "clubs", ghettos, and segregation, whose archetypal example is Becker's (1973) theory of marriage. Segregation arises when complementarities are prominent, so that people want to end up in the same "club" or firm, as the "best" ones. This is often the case in problems of local public finance or peer group effects (as in Epple and Romano (1998), Bénabou (1993,1996), Fernandez and Rogerson (1996), or Rioux (1999)). As in that literature, strong spillovers (that in our case show up in the form of convexity rather than complementarity) lead to segregation. The present paper, however, is more concerned with the determination of earnings and establishes general results regardless of the shape of the spillover.

Finally, the present paper is also related to the literature on "superstars", following Rosen (1981), who, in a seminal paper, established that the large income for superstars results from their ability to cover a larger market. The present paper differs from Rosen's superstars model in that spillovers play a key role in determining the distribution of income, while in Rosen's paper the key mechanism rather came from the ability of talented individual to cover a larger market, much in the fashion that, under imperfect competition, firms

with lower costs have more customers.⁴ Kremer's (1993) o-ring production function has more in common with the present paper. In both papers firms consist of a fixed number of workers⁵ and workers exert spillovers over each other. Here, however, we allow for a much more general pattern of spillovers. Both the o-ring production function and the ideas/networks application that we develop are special cases of that general analysis. In fact, in our case it is the best workers who exert positive spillovers, while in Kremer's model it is the worst performer who exerts negative spillovers.⁶

Despite the generality of our results, there is one aspect of that literature that is not captured by our model, namely the hierarchical one. Many authors, such as Calvo and Wellisz (1979), Rosen (1982), or, more recently, Garicano (1998), have developed models which predict that more talented people are more likely to be "above" in hierarchies and that span of control is increasing with talent. This is clearly not captured in our model where tasks are modelled as symmetrical; as we briefly argue, to capture these aspects would require a production function where in addition to the mean, higher moments of the intra-firm distribution of skills would enter.

The paper is organized as follows: in section 2 we establish our basic general results regarding existence and efficiency of equilibrium, as well as the pattern of segregation and the shape of the wage distribution. In section 3, we show how this framework can be applied to ideas and information technology. In section 4, we perform comparative statics exercises, looking at the impact on the distribution of income and the pattern of segregation of an increase in average worker quality, and improvement in information technology, and increased mobility in goods and ideas. These exercises are

⁴Murphy, Shleifer, and Vishny (1991) have applied Rosen's ideas to understand the interaction between economic growth and the allocation of talent across different activities. This is a very different route from the one pursued here, although the two kinds of considerations could be merged.

⁵Or, more fundamentally, the number of workers enters the production function separately from the aggregate labor input.

⁶Furthermore, the correlation between productivity and creativity plays a key role, in our model, in shaping the structure of returns to scale, while skill has only one dimension in the O-ring case.

motivated by the above mentioned debate on the rise of inequality. Section 5 briefly discusses potential extensions of the model, while section 6 contains concluding comments.

2 A model of the labor market under intra-firm spillovers

2.1 Basic setup and equilibrium definition

We consider an economy populated by a continuum of workers who differ by their skill level y , which is distributed over the interval $I = [y_0, y_1]$ with a measure μ , which has full support. The total mass of workers is normalized to 1. Firms freely enter the market and consist of a mass s of workers. Therefore there will be a mass $1/s$ of firms in equilibrium.⁷

A firm's total output is given by $a(\bar{y})$, where \bar{y} is the average skill level of its workers. Therefore, we assume that skills can be defined in such a way that the firm's output only depends on the aggregate skill level within the firm. In the absence of spillovers, the firm's output is simply the sum of the contributions of each individual worker, which does not depend on his colleagues' characteristics. Consequently, by redefining the skill level as being productivity, we get that the $a(\cdot)$ function is then linear. If, on the other hand, there are intra-firm spillovers, then the $a(\cdot)$ function will be nonlinear and this is the case of interest.

The production function exhibits increasing returns to skills if $a'(y)y > a(y)$. This condition is different from the condition for increasing *marginal* returns to skills, which is given by $a''(\cdot) > 0$. In our application below the first condition is always satisfied while the second may not hold. For the time being we do not make any assumption about the returns to skills, just assuming that $a(\cdot)$ is continuous and twice differentiable.

⁷More generally, the assumption of a fixed number of employees within each firm could be relaxed by specifying a production function which depends on both the number of workers and their aggregate skill level. This is discussed below in section 5.

An equilibrium is characterized by a wage schedule $w(y)$ which tells us how much a worker of skill y will be paid, and by an assignment of workers to firms such that all workers are assigned, firms maximize profits and no potential entrant could make strictly positive profits. The assignment can be represented by a mapping η which tells us the skill level of worker i in firm k . The definition below uses such a representation to precisely state the conditions for an equilibrium.

DEFINITION 1: Let M be the set of mappings from $[0, s]$ to I . Let m be the Lebesgue measure on $[0, s] \times [0, 1/s]$. An equilibrium consists of (i) a mapping η from $[0, s] \times [0, 1/s]$ to I , $(i, k) \rightarrow y_{ik}$ (worker assignment) and (ii) a mapping w from I to R , $y \rightarrow w(y)$ (the wage schedule) such that:

a. $\forall k \in [0, 1/s]$, $a(\bar{y}_k) = \int_0^s w(y_{ik})di$, where $\bar{y}_k = (\int_0^s y_{ik}di) / s$ is firm k 's average skill level.

b. $\forall \nu \in M : i \rightarrow \hat{y}_i$, $a(\int_0^s \hat{y}_i di / s) \leq \int_0^s w(\hat{y}_i) di$.

c. $\forall S \subset I$, $\mu(S) = m(\eta^{-1}(S))$.

The first property tells us that because of the free entry condition, existing firms make zero profits. The second condition says that potential entrants cannot make positive profits, otherwise they would indeed enter. The third condition is the full employment condition for any type of worker, which says that the distribution of y implied by the assignment of workers matches the actual one.

2.2 The structure of earnings

The next task is to characterize equilibrium and prove its existence. To do so we proceed in several steps. The next proposition tells us how to recover the wage schedule given the initial configuration of firms.

PROPOSITION 1 - The equilibrium wage schedule is such that:

(i) $w(y) \geq a(\bar{y}_k)/s + a'(\bar{y}_k)(y - \bar{y}_k)/s$, $\forall y \in I$, $\forall k \in [0, 1/s]$

(ii)

$$w(y) = a(\bar{y}_k)/s + a'(\bar{y}_k)(y - \bar{y}_k)/s, \quad (1)$$

if firm k employs some workers of type y .⁸

(iii)

$$w(\bar{y}_k) = a(\bar{y}_k)/s$$

PROOF: To compute their desired optimal employment structure firms maximize

$$\max_{\omega(\cdot, \cdot)} a \left(\int_I \omega(y) y d\mu / s \right) - \int_I \omega(y) w(y) d\mu$$

subject to

$$\omega(y) \geq 0, \forall y \in I$$

and

$$\int_I \omega(y) d\mu = s \tag{2}$$

That is, they elect how much of each type they want to hire, subject to the constraint that their total size is s . A type y is employed in that firm iff $\omega(y) > 0$.

The first order condition with respect to any $\omega(y)$ is therefore

$$a'(\bar{y}_k) y / s - w(y) \leq \lambda_k \tag{3}$$

where λ_k is the Lagrange multiplier⁹ associated with (??), and where equality holds if $\omega(y) > 0$. Integrating both sides over all employees of firm k and making use of the zero profit condition, we get that

$$\lambda_k = \frac{a'(\bar{y}_k) \bar{y}_k - a(\bar{y}_k)}{s}.$$

⁸Formally, this means that $k \in P_2(\eta^{-1}(y))$, where P_2 is the projection operator from $[0, s] \times [0, 1/s]$ to $[0, 1/s]$.

⁹One cannot readily apply the Kuhn-Tucker theorem because $a(\cdot)$ may not be concave. However, consider a type y employed by firm k and a replacement of a small mass of that type by another type y' . For the firm to be at its optimum it must be that such a change does not increase profits, which is equivalent to

$$a'(\bar{y}_k) \frac{y'}{s} - w(y') \leq a'(\bar{y}_k) \frac{y}{s} - w(y)$$

If y' is also employed in positive quantity by the firm then the converse must also hold, implying that the quantity $a'(\bar{y}_k) \frac{y}{s} - w(y)$ is the same for all workers employed by the firm. Denoting that quantity by λ_k and calling it the Lagrange multiplier we get back to (??).

Substituting into (??) completes the proof of (i) and (ii). To prove (iii), note that (i) applied to $y = \bar{y}_k$ implies $w(\bar{y}_k) \geq a(\bar{y}_k)/s$. Assume (iii) does not hold; then it implies $w(\bar{y}_k) > a(\bar{y}_k)/s$ and that firm k does not employ any worker of type $y = \bar{y}_k$. But then, any firm that employs workers of type \bar{y}_k could strictly increase its profits by replacing them by the same mix of workers as those employed in firm k . This would not affect its average productivity and yield an average labor cost equal to $a(\bar{y}_k)/s + a'(\bar{y}_k)(\bar{y}_k - \bar{y}_k)/s = a(\bar{y}_k)/s < w(\bar{y}_k)$. Consequently, no firm would employ type \bar{y}_k workers, which contradicts the equilibrium assumption. Thus, (iii) must hold. \square

COROLLARY - The wage schedule $w(y)$ is convex.

PROOF - Proposition 1 implies that $w(\cdot)$ is the maximum of a set of linear functions. \square

The RHS of (??) is firm's i willingness to pay for any type (y). It must be equal to its wage if the firm employs that type, and lower than its wage if it does not. The interpretation of (??) is as follows. Wages paid by firm k can be decomposed as $w(y) = a'(\bar{y}_k) y/s - \lambda_k$. The first term $a'(\bar{y}_k) y/s$ is the marginal product of an individual of type y if he works in firm k . It is the only term that would remain under constant returns to scale. The second term $\lambda_k = \frac{a'(\bar{y}_k)\bar{y}_k - a(\bar{y}_k)}{s}$ is the shadow price of bodies in firm k , or equivalently the price that anyone has to pay to be a member of that firm (or network). This price is positive if there are increasing returns and negative if there are decreasing returns. Under constant returns it is exactly equal to zero and people are just paid their marginal product. If individuals could be hired in addition to those already in the firm, they could simply be paid their marginal product, without affecting others. But by joining a firm with a fixed number of employees, the individual is filling a slot, reducing the number of slots available for his other colleagues. In equilibrium he will compensate them by transferring to others the value of a slot λ_k . That value is equal to the reduction in output per capita associated with the reduction in the firm's work force by one unit. This is what the other colleagues lose, on average, if the worker gets his full marginal product; to them, this is equivalent to

having one less person in the firm. An important aspect of (??) is that under increasing returns, there exists worker types such that the willingness to pay for those workers by firm k is *negative*. This arises for y low enough relative to \bar{y}_k .

