Technological Changes and Labor Market Segregation

Li, Hao
University of Toronto

Wing Suen
University of Hong Kong

June, 2001

Abstract: International evidence has emerged that workers are increasingly segregated by skill across firms. Using a one-sided assignment model where workers form manager-assistant pairs, with both complementarity and imperfect substitution between skills of the manager and the assistant, Kremer and Maskin (1996) argue that increasing labor market segregation can be explained by increases in the mean and dispersion of workers' skill distribution. We demonstrate that their result depends on the assumption that the skill distribution is sufficiently tight. Changes in the skill distribution of workers can have little effects on the degree of labor market segregation when the distribution is wide. Instead, increasing segregation can be explained by changes in technology that makes output of a firm more sensitive to skills of all workers. Essentially, labor market becomes more segregated as complementarity between high skill workers becomes more important than efficient allocation of scarce supply of high skill workers. Such technological changes can also explain the trend of increasing wage inequalities.

Note: This work is preliminary; do not quote without permission.

1. Introduction

Two noticeable recent trends in labor markets are increasing wage inequality between high skill and low skill workers, and increasing segregation of workers by skill into different firms. Empirical evidence of the first trend is well-documented; see e.g., Katz and Murphy (1992). Evidence for the trend of increasing segregation can be found in Kremer and Maskin (1996) and the references therein. Many economists and social commentators have also argued that increasing wage inequality and increasing labor market segregation reinforce each other.[1]

Increasing wage inequality is often explained by skill-biased technological changes that raise the effective skill-level of high-skill workers (see, e.g., Kahn and Lim, 1994). In particular, the popular view is that recent advances in information technology have increased the demand for high skill workers relative to low skill workers. However, since such skill-biased technological changes presumably occur at a typical establishment (plant), one would expect much of the resulting increases in wage inequality to be within-establishment. A recent paper by Dunne, Foster, Haltiwanger and Troske (2000) shows that this is not the case. Their establishment data indicate that virtually the entire increase in overall dispersion in hourly wages for U.S. manufacturing workers from 1975-92 is accounted for by the between-plant components.

Skill-biased technological changes by themselves cannot explain such empirical findings, nor can they explain the trend of increasing segregation in the labor market. A matching framework, where workers of different skills choose to match with each other to form firms (production teams), is necessary. Ignoring issues regarding changes in the firm size, we can treat segregation as equilibrium outcome resulting from pairwise matching among workers. Each firm (pair of workers) has two production tasks, one for "manager" and the other for "assistant." If skills for the two tasks are perfect substitutes, then the usual assumption of complementarity between tasks, that the same marginal increase in the skill of a worker (regardless of whether he is the manager or the assistant) increases

[1] See the references in Kremer (1997). His empirical investigation casts some doubt on the magnitude of such interactions.

output more if his coworker has higher skill, suggests complete segregation: highest skill workers are matched among themselves, and so are workers of lower skills (Kremer, 1993). If the skills for the two tasks are imperfect substitutes, that is, if output increases more with an increase in the manager's skill than with the same increase in the assistant's skill, optimal matching becomes a non-trivial problem. Workers of the lowest skill are undoubtedly assistants and those of the highest skill managers, but workers of middle skills can be either assistants or managers.

Kremer and Maskin (1996) use the above matching framework to argue that increasing segregation can be explained by increases in the mean and dispersion of the skill distribution in the economy. We agree that there is evidence that both the mean and dispersion have increased modestly in recent decades (see, e.g., Juhn, Murphy, and Pierce, 1993), and that optimal matching pattern generally depends on the skill distribution. However, our result suggests that skill distribution is not a significant factor in segregation and inequality if it is continuous and sufficiently wide.[2] We pursue the alternative direction recognized by Kremer and Maskin in arguing that increasing segregation is explained by changes in technology. In our model with homogeneous production functions, skill levels are defined relative to output elasticities in the production function. Skill distributions are wide when the ratio of the highest skill level to the lowest is large. We recognize that whether skill distributions are sufficiently wide in actual applications is an empirical issue that needs to be resolved, but we believe that it is important to understand the impact of technological changes on segregation and inequality.

Assuming skill distributions that are continuous and sufficiently wide, we show that the degree of segregation instead depends on the trade-off faced by marginal workers who are indifferent between being a manager of a lower skill worker and an assistant to higher skill worker. This trade-off is in turn characterized by the production technology in terms of the comparison between the degree of complementarity between the two tasks and the degree of imperfection in substitution between the manager's skill and the assistant's skill.

[2] The discrete distributions assumed in Kremer and Maskin do not seem to leave enough matching choices for different degrees of segregation. In an early version of their paper, Kremer and Maskin assume continuous skill distribution, but the range of the distribution is narrow and marginal effects of changes in the distribution on segregation may be misleading.

We argue that increasing segregation in recent years can be explained by changes in technology that make output of a firm more sensitive to skill of both the manager and the assistant. Evidence of such technological changes is clear in wide-spread adoption of new information and communication technologies. Managers as well as assistants must master these new technologies in order to gain advantage in economic competitions. As a result, skill requirements for managers and assistants become relatively similar. Increasing segregation of workers by skills is an efficient response to such technological changes, because complementarity between the managerial and assisting tasks becomes more important as relative difference in skill requirements shrinks.