2.3 Segregation and workers' assignment

The next issue is: how does the equilibrium look like in terms of the equilibrium distribution of firm's average skill levels and in terms of the assignment of workers to these firms? The next proposition characterizes the equilibrium.

PROPOSITION 2 — Equilibrium can be characterized as follows:

There exists a partition of I into adjacent intervals $I_\alpha \subset I$, indexed by $\alpha \in A$, such that

(i) For any firm k there exists $\alpha(k) \in A$ such that all workers of firm k come from $I_\alpha : \forall k \in [0, 1/s], \exists \alpha, \eta([0, s], k) \subset I_\alpha$

(ii) The wage schedule $w(y)$ is linear over $I_\alpha : \forall \alpha \in A, \forall y \in I_\alpha, w(y) = \omega_\alpha + \delta_\alpha y$

(iii) For any firm k recruiting in I_α , the wage schedule is tangent to the average output schedule at $y = \bar{y}_k : \delta_{\alpha(k)} = a'(\bar{y}_k)/s; \omega_{\alpha(k)} = [a(\bar{y}_k) - a'(\bar{y}_k)\bar{y}_k]/s$

(iv) The wage schedule is convex

(v) $w(y) \geq a(y)/s, \forall y \in I$.

(vi) The mean of y over I_α is a convex combination of the mean skill level of the firms that recruit in I_α .

(vii) supply equals demand, i.e. condition (c) in definition 1 holds.

PROOF. We must prove that these conditions are necessary and sufficient for equilibrium. Let us prove that they are sufficient first.

Over the interval in which a firm hires people, wages are linear. Aggregating them using (ii) and (iii) implies that a firm's average labor cost is $a(\bar{y}_k)/s$. Thus existing firms make zero profit.

As the wage schedule is convex ((iv)), an entrant with average skill \hat{y} cannot do better than hiring workers who all have the same skill \hat{y} . But then his profits are given by $a(\hat{y}) - s w(\hat{y})$, which by (v) cannot be positive. Thus

properties (a) and (b) in definition 1 hold. Property (c) holds because of (vii).

This proves sufficiency.

Let us now prove that these conditions are necessary. (iv) must hold because of the corollary to proposition 1. (vii) must hold because of definition 1. If (v) were violated for some \hat{y} , an entrant hiring s workers of type \hat{y} would make strictly positive profits, violating (b) in definition 1.

Next, consider two existing firms k and l , with different means $\bar{y}_k < \bar{y}_l$. We prove that the most skilled employee of firm k , of skill $y_k^+ \geq \bar{y}_k$ must be less skilled than the least skilled employee of firm l , of skill $y_l^- \leq \bar{y}_l$. To see this, note that proposition 1 implies

$$sw(y_l^-) = a(\bar{y}_l) + a'(\bar{y}_l)(y_l^- - \bar{y}_l) \geq a(\bar{y}_k) + a'(\bar{y}_k)(y_l^- - \bar{y}_k) \quad (4)$$

$$sw(y_k^+) = a(\bar{y}_k) + a'(\bar{y}_k)(y_k^+ - \bar{y}_k) \geq a(\bar{y}_l) + a'(\bar{y}_l)(y_k^+ - \bar{y}_l) \quad (5)$$

Taking the difference between the left hand side of (4) and the right hand side of (5), and making use of these inequalities, we get

$$(a'(\bar{y}_l) - a'(\bar{y}_k)) \cdot (y_l^- - y_k^+) \geq 0.$$

Thus, either $y_l^- \geq y_k^+$ or $a'(\bar{y}_l) \leq a'(\bar{y}_k)$. But we have, again from proposition 1:

$$sw(\bar{y}_l) = a(\bar{y}_l) \geq a(\bar{y}_k) + a'(\bar{y}_k)(\bar{y}_l - \bar{y}_k)$$

$$sw(\bar{y}_k) = a(\bar{y}_k) \geq a(\bar{y}_l) + a'(\bar{y}_l)(\bar{y}_k - \bar{y}_l)$$

these two inequalities imply that

$$a'(\bar{y}_k) \leq \frac{a(\bar{y}_l) - a(\bar{y}_k)}{\bar{y}_l - \bar{y}_k} \leq a'(\bar{y}_l)$$

Thus the only way we can have $y_l^- < y_k^+$ is if $a'(\bar{y}_k) = \frac{a(\bar{y}_l) - a(\bar{y}_k)}{\bar{y}_l - \bar{y}_k} = a'(\bar{y}_l) = \delta$ exactly. In that case the two firms offer the same wage schedule and can recruit within the same cluster.¹⁰ Furthermore, all firms with mean

¹⁰This is what we call a *dual cluster*.

skill level between \bar{y}_k and \bar{y}_l will also offer the same wage schedule. Let p index such a firm. Assume first that $a(\bar{y}_p) > a(\bar{y}_k) + \delta(\bar{y}_p - \bar{y}_k)$. Then by hiring workers of types \bar{y}_k and \bar{y}_l in proportions $(\bar{y}_l - \bar{y}_p)/(\bar{y}_l - \bar{y}_k)$ and $(\bar{y}_p - \bar{y}_k)/(\bar{y}_l - \bar{y}_k)$ respectively, the firm gets a total labor cost equal to $a(\bar{y}_k) + \delta(\bar{y}_p - \bar{y}_k)$, thus obtaining strictly positive profits, which contradicts the equilibrium assumption. Next, note that $a(\bar{y}_p) < a(\bar{y}_k) + a'(\bar{y}_k)(\bar{y}_p - \bar{y}_k)$ cannot hold, as it contradicts (i) in proposition 1, since by (iii) we have that $a(\bar{y}_p)/s = w(\bar{y}_p)$. Thus we necessarily have

$$a(\bar{y}_p) = a(\bar{y}_k) + \delta(\bar{y}_p - \bar{y}_k)$$

Finally, note that if $\delta \neq a'(\bar{y}_p)$, there exists some y in the neighborhood of \bar{y}_p such that $a(y) > a(\bar{y}_k) + \delta(y - \bar{y}_k)$. The above argument can then be used to show that an entrant with an average skill level equal to y would make strictly positive profits. Consequently one must also have $a'(\bar{y}_p) = \delta$.

To summarize: Either two firms offer the same wage schedule, in which case so do all firms in between, or they must obey the "sorting property": the most skilled person employed by firm k cannot be more skilled than the least skilled person employed by firm l .

The preceding results imply that if we order firms by increasing values of \bar{y}_k , they can be grouped in clusters that hire within the same adjacent intervals. To construct these intervals, group all firms into subsets of "consecutive" firms offering the same wage schedule. For any such subset α , define $I_\alpha = \bigcup_{k \in \alpha} \eta([0, s], k)$. As the distribution of types has full support, these sets cover the whole interval I . Furthermore, to match the above sorting property they must be intervals.

Within each I_α , the wage schedule is linear since all firms offer the same one, and tangent to each firm's mean value of y by construction. This proves (i), (ii), and (iii). To prove (vi), note that firms recruiting in I_α must hire the whole supply of workers in I_α . Thus the employment weighted average of the mean skill in each firm must be equal to the mean of y over I_α .

Q.E.D.

Proposition 2 is very useful to understand the structure of equilibria. The typical structure is illustrated on Figure 1, which shows us a map of how the economy organizes itself into various zones with very different properties. The average output schedule, which gives us $a(y)/s$ as a function of the firm's average skill level y , and is represented by the dotted curve, will in general have convex parts and concave parts. In the concave zones, or "humps" there are local decreasing returns to average human capital, which create a force for agglomeration of different people within the same firm. In the convex zones there are local increasing returns, which create a force for segregation, i.e. matching of people with similar levels of human capital within the same network.

As can be seen on figure 1, equilibrium is characterized by a map of segregated clusters. The relevant intervals are determined by the points where the wage schedule $w(y)$, represented by the plain curve, has a kink. Because of convexity, wages are always more sensitive to skill when one moves up the distribution of income. Both that property and the segregation property are quite general and do not depend on $a(.)$ being convex throughout. If $a(.)$ were concave there would be a single cluster and a linear wage schedule, but typically it is enough for it to be convex over some interval to generate segregation and increased steepness of the wage schedule.

The following taxonomy emerges:

1. In many cases there will arise "unitary zones", i.e. intervals of values of y in which all people are hired in a single type of heterogeneous firm. By single *type* we mean that they all have the same mean skill level, although they clearly can achieve that mean using different workers, provided they only hire from the relevant cluster. These firms are segregated in the sense that two firms with different mean skill levels hire from different intervals; but within each interval there is complete agglomeration in the same firm types. Each interval roughly corresponds to a hump or a concave zone. Indeed, (iii) and (v) in proposition 2 imply that for any firm that employs several types of workers the average output schedule is locally concave at

that firm's mean skill level.¹¹ The wage schedule paid by the firm is tangent to that hump at precisely the mean of the distribution of human capital over the corresponding interval. Since in equilibrium all workers in that interval must be hired by firms of the same type, their common mean skill level must be equal to the population mean of that cluster. The mass of firms of a given type is then determined so as to ensure that all workers in the corresponding interval are employed. On figure 1A we have represented a unitary zone covering two humps, and on figure 1B we have two consecutive unitary zones corresponding to two consecutive humps. Therefore, not all humps generate a separate cluster, as workers in a hump may be absorbed in a firm corresponding to a bigger hump.¹² The number of different unitary zones is not necessarily equal to the number of humps; it all depends on the underlying distribution of workers. However, the number of unitary zones cannot exceed the number of humps.

2. There may be a "hypersegregated", or "assortative" zone where there is a continuum of firm types each hiring a single type of workers. In this zone, where the output per capita schedule is necessarily convex, the wage schedule exactly matches the output per capita schedule. As illustrated on figures 1C and 1D, this typically happens when the wage schedules of two consecutive humps cut the average output schedule before cutting each other. In terms of proposition 2, in a hypersegregated zone, each skill level defines a cluster, which is reduced to one point.

3. Finally, there may exist "dual clusters", where two different types of firm hire from the same pool of workers, offering the same wages. This is illustrated on figure 1E. The interval of workers hired in a dual cluster roughly covers the two humps corresponding to each firm type. This corresponds to the case where the interval I_α contains more than one type of firm. In that

¹¹This property can also be directly proved using the second-order condition of the firm's optimization problem discussed in proposition 1.

¹²This will occur if these workers are not too far away not too numerous, so their participation in the cluster is compatible with that interval's mean skill level being located in the zone of the other hump.

case, the two firms may not be segregated: some people employed in the high productivity firm may be less skilled than some people employed in the low productivity firm. There are many possible equilibrium distributions of workers between the two firms; the only constraints are that everybody is employed and that the mean skill of each network corresponds to the tangency point. While in principle there may be more than two different types of firm in I_α , generically this will never arise.

Now that we understand the structure of equilibria, we proceed and analyze its efficiency and existence.

2.4 Efficiency and existence

One interesting question is: is the equilibrium efficient? Intuitively, the answer should be yes. There is no market failure and the spillovers exerted by people on each other through ideas are entirely internalized by firms, and *in fine* reflected in the wage structure. The next proposition proves that this is indeed the case.

PROPOSITION 3 - An allocation is an equilibrium if and only if it maximizes total output.

PROOF - First, let us take an allocation that maximizes total output and prove that it may be supported by an equilibrium. Let us normalize s to 1 to save on notation.

Let k, l , be two types of firms in that allocation, with corresponding mean skill levels \bar{y}_k, \bar{y}_l . Let μ_k (resp. μ_l) be the measure representing the distribution of people working in firms of type k (resp. l).

Consider the following reallocation of output: take a distribution of people of total infinitesimal mass ε and mean y out of firms of type k and allocate it to firms of type l in such a way that all firms of type k (resp. l) keep having the same common mean.

Then the new mass of type k firms is $\mu_k(I) - \varepsilon$ while their new mean skill level is $\bar{y}_k + \frac{\varepsilon}{\mu_k(I)}(\bar{y}_k - y)$. Similarly, the new mass of type l - firms is $\mu_l(I) + \varepsilon$ and their new mean skill level is $\bar{y}_l + \frac{\varepsilon}{\mu_l(I)}(y - \bar{y}_l)$.

The contribution of these two types of firms to total output is therefore

$$[\mu_k(I) - \varepsilon\mu_k(I)] a(\bar{y}_k + \frac{\varepsilon}{\mu_k(I)}(\bar{y}_k - y)) + [\mu_l(I) + \varepsilon] a(\bar{y}_l + \frac{\varepsilon}{\mu_l(I)}(y - \bar{y}_l))$$

This cannot exceed the original contribution, equal to $\mu_k(I)a(\bar{y}_k) + \mu_l(I)a(\bar{y}_l)$. Using a first order Taylor expansion, we see that for ε small enough this is equivalent to

$$a(\bar{y}_l) + (y - \bar{y}_l)a'(\bar{y}_l) \leq a(\bar{y}_k) + (y - \bar{y}_k)a'(\bar{y}_k) \quad (6)$$

This must hold for any pair (k, l) and any y employed by some type k -firm.

Another option is to create a mass ε of firms with average skill level exactly equal to y . The contribution of total output to these firms and type k firms must then be equal to

$$[\mu_k(I) - \varepsilon\mu_k(I)] a(\bar{y}_k + \frac{\varepsilon}{\mu_k(I)}(\bar{y}_k - y)) + \varepsilon a(y)$$

This cannot exceed the original contribution of type k firms, i.e. $\mu_k(I)a(\bar{y}_k)$. Taking again a first-order Taylor expansion we see that this is equivalent to

$$a(y) \leq a(\bar{y}_k) + (y - \bar{y}_k)a'(\bar{y}_k) \quad (7)$$

It is then not difficult to see that by taking $w(y) = \max_k a(\bar{y}_k) + (y - \bar{y}_k)a'(\bar{y}_k)$ this allocation supports an equilibrium. Equation (??) and our definition of $w(y)$ guarantee that (i)-(iv) hold in proposition 2, while (v) holds because of (??) and (vi)-(vii) hold by construction.

Conversely, consider an equilibrium. Because of free entry, total output Y must be equal to total wages. Consider any other allocation. Let \tilde{y}_k be the mean skill level of firm k in that allocation, $k \in [0, 1]$. Let $d\mu_k$ be the distribution of workers hired by firm k . Then free entry implies $a(\tilde{y}_k) \leq \int w(y)d\mu_k$. Aggregating over firms yields $\int_0^1 a(\tilde{y}_k)dk \leq \int \int w(y)d\mu_k dk = \int w(y)d\mu = Y$. Thus, total output cannot exceed the equilibrium one.

Q.E.D.

Since total output is bounded, there will always exist an allocation that maximizes it. Therefore, there always exists an equilibrium and it is typically "unique"¹³.

In particular, there may well be increasing returns to skills over some range, and this does not prevent an equilibrium from existing. Why? The main difference between that model and a standard production function with increasing returns to *scale* is that it is impossible for a firm to take advantage of increasing returns by simply replicating itself. The size of a firm is fixed by assumption, so we have increasing returns to worker quality but not to scale. This has important consequences for equilibrium. Under economies of scales, replication always leads to an increase in profits because revenues increase more than proportionally to size, while costs increase only proportionally. Hence an equilibrium with zero profits cannot exist since a deviator, by increasing its size, always gets positive profits. Here to increase worker quality we need to hire a different *mix* of people, which bids up the wages of the most skilled, thus defeating, in equilibrium, the firm's attempt to increase its profits. In other words, while in the standard case increasing returns makes equilibrium incompatible with perfect competition, this is not the case here.

2.5 Examples and counter-examples

The above framework applies to a variety of setups, including the one we shall see in the next section. One example is Kremer's (1993) o-ring production function, where a firm's output is given by $\prod_{i=1}^n h_i^{\varepsilon/n}$, where each firm has n members and h_i is the human capital of individual i , as defined as the probability that the corresponding task is performed properly. This can be rewritten as $\exp(\varepsilon\bar{y})$, where an individual i 's "skill level" is defined as $y_i = \ln h_i$, and \bar{y} is the firm's average skill level. Hence it corresponds to the case where our $a(\cdot)$ is exponential. The average output schedule is then clearly convex throughout the whole interval of values of y , meaning that the

¹³In the sense that the clusters and the distribution of firms' average skill levels are uniquely determined.

equilibrium is always hypersegregated, or assortative (Figure 2). Firms only employ workers of one type, and there is a continuum of firms indexed by the worker type that they hire, with a density of firms proportional to the density of workers. The wage of any worker type is simply equal to average output in the firms that employ this type of workers.

By contrast, the above model cannot be applied when one cannot write output as a function of a single aggregate of the firm's employees' skill levels. This will be the case whenever there exists asymmetries across workers within the firm, for example if they perform different tasks. This is the case in the important class of hierarchical models where the typical prediction is that the most talented workers will be assigned to supervisory tasks with the span of control increasing with talent (See Calvo and Wellisz (1979) and Rosen (1982)). In this case total output will not only depend on the average skill level within the firm but also on some higher moments. Consider the case, for example, where a fraction θ of the firm's employees are assigned to supervisory tasks, the remaining workers being assigned to production task. Assume output is proportional to $\bar{y}_s \bar{y}_p$, where \bar{y}_s is the average skill level of supervisors and \bar{y}_p the average skill level of production workers. Then, if $\theta < 1/2$, and if the firm hires from a uniform distribution over $[\bar{y}_k - \sigma, \bar{y}_k + \sigma]$, one can show that if the best workers are allocated to supervisory tasks output can be written as $\bar{y}_k^2 + \sigma(1 - 2\theta)\bar{y}_k - \sigma^2\theta(1 - \theta)$. Thus it is now a function of both average skill and its standard deviation, $a(\bar{y}, \sigma)$, which satisfies $a'_1 > 0$, $a''_{12} > 0$, and $a'_2 \geq 0$.

We will briefly discuss the implication of higher order moments appearing in the production function in section 5 below.

3 Application: ideas, networks, and information technologies

We now apply the above framework to the issue that was mentioned in the introduction, namely the effect of advances in information technology on the

distribution of income. We first derive the properties satisfied by the $a(\cdot)$ function corresponding to that specific application, and then, in the next section, study how the distribution of income changes when the parameter that characterizes the efficiency of information technology shifts; we also do other comparative statics exercise, looking at the impact of the supply of skills and of international trade.

3.1 Basic setup

As above the economy is populated by a continuum of agents, whose skill is given by a number $y \in I = [y_0, y_1]$. The distribution of y is still given by measure μ , and the total mass of workers still normalized to 1. An agent with skill y has a physical productivity equal to y , and at the same time an ability to have ideas h (referred to as "creativity"). We assume that $h = cy + b$. Therefore, productivity and creativity are perfectly correlated.

As above each firm hires a mass s of workers. Firms are two things. First they are firms, i.e. they hire people and sell output, maximizing profits. Second they are networks, i.e. they are the space over which informational spillovers take place, i.e. an idea can be applied at no cost to all the workers of the firm. Firms and networks coincide, so that spillovers are internal to the firm. As already mentioned, this coincidence is what we would expect to arise in equilibrium following a Coasian argument¹⁴. It might be interesting, however, to allow for the frontiers of firms to differ from those of networks.

Hence, the size of the firm s is also the range over which ideas can be spread at no cost. Improvements in information technology can therefore be represented by an increase in s : the larger the network, the larger the number of workers who can benefit from a given idea.

The timing of events is as follows:

1. Firms freely enter the market.
2. The labor market operates and hiring takes place. Firm maximize their expected profit. This yields an equilibrium wage schedule $w(y)$. Note

¹⁴Coase (1995).

that wages are set prior to production.

3. A finite number N of randomly drawn workers have an idea. Each firm makes use of its workers' ideas to improve production. Production takes place and wages are paid. We assume that N is large, so that each firm considers that among its employees there will be exactly $n = sN$ workers that will have an idea.

We now describe the idea process and the production process. Each idea is represented by a random variable $z \in [0, 1]$. It is drawn from a cumulated distribution function $Q(z, h)$ where h is the creativity of the worker who has the idea. More creative people have better ideas. To represent that, we assume that as h increases Q changes in such a way that the new distribution dominates the old one in the first-order stochastic dominance sense. That is, we assume $\partial Q / \partial h < 0$.

A very convenient specification is $Q(z, h) = F(z)^h$. It guarantees that an increase in h shifts Q downwards while preserving the boundary conditions $Q(1, h) = 1$ and $Q(0, h) = 0$. Furthermore, it has very convenient analytical properties.

Then, if an idea z is applied to a given network, total output in that network is given by:

$$u(z) \int_0^s y_{ik} di = u(z) s \bar{y}_k,$$

where u is an increasing function of z and, as above, y_{ik} denoted the skill level of worker i in firm k , while \bar{y}_k is the average productivity of network k 's workers. Thus the production function is linear in individual productivity, while any employee's idea can be applied by all workers in the network, so that it acts as a multiplicative shift to the production function. Below we shall correctly refer to $u(z)$ as the firm's *total factor productivity*.

We assume that ideas are not cumulative, that is any network will apply the best idea among its employees. An idea is useless even if it is only marginally worse than another one that can be applied over the same network, while productivities add up so that if a worker's productivity is marginally lower than another's, so are their marginal products.

The number of ideas in a network is n . The cumulative probability distribution of the best idea is then, conditional on the set of workers who have ideas:

$$\Phi_N(z, \{h_i\}) = \prod_{i=1}^n Q(z, h_i) \quad (8)$$

Using our specification we can see that this is equal to $F(z)^{\sum_i h_i}$. Applying the law of large numbers this is approximately equal to $F(z)^{n\bar{h}_k}$, where \bar{h}_k is the firm's average creativity.

Thus we have the convenient property that the only thing that matters for the distribution of the best idea is the firm's average creativity, irrespective of how it is distributed across workers. This makes it possible to apply the results of the previous section.

3.2 The output schedule

To do so, all we have to do is to compute how a firm's expected output depends on its worker's average quality. We get

$$\begin{aligned} Eu(z)s\bar{y}_k &= \left[\int_{z=0}^{z=1} u(z)d(F(z)^{n\bar{h}_k}) \right] s\bar{y}_k \\ &= \left[u(1) - \int_0^1 u'(z)F(z)^{n\bar{h}_k} dz \right] s\bar{y}_k, \end{aligned}$$

where we have integrated by parts. It can be shown that the RHS is *concave* in \bar{h}_k : the marginal return to creativity is decreasing, reflecting our assumption that ideas are bounded. An infinitely creative person is simply sure to have the best idea, but that best idea only boosts output by a finite amount.

The following proposition then characterizes how expected output depends on average skill.

PROPOSITION 4 - A network k 's expected output can be written as

$$E_k(Y) = \phi(s\bar{h}_k)s\bar{y}_k,$$

where $\phi(\cdot)$ is increasing and concave.

PROOF: We have $\phi(s\bar{h}_k) = \left[u(\bar{z}) - \int u'(z)F(z)^{n\bar{h}_k} dz \right]$. As $n = sN$, this is a function of $s\bar{h}_k$. Taking derivatives we get that

$$\phi' = -N \int u'(z)F(z)^{n\bar{h}_k} \ln F(z) dz > 0, \quad (9)$$

$$\phi'' = -N^2 \int u'(z)F(z)^{n\bar{h}_k} (\ln F(z))^2 dz < 0. \quad (10)$$

Q.E.D.

Because h is a linear function of y , expected output is a function of the firm's average skill only. That is,

$$E_k(Y) = \phi(s(c\bar{y}_k + b))s\bar{y}_k = a(\bar{y}_k).$$

This equation defines total output as a function of average worker quality so that the previous section's analysis applies. If creativity and productivity are positively correlated ($c > 0$), then this function exhibits increasing returns to worker quality: $ya'(y) > a(y)$, although marginal returns can be either decreasing or increasing. An increase in productivity by a factor λ increases output by more than λ , because these people will also have better ideas. In the less intuitive case where $c < 0$, then there are decreasing returns.

Note that $a(\cdot)$ would be concave if people differed in their creativity h but not in their productivity y . That is, skill is now redefined as creativity only and total output is now equal to $a(h) = \phi(h)$, where ϕ is concave, as we have seen. This concavity reflects the fact that our production process is in some sense the opposite of the o-ring one. In the o-ring case, output is determined by the worst performance, which leads to a convex dependence in average worker quality.¹⁵ Here it is determined by the best idea, which leads to a concave dependence of total factor productivity on average worker quality. Therefore, if y is the same for all workers, according to Proposition 2

¹⁵To get back to the o-ring case, just change the sign of $u'(z)$ in equations (??) and (??): if u is decreasing in z , it is the worst idea that we are picking up, not the best one. In that case it is clear that ϕ is convex rather than concave—implying that convexity is a property inherent to the worst performer's dominance and does not depend on the specific functional form chosen by Kremer.

there will be a single type of network in equilibrium, and this type's average skill level will be equal to the population average—as in Figure 3. Thus, despite the fact that it is only the best idea that is being applied, the positive spillovers associated with the spreading of ideas lead, to the extent that ideas are bounded, to integrating heterogeneous people into the same networks.

Things are more complicated, however, in the more interesting case where creativity is correlated with productivity. In such a case, the total output function is given by $a(y) = \phi(s(ay + b))sy$, and may be either concave or convex. In principle, thus, many configurations may arise, including dual clusters. However, a plausible configuration is an S-shape form for $a(y)$, i.e. convexity for low values of y and concavity for high values of y . This would be the case, for example, if there were only two possible ideas, a good one z_1 , and a bad one z_0 . Let P be the probability of having a bad idea. Then expected output would then be equal to, as a function of average productivity y :

$$a(y) = sy \left(u(z_1) - P^{s(ay+b)/N} [u(z_1) - u(z_0)] \right). \quad (11)$$

It is straightforward to check that this is first convex and then concave as y rises. More generally, average output can be written as an integral, over various ideas, of expressions such as (??). So it is plausible that it will have the same S-shape, although one can construct examples where it is convex throughout.

This S-shape pattern reflects the fact that at low levels of worker quality, the best idea is very sensitive to the average quality of the workers employed by the firm. By contrast, at high levels of worker quality, one is almost certain to get the best idea, so that expected output increases almost linearly with average worker quality. In between, there is a zone where the elasticity of average output to skills falls as one quickly reaches the zone where the idea potential is almost exhausted.

As illustrated on figure 4, in equilibrium the labor market typically segregates itself into a hypersegregated zone of low skill workers and a unitary

network type of high skill workers¹⁶. Hence, people at the bottom of the income distribution work in homogeneous firms employing similarly skilled workers, while at the top they work in diversified networks where there is, in some sense, a complementarity between the most able workers, whose value is in the ideas they have, and the least able ones, whose value is mostly in their contribution to output.

As the non linearity in production entirely comes from the existence of ideas, it is interesting to reinterpret the wage equation in the following way. Since $a(y) = \phi(s\bar{h}_k)y$, we can apply Proposition 1 to compute wages in firm k , getting:

$$w(y) = \phi(s\bar{h}_k)y + \phi'(s\bar{h}_k)s\bar{y}_k(h - \bar{h}_k). \quad (12)$$

The interpretation we gave of equation (??), decomposing it into a marginal return to skill and the shadow price of bodies, still holds. But there is another, interesting interpretation that can be given to (??). The first term $\phi(s\bar{h}_k)y$ is the worker's return to his productivity, which is simply equal to his physical product in his firm ignoring his intellectual contributions. The second term $\phi'(s\bar{h}_k)s\bar{y}_k(h - \bar{h}_k)$ is a "bonus" paid for creativity. Because of the zero profit condition this bonus has to average to zero, and we can see that it is proportional to the *deviation* between this individual's creativity and average creativity in the firm. Thus people earn more (less) than their physical marginal product depending on whether they are more (less) creative than their firm's average. This way, low creativity people are penalized for the fact that they occupy a job that might be instead held by somebody equally productive but with better ideas. Hence, the wage structure rewards *absolute* productivity, but *relative* creativity, where the basis of comparison is average creativity in the same network.

¹⁶If the line tangent to the average output schedule at the population average of y intersects the vertical axis above the average output schedule, then there is a single cluster which also covers the zone where the average output schedule is convex.

4 The impact of technical change on segregation and income distribution

In this section we will discuss how technical change —broadly defined— affects the distribution of income in our model. We shall assume that $\phi(\cdot)$ is such that the average output schedule $a(\cdot)$ has the S-shape studied in the previous section, so that worker assignment is summarized by Figure 4. More specifically, we assume that $\phi(\cdot)$ has the convenient exponential form:

$$\phi(x) = A(1 - \beta e^{-\alpha x}).$$

This specification is exact in the special case where there are only two possible ideas. It yields the desired S-shape for $a(\cdot)$.

4.1 An improvement in the quality of the workforce

We start with the simplest experiment, namely a change in the underlying distribution of skills. The average output schedule is independent of that distribution, and therefore does not shift. Equilibrium changes, however, because the initial distribution of network types no longer matches the distribution of agents. Figure 5 illustrates what happens when skills improve in the new distribution. The mean skill level of the initial unitary network increases beyond its original value. To restore equilibrium unitary networks must have a greater average skill level, which, given the local concavity of the average output schedule, implies that these networks hire a more diversified set of people. Thus the networks that hire the top of the distribution of income broaden the range of skills they are hiring from. As ideas are "cheaper" they are more willing to hire low quality people. The least skilled, however, remain in hypersegregated networks and gain nothing from the change, as they do not cooperate with more skilled workers in production. Next, there exists a range of skills that were previously confined to hypersegregated networks and can now join a unitary network. These people enjoy a large wage gain. So do those who were previously employed in the unitary network

but relatively unskilled relative to that network's average. Finally, the most skilled suffer from the reduction in the value of creativity and experience a wage drop. Overall, society is less segregated and more equal, but there remains a mass of poor workers who do not benefit from the change, because segregation prevents them from benefitting from the greater flow of good ideas. Indeed, the gap between that group of people and the average income of the rest of the population has increased.

4.2 An increase in network size s

An increase in network size s means that the range of spillovers is greater. We can take it as representing an improvement in information technology. Formally, it is somewhat similar to an improvement in the quality of the workforce, as in both cases a network recruiting any given set of workers will have better ideas. However, the analysis is slightly more complex here, as the average output schedule shifts.

The following proposition tells us that the response of inequality to network size is typically hump shaped.

PROPOSITION 5 - Let $w'(y; s)$ be the slope of the wage schedule at point y . Let $INEQ = \int w'(y, s)d\mu$, an inequality measure equal to the average of the absolute return to skill over I . Then

(i) There exists s_0 and s_1 such that

$$\begin{aligned} s \leq s_0 &\implies \frac{\partial INEQ}{\partial s} > 0; \\ s \geq s_1 &\implies \frac{\partial INEQ}{\partial s} < 0. \end{aligned}$$

Therefore, inequality rises with s for s small enough but falls for s large enough.

PROOF - The slope of the wage schedule for a firm with type \bar{y}_k is given by

$$w'(y; s) = \phi(s(c\bar{y}_k + b)) + \phi'(s(c\bar{y}_k + b))cs\bar{y}_k$$

Taking derivatives with respect to s we get that

$$\frac{\partial}{\partial s} w'(y; s) = (2c\bar{y}_k + b)\phi' + (c\bar{y}_k + b)cs\bar{y}_k\phi'', \quad (13)$$

where derivatives of ϕ are computed at $c\bar{y}_k + b$. As $\phi'' = -\alpha\phi'$, this has the same sign as $(2c\bar{y}_k + b) - \alpha(c\bar{y}_k + b)cs\bar{y}_k$, an expression which we denote by $K(\bar{y}_k, s)$. This expression is quadratic, concave in \bar{y}_k and its maximum value is $1/(\alpha s) + b - b^2\alpha s/4$. This is clearly negative for s large enough, and one then has $\frac{\partial}{\partial s} w'(y; s) < 0 \forall y$, implying $\frac{\partial INEQ}{\partial s} < 0$. This proves statement (ii). As for the minimum value of $K(\bar{y}_k, s)$ as \bar{y}_k increases, since $K(y, s)$ is concave in y it is given by $\min\{K(y_0, s), K(y_1, s)\}$, which is always positive for s small enough. In that case we have $\frac{\partial}{\partial s} w'(y; s) > 0 \forall y \in I$, implying $\frac{\partial INEQ}{\partial s} > 0$, which proves statement (i). Q.E.D.

Hence, contrary to some conventional wisdom, improvement in information technology does not necessarily increase inequality. This is only so below a certain level of technology. This nonmonotonicity is the result of conflicting effects. On the one hand, the most creative workers can spread their ideas over a larger network of economic activity when s increases. This tends to increase their wages relative to others. On the other hand, when networks get larger, the ideas of a given, highly creative worker, are less valuable because it is more likely that somebody else in the network would have had an idea almost as good. That is, past some large network sizes superstars end up competing against each other which eventually pushes down their wages relative to other workers.

What happens to inequality *within* the unitary network? As it is a network with high skill workers the above proposition may not imply that inequality increases within that network. More precisely, the fact that a unitary network may only arise at a mean firm skill level such that $a(y)$ is locally concave puts restrictions on the sign of the RHS of (??). In fact, one can show that if $b < 0$ there exists a zone where the slope of the wage schedule offered by the unitary network may increase. In this case, the model predicts that intra-firm inequality may increase over some range. By contrast, if $b > 0$

this cannot happen.¹⁷ The unitary network is then always in the zone where it has enough ideas so that an increase in network size lowers the return to skills.

Figure 6 summarizes these results. On the horizontal axis there is the average skill level of a firm y . On the vertical axis there is network size s . The first frontier HH is the frontier between hypersegregated and unitary firms.¹⁸ The second one II is the frontier between the zone where local inequality, as measured by $w'(\bar{y})$, increases with s , and the zone where it is reduced.¹⁹ If $b < 0$, II is above HH, so that inequality may increase within a unitary network as s rises. If $b > 0$ II is below HH, so this cannot happen.

Another interesting question is whether an increase in network size could actually make some people poorer. Here, this can only happen at the top of the distribution of income, not at the bottom. Thus, despite the possible increase in the steepness of the wage schedule offered by the unitary network, technical progress in the form of increased network size cannot impoverish the poorest. This is summarized in the following "Anti-Marxist" result:

PROPOSITION 6 (No impoverishment) - Remember that y_0 is the lowest value of y . Then $\partial w(y_0)/\partial s > 0$.

PROOF - If y is in a hypersegregated zone, wages in a firm employing

¹⁷Let us prove this claim. The unitary network has a \bar{y}_k such that $a(\cdot)$ is locally concave, or equivalently

$$2\phi' + cs\bar{y}_k\phi'' < 0,$$

which may be rewritten

$$\bar{y}_k > \frac{2}{\alpha cs}$$

Computing the expression $(2c\bar{y}_k + b) - \alpha(c\bar{y}_k + b)cs\bar{y}_k$ at $\bar{y}_k = \frac{2}{\alpha cs}$ we simply get $-b$. So, if $b > 0$, $\partial w'(y; s)/\partial s < 0$ over the whole range of values of \bar{y}_k where the unitary network may arise, while if $b < 0$ for some distributions μ one may see a unitary network with a value of \bar{y}_k such that $\partial w'(y; s)/\partial s > 0$.

¹⁸It is defined by $s = 2/(\alpha cy)$. The previous footnote implies that above that locus, the average output schedule is locally concave at y , a necessary condition for a firm with average skill y to define a unitary zone. Conversely if that condition holds we can always pick up an distribution μ such that the unitary zone has a mean precisely equal to y .

¹⁹This frontier is defined by $s = (2cy + b)/[\alpha(cy + b)cy]$.

workers of type y are just equal to output per capita, which obviously increases with s as these people, when more numerous, will share better ideas. If y_0 is included in the unitary network, then this network must cover the whole interval of skill levels. Then its mean skill level is equal to the population one, \bar{y} , which is unaffected by an increase in s . Next, we can compute

$$\partial w(y)/\partial s = y\bar{h}_k\phi'(s\bar{h}_k) + \phi'(s\bar{h}_k)\bar{y}_k(h - \bar{h}_k) + \phi''(s\bar{h}_k)s\bar{y}_k\bar{h}_k(h - \bar{h}_k)$$

Because the average output schedule must be concave at $y = \bar{y}_k$, we must have

$$2\phi'(s\bar{h}_k) + sa\bar{y}_k\phi''(s\bar{h}_k) \leq 0,$$

implying, as long as $h < \bar{h}_k$:

$$\partial w(y)/\partial s \geq [h\bar{y}_k + \bar{h}_k(\bar{y}_k - y)]\phi' > 0.$$

. Q.E.D.

4.3 International trade

Although the recent empirical literature on the rise in inequality does not ascribe a big role to foreign trade, it is interesting to discuss how "globalization" may affect the distribution of income in our model.

This economy has a continuum of skill levels and only one good. Therefore, as long as labor is not mobile, international trade should not affect the distribution of income. However, things are different if ideas, in addition to goods, are mobile, that is, if information technology allows ideas to spread across borders. A multinational network can then take advantage of cross-country differences in relative wages to structure its workforce optimally. A country's distribution of income will then be determined by the world distribution of skills. What happens, then, if a high-skill country trades with a low-skill one? It is as if the distribution of skills in both countries suddenly became equal to the world's distribution of skills. To maintain equality between supply and demand, some multinational networks, that have the same

skill composition of similar national networks, will arise. So, in the high-skill country, it is as if there was a reduction in the quality of the workforce; that is, the reverse of the experiment we made above.

Overall, the effects are qualitatively similar to the ones predicted by the Stolper-Samuelson theorem, with two important provisos. First, the crucial mechanism is the mobility of ideas, rather than goods. Second, openness affects the pattern of segregation, and workers who remain in the hypersegregated zone are not affected by the change.

Finally, note that in the model "globalization" affects the distribution of income despite having no effects on relative prices, since there is only one good. This may lead to a reconsideration of the role of international trade in inequality, since it was dismissed empirically precisely because the rise in inequality was not associated with adequate movements in relative prices²⁰.

5 Extensions

In this section we briefly discuss how relaxing some of the assumptions of our models may affect our results. The first two extensions are concerned with the specific ideas/networks application, while the last two deal with the more general framework.

5.1 More than one good

The interpretation of the results in terms of income distribution may be slightly more complex if there is more than one good in the economy. Above we considered that people produced a single homogenous good. We were therefore looking at changes in the distribution of wages expressed in terms of that good (call it good A), rather than in terms of welfare. If there are other goods, and if these goods are less intensive in information technology, then an improvement in IT will lower the relative price of good A.²¹ It is

²⁰See Laurence and Slaughter (1993), for example.

²¹See Cohen and Saint-Paul (1994) for an analysis of the implications of asymmetrical technical change on wages and employment dynamics.

now possible for those workers who gain little when their wage is expressed in terms good A to actually lose in welfare terms. Consider for example an improvement in the economy’s average skill level. We have shown that if the least skilled workers were in a hypersegregated zone, their wages—in terms of good A— did not change. If the relative price of good A falls then their real wage will actually fall, and reallocation of these workers to other sectors may only partially mitigate that phenomenon. The same phenomenon may arise in the case of an improvement in information technology if the rise in the poorest’s wage in terms of good A is small relative to the fall in the price of that good.

5.2 Unbounded ideas

The results we established in the previous section depend on the assumption that output is S-shaped in average skill, which came from the concavity of $\phi(\cdot)$. Such concavity depended on the assumption that ideas are bounded—more precisely that their effect on total factor productivity was bounded—, which implied that as workers got more talented and/or network size increased, the return to creativity eventually fell. Things may be different, however, if ideas are unbounded, in which case the output schedule may be convex throughout. An easy way to formalize unbounded ideas is simply to assume that the $u(\cdot)$ function has a vertical asymptote at $z = 1$, implying that as z converges toward 1, total factor productivity goes to infinity. In that case integration by parts is no longer possible, implying that one cannot prove that $\phi(\cdot)$ is concave.

Therefore, if ideas are unbounded, the zone where inequality falls with network size may disappear.

5.3 Endogenous firm size

Our results clearly depend on the assumption of a fixed firm size; however, the framework can be extended to a variable size. Going back to our general model, we assume that total output is a function of both the number of

workers and their average skill. Let us denote by $a(\bar{y}, s)$ such a function. Then one has to add to the analysis the determination of optimal firm size. One can check by extending the proof of Proposition 1 that the corresponding first-order condition is

$$\frac{\partial a}{\partial s}(\bar{y}_k, s) = \frac{a(\bar{y}_k, s)}{s} - \frac{1}{s} \bar{y} \frac{\partial a}{\partial \bar{y}}(\bar{y}_k, s) \quad (14)$$

This equation determines the firm's optimal size as a function of its average level of human capital. Note that the rest of the analysis is unchanged, in particular the wage schedule offered by a firm still obeys (??), so that our basic results are still likely to hold, allowing for variations in firm size.

The right hand side of (??) is nothing but $-\lambda_k$, the opposite of the shadow price of bodies. Hence, if size is determined optimally, the shadow price of bodies will be equal to the marginal cost of increasing the size of the firm, $-\frac{\partial a}{\partial s}(\bar{y}_k, s)$. Under increasing returns to worker quality the firm will end up in a zone where $\partial a / \partial s < 0$, meaning that an increase in the number of employees will reduce net output. For this to occur it must be that there exists a cost of increasing the size of the firm, say a setup or investment cost, which ends up increasing more than the direct contribution of the extra employees to output.

Equation (??) implies that firms that have the same type, i.e. the same average skill level, will also have the same size, but that size will in general differ across firms. It sounds reasonable to think that firms with a greater average skill level will be larger; this is what happens in Kremer (1993), and is in accordance with the often found result that more talented people have a greater span of control (as in Rosen (1982)). However it all depends on cross derivatives, and on whether positive spillovers are stronger in firms with more talented people. If this is true then the right hand side of (??) will become more negative as \bar{y}_k gets larger, implying that these firms will have a bigger marginal cost of size and therefore will typically²² be larger.

²²By "typically" we actually mean provided the cross derivative $\partial^2 a / \partial \bar{y}_k \partial s$ is not too negative.

In the case of our idea/networks model size can be made variable by deducting a convex setup cost from average output, so that we would have

$$a(\bar{y}, s) \equiv \phi(s(c\bar{y} + b))s\bar{y} - \gamma s^2/2.$$

One can then easily check that the first-order condition determining optimal firm size is equivalent to:

$$\frac{\gamma s}{2} = \phi(s(c\bar{y} + b))\bar{y} + 2sc\bar{y}^2\phi'(s(c\bar{y} + b))$$

It is very likely that this defines a positive relationship between s and \bar{y} , but if ϕ'' is negative enough this may not happen.

5.4 Higher moments

Another important assumption is that the effect of the intra-firm distribution of skills on output can be represented as depending only on an aggregate, additive, index of worker's skills. The results therefore depend on how good that assumption is as an approximation. In more general cases, higher moments of the firm's distribution of skills will intervene, and the results can be substantially changed. Consider for example what happens when the second moment as well as the first enters the production function, i.e. when output in firm k is given by $a(\bar{y}_k, \mu_k)$, where μ_k is the average square skill in the firm, i.e., using the notations of definition 1, $\mu_k = \int_0^s y_{ik}^2 di$. Then one could check that firm k 's willingness to pay for worker y is given by

$$\frac{a(\bar{y}_k, \mu_k)}{s} + \frac{\partial a(\bar{y}_k, \mu_k)}{\partial \bar{y}_k} \frac{y - \bar{y}_k}{s} + \frac{\partial a(\bar{y}_k, \mu_k)}{\partial \mu_k} \frac{y^2 - \mu_k}{s}$$

This is now quadratic rather than linear, and the analysis is far more complex given that the type of a firm is now characterized by two parameters rather than one. It will still be true that the wage schedule is convex if $\partial a/\partial \mu > 0$, but even then, workers may not be assigned to firms in a segregated fashion, since two quadratic curves may cross more than once.

6 Conclusion

The framework developed here is flexible and general enough to be applicable to the analysis of people's assignment and segregation in a variety of settings. Potential examples include schools, neighborhoods and ghettos, and social networks.

On the other hand, as any model, it has some limitations. In particular, as we have seen, in order to capture the hierarchical aspects of the firm's organization, one would need a more flexible specification of spillovers which would make use of higher moments of the intra-firm distribution of skills.

Our application to information technology has yielded a variety of insights; let us insist on two of them, which we believe are most relevant. First, increases in network size have both positive and negative effects on inequality. Under reasonable assumptions, we find that inequality typically rises and then falls as network size increases. This is somewhat consistent with the observation that after years of an upward trend, inequality seems to be falling again in the nineties. Second, in many cases the bottom of the distribution of income was locked into a "hyper-segregated" zone; i.e. they were not interacting with high skill workers in production. This prevented them from benefitting from an increase in the supply of skilled workers, and they were likely to benefit less than other workers from an improvement in information technology. Their real wage could even fall if one considered a multi-good world.

References

- Aghion, Philippe and Peter Howitt (1998) *Endogenous growth theory*, MIT Press
- Becker, Gary (1973) "A Theory of Marriage: Part I", *Journal of Political Economy*, July/August, 81 (4), 813-846
- Bénabou, Roland (1993) "Workings of a city: location, education, and production", *Quarterly Journal of Economics*, 108 (3), 619-652
- (1996), "Heterogeneity, Stratification, and Growth", *American Economic Review* 86 (3), 584-609
- Bound, John and George Johnson (1992) "Changes in the structure of wages in the 1980's: an evaluation of alternative explanations", *American Economic Review*, 82 (3), 371-392
- Calvo, Guillermo and Stanislaw Wellisz (1979) "Hierarchy, Ability, and Income Distribution", *Journal of Political Economy* 87 (5), 991-1010
- Coase, Ronald (1995) "The nature of the firm" in Steven Medema, ed., *The legacy of Ronald Coase in economic analysis*. Vol. 1. Aldershot, U.K.: Elgar, 5-24.
- Cohen, Daniel, and Gilles Saint-Paul (1994) "Uneven technical progress and job destruction", CEPR working paper
- DiNardo, John, and Jorn-Stephen Pischke (1996), "The return to computer use revisited: have pencils changed the wage structure too?" *Quarterly Journal of Economics*, 112 (1), 629-51
- Epple, Dennis and Richard Romano (1998) "Competition between Private and Public Schools, Vouchers, and Peer-Group effects", *American Economic Review*, 88 (1) 33-62
- Fernandez, Raquel, and Richard Rogerson (1996) "Income distribution, communities, and the quality of public education", *Quarterly Journal of Economics*, 111 (1), 135-164
- Frank, Robert H. and Philip J. Cook, (1995), *The winner takes-all society*. New York; London and Toronto: Simon and Schuster, Free Press, Martin

Kessler Books.

Garicano, Luis (1998), *Knowledge-based hierarchies*. Ph. D. Thesis, University of Chicago.

Grossman, Gene and Elhanan Helpman (1992), *Innovation and Growth in the Global Economy*, MIT Press

Juhn, Chinhui, Kevin M. Murphy and Brooks Pierce, "Wage Inequality and the Rise in the Returns to Skill", *Journal of Political Economy*, 101 (3), 410-442

Katz, Lawrence and Kevin M. Murphy, "Changes in relative wages, 1963-1987: Supply and demand factors", *Quarterly Journal of Economics* 107 (1), 35-78.

Kremer, Michael, (1993), "The o-ring theory of economic development", *Quarterly Journal of Economics*; 108 (3), 551-75

Krueger, Allan (1993), "Have computers changed the wage structure?" *Quarterly Journal of Economics* 108 (1), 33-60

Lawrence, Robert Z. and Matthew J. Slaughter (1993) "International trade and American wages in the 1980s: giant sucking sound or small hiccup?", *Brookings Papers on Economic Activity*, microeconomics, 161-210.

Levy, Frank and Richard Murnane (1992), "U.S. earnings levels and earnings inequality: a review of recent trends and proposed explanations", *Journal of Economic Literature* 30 (3), 1333-1381

Lucas, Robert E. (1988) "On the mechanics of economic development", *Journal of Monetary Economics*, 22 (1), 3-42

Murphy, K., A. Shleifer, and R. Vishny, (1991) "The allocation of talent: implications for growth", *Quarterly Journal of Economics*, 106 (2) 503-30

Rioux, Laurence (1999) *L'importance des qualifications dans les nouvelles organisations: quelles conséquences pour le marché du travail?*, Ph. D. Thesis, Ecole des hautes études en sciences sociales, Paris.

Romer, Paul (1990) "Endogenous technical change", *Journal of Political Economy*, 98 (5), S71-102

Rosen, Sherwin (1981), "The economics of superstars", *American Eco-*

conomic Review 71 (5) 845-58

————— (1982) "Authority, Control, and the Distribution of Earnings", *Bell Journal of Economics*, 13 (2), 425-31

———, ——— (1996), "Review of 'The winner takes-all society'", *Journal of Economic Literature* 34 (1) 133-135

Sattinger, Michael (1993) "Assignment models of the distribution of income", *Journal of Economic Literature* 31 (2) 831-880

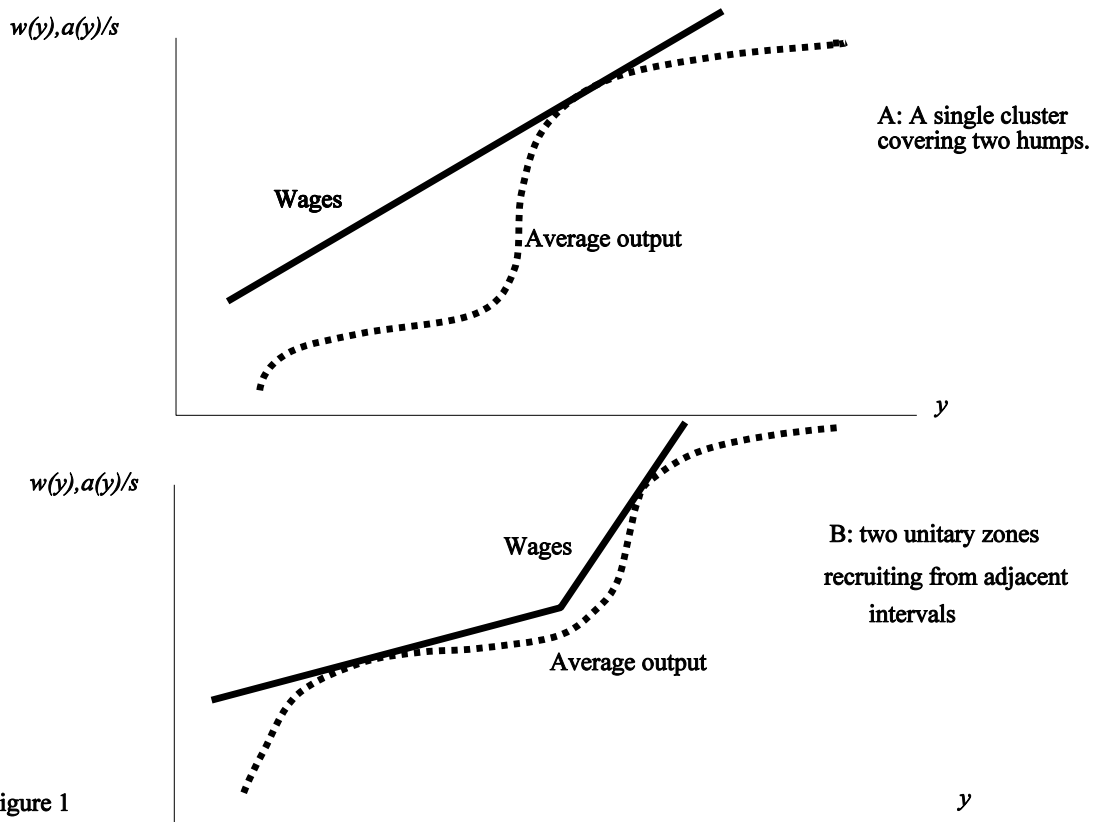


Figure 1

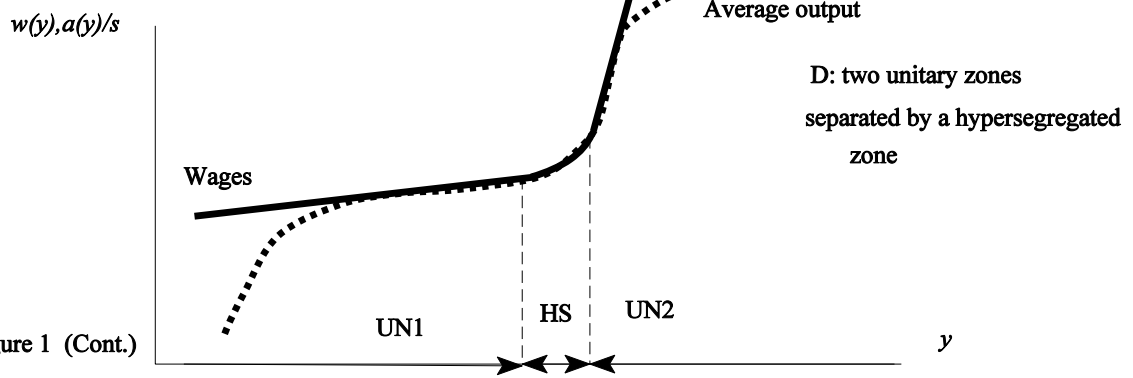
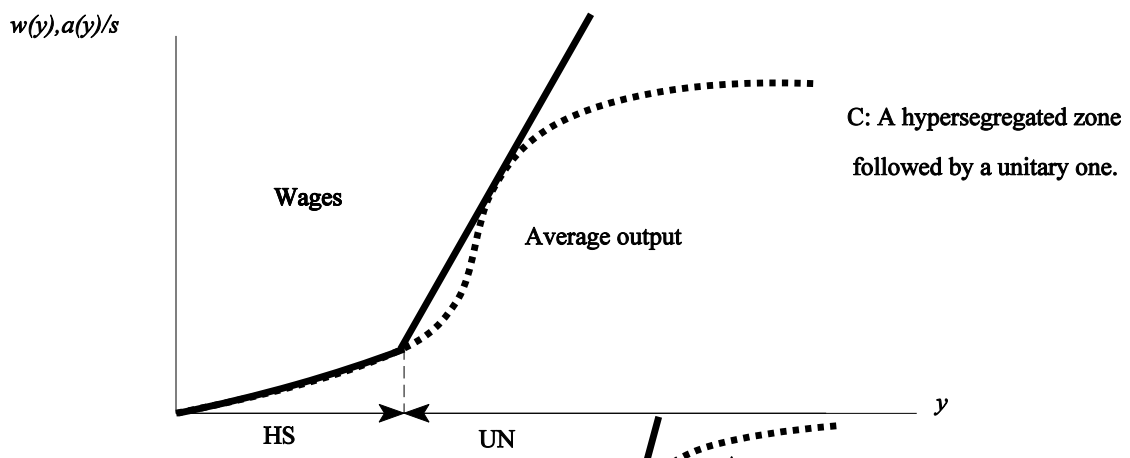


Figure 1 (Cont.)

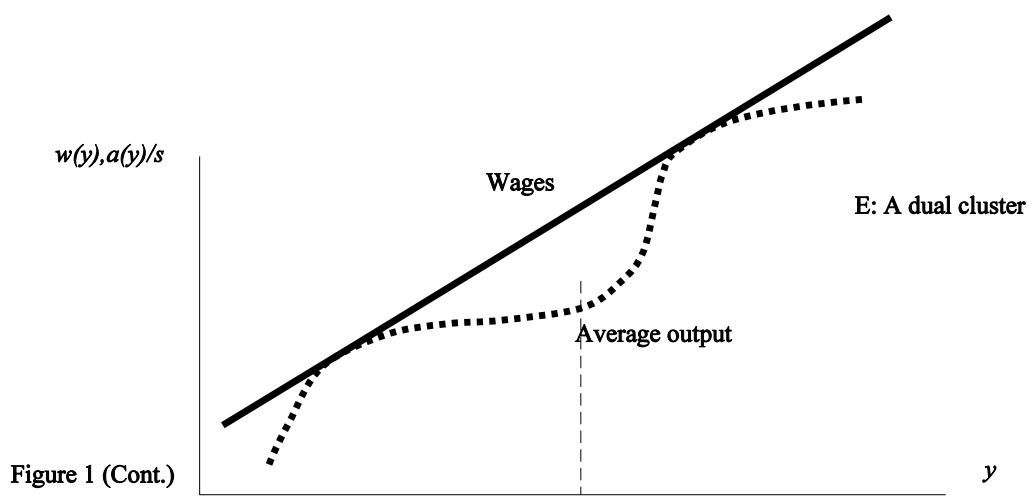


Figure 1 (Cont.)

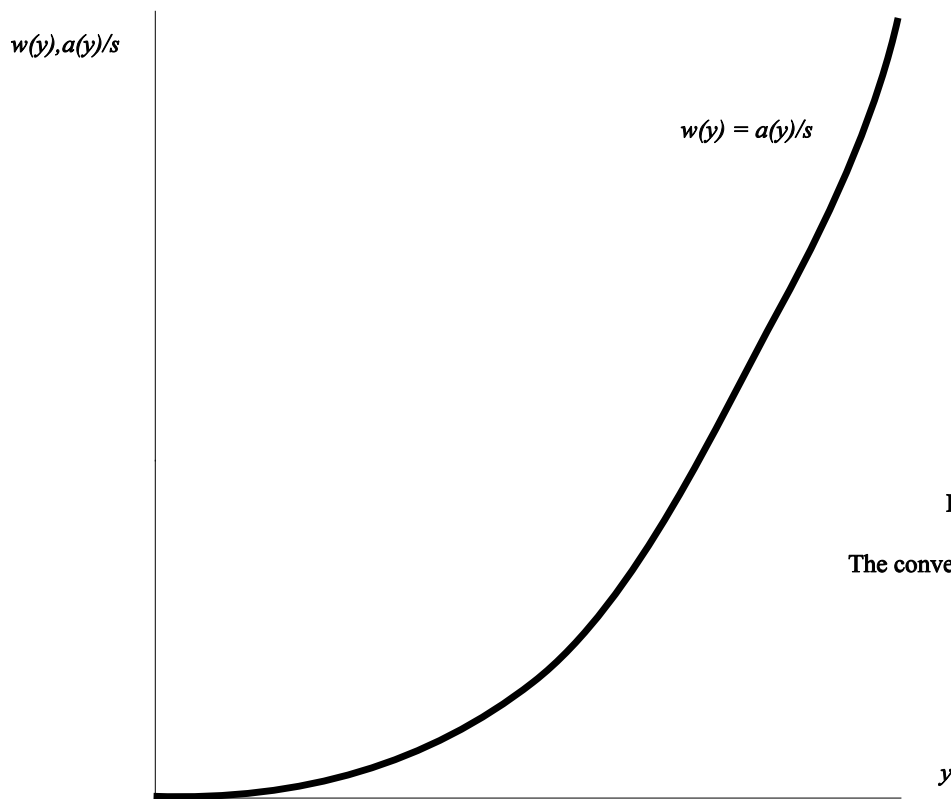


Figure 2
The convex case (O-ring)

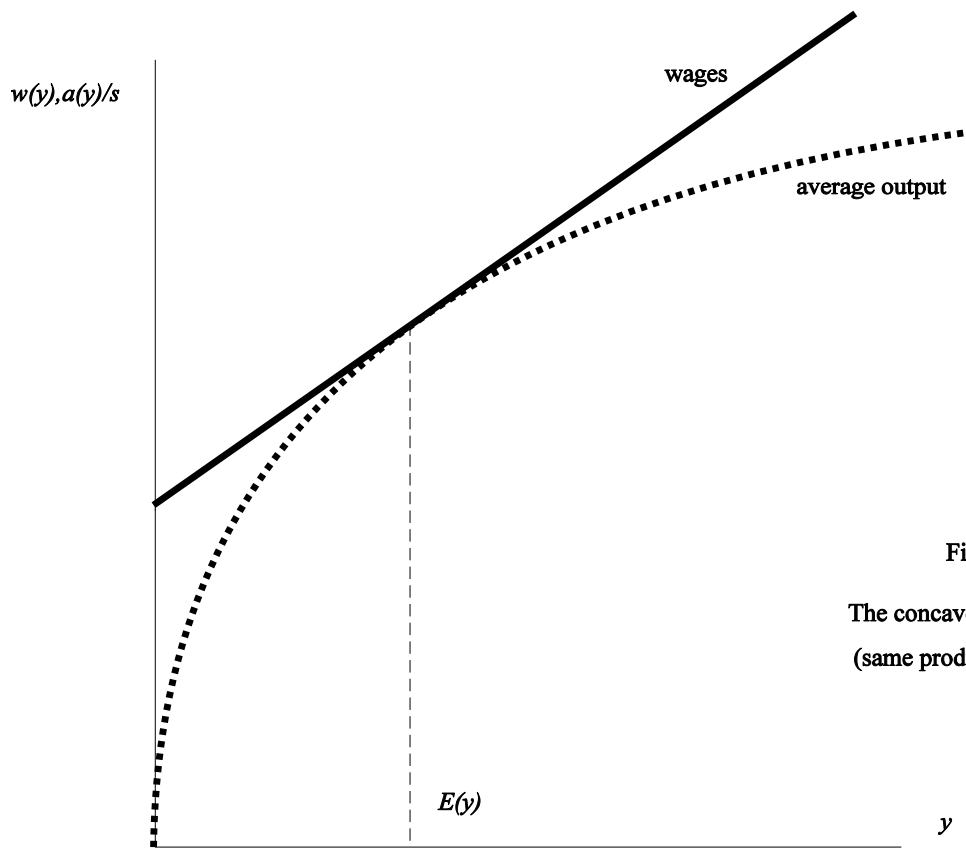


Figure 3
The concave case
(same productivities)

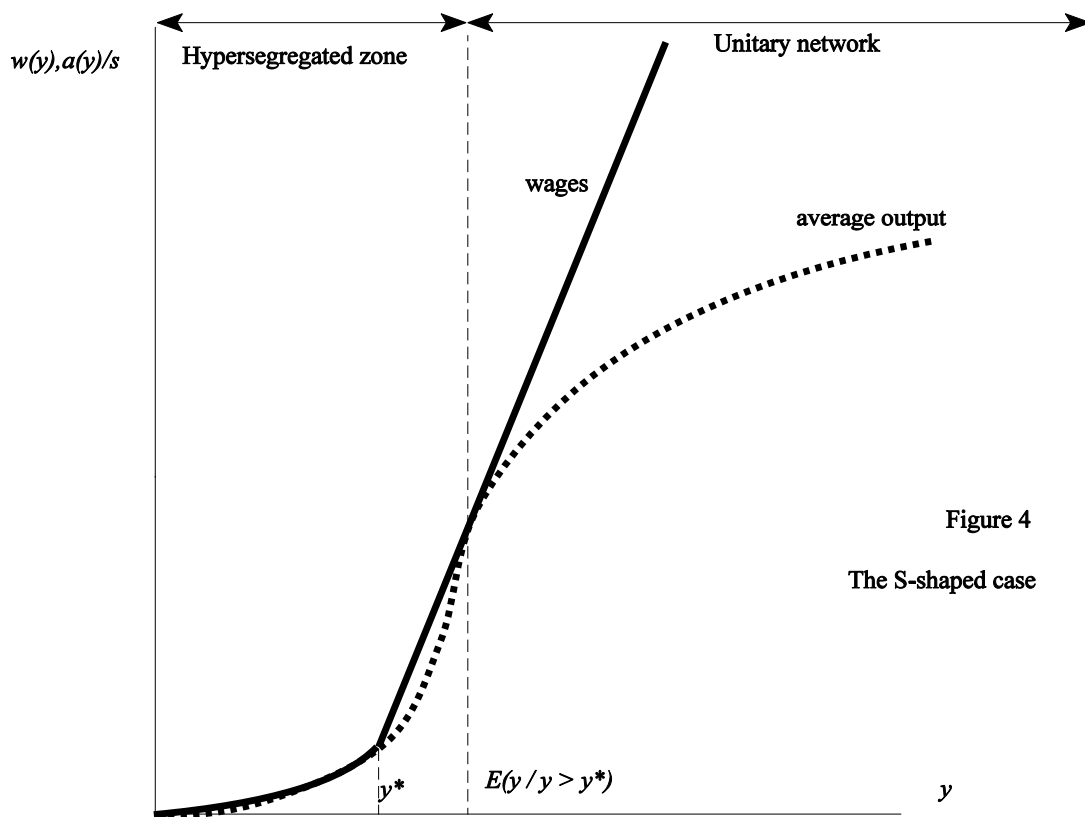


Figure 4
The S-shaped case

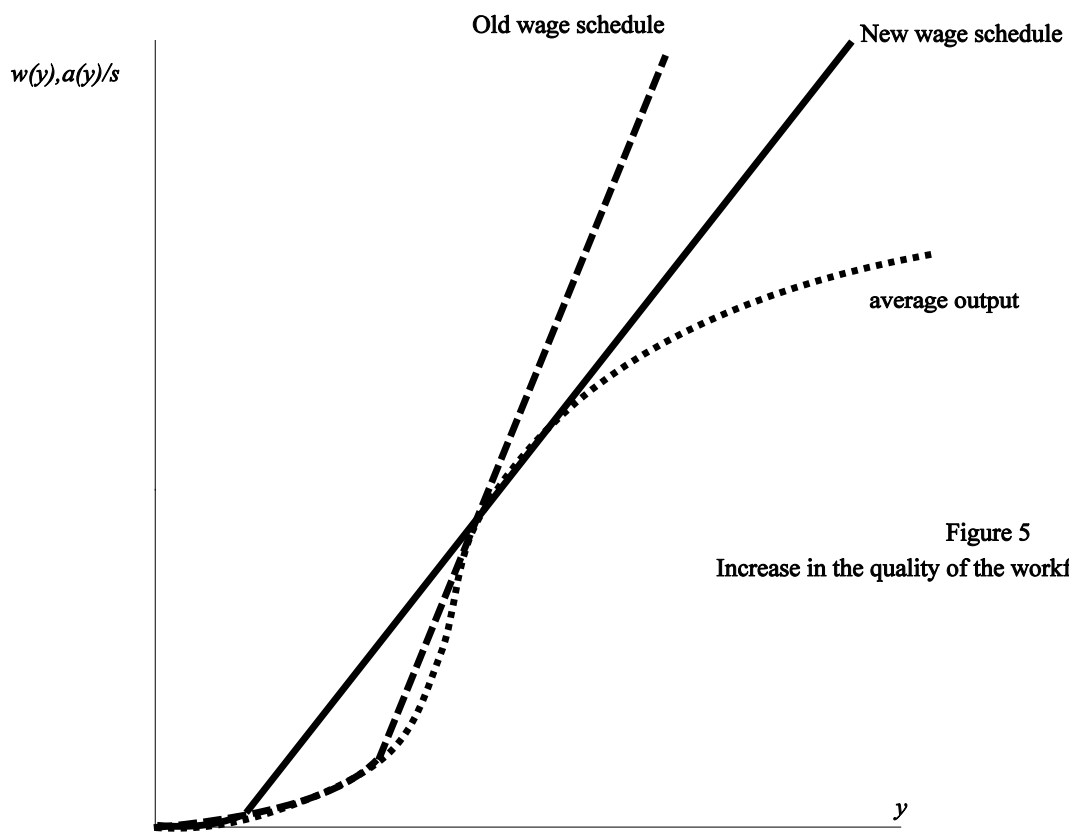


Figure 5
Increase in the quality of the workforce

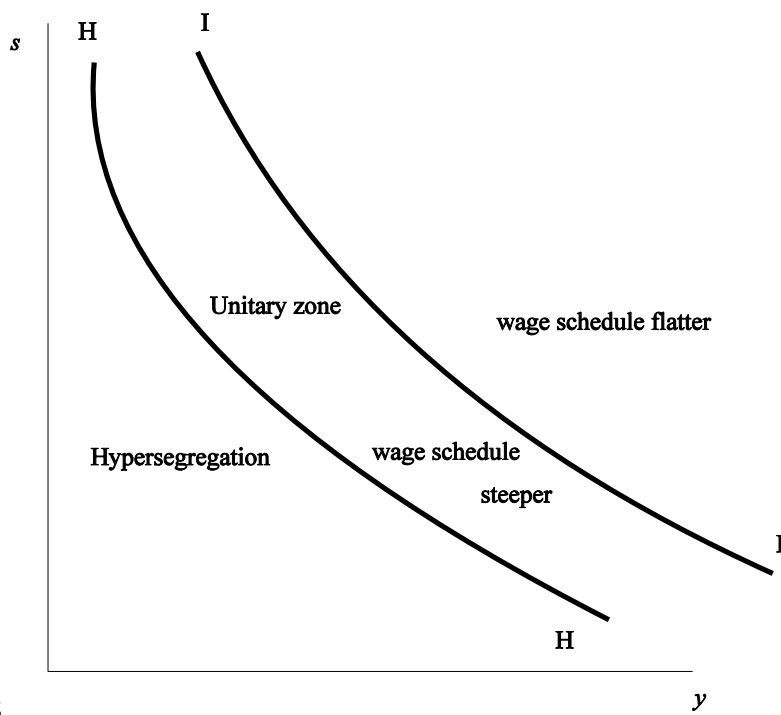


Figure 6

Effect of an increase in s on the slope of the wage schedule in the unitary zone. ($b < 0$)