Our model suggests that increasing segregation and increasing wage inequality are consistent and both can be consequences of technological changes that increase output elasticities to skills of manger and assistant. As output elasticities increase, wage as a function of skill also becomes more elastic. While average wage rates may increase, it is workers with higher skills that see greater increases in their wages, and economy-wide wage inequality rises. However, within-firm wage inequality is unaffected by increases of output elasticities to skills of manager and assistant. This is because of the accompanying trend of increasing segregation of workers by skills into different firms. Workers of higher skills earn more, and at the same time their coworkers also have higher skills and earn more. Greater wage inequality is driven by increasing segregation as complementarity becomes more important.

An implication of the contrast between increasing wage inequality across firms and constant wage inequality within-firm is that segregation index should be constructed by using correlation of within-firm education level or other indicators of skill such as job experience, rather than correlation of within-firm wage rates. Indeed, the data assembled by Kremer and Maskin (1996) show that increases in correlation of within-firm wage rates is much less significant than increases in correlation of within-firm education level or job experience. For example, using French data from Kramarz, Lollivier, and Pele (1996), Kremer and Maskin find that from 1986 to 1992, correlation of within-firm wage rates increased from 0.36 to 0.44, while correlation of experience and correlation of seniority within firm exhibited much greater proportional increases from 0.11 to 0.16 and 0.24 to

0.31 respectively.[3]

While we explain increasing labor market segregation and increasing wage inequality as consequences of technological changes that increase output elasticities to both skills of manger and skills of assistant, we admit that changes in the relative elasticities may have occurred too. But it is not obvious in which direction the latter changes have taken place. The popular view of skill-biased technological changes would suggest that output has become relatively more sensitive to the skill of managers, while the total elasticity has changed little. In this case, our model would predict decreasing instead of increasing segregation, because complementarity between manager's skills and assistant's skills becomes less important relative imperfect substitution between the skills at the two tasks. The counter-factual prediction of our model suggests that the driving force behind increasing segregation and increasing inequality is not technological changes biased in favor of skilled workers, but across-the-board increases in skill requirements.

This paper is organized as follows. Section 2 we first use an example of two skill levels to illustrate the point that optimal matching depends on the production technology in terms of how complementarity between manager's skill and assistant's skill compares with imperfect substitution between the two. Then we provide a formulation of the one-sided assignment model with a discrete number of skill levels. Simulation results demonstrate that when the skill distribution is wide enough, the optimal matching pattern is little affected by increases in the mean and dispersion of the distribution, contrasting with Kremer and Maskin's (1996) findings. Section 3 provides some rough characterization results for the case of continuous skill distributions. These results exploit the trade-off faced by workers who can be both manager of a lower skill worker and an assistant to higher skill worker, which is in turn characterized by the production technology in terms of the degree of complementarity between the two tasks and the degree of imperfect substitution between the two. We also apply the results to the case of tight skill distributions considered

[3] Changes in within-firm wage inequality can be accommodated in our model. For example, decreasing within-firm inequality can be explained by relative decreases in output elasticity to the skill of manager compared to elasticity to the skill of assistant. This kind of technological change also increases segregation of workers by skills, but has no effect on across-firm wage inequality as long as overall output elasticity does not change.

by Kremer and Maskin (1996). Section 4 provides a special case that demonstrates how the optimal matching depends on the production technology when the skill distribution is not a factor. The degree of segregation depends positively on the overall output elasticity to skills of manager and assistant, and negatively on the elasticity to skill of manager relative to the elasticity to the skill of assistant. Other comparative statics results of this special case are shown to be consistent with the empirical evidence cited above. Section 5 concludes.

2. Simulation Results with Discrete Skill Distributions

Let us start with the simple example with two-point skill distributions from Kremer and Maskin (1996) to explain the basic "one-sided" assignment framework. Imagine two levels of skill, H and L with $H > L$. For now, let's assume equal and even number of workers of each skill level. Production in a firm (a pair of workers) requires performance of two tasks, assistance and management. Output is given by $F(A, B)$, if a worker with skill A is assigned with the management task and a worker with skill B are assigned with the assistance task.[4]

The optimal matching pattern in this example is either "segregated" (workers with equal skills matched together) or "mixed" (high skill workers matched to low skill workers). Whether segregated or mixed matching is optimal in terms of maximizing total output depends on comparison between two opposite forces. One is complementarity between the managerial and the assistance tasks, captured by the assumption that $F_{12}(A, B) > 0$ for all A and B. Complementarity means that a high skill worker is more productive when matched to another high skill worker. Complementarity favors segregated matching. The opposite force is imperfect substitution between the skills for the two tasks, captured by the assumption that $F(A, B) < F(B, A)$ for all $A > B$. Imperfect substitution means that within a firm the worker with more skill should be the manager, and therefore favors mixed matching.

[4] Although we interpret $F(A, B)$ as quantity of output of a homogeneous good, a quality interpretation of $F(A, B)$ is equally valid.

If skills for the two tasks are perfect substitutes, then complementarity between the two tasks implies that segregated matching always dominates mixed matching, regardless of the skill difference. On the other hand, when the production exhibits no complementarity, then imperfect substitution between the two tasks implies that mixed matching always dominates segregated matching, regardless of the skill difference. These two claims can be seen as follows. Complementarity implies that

$$F(H,H) - F(H,L) = \int_L^H F_2(H,B)dB > \int_L^H F_2(L,B)dB = F(L,H) - F(L,L).$$

The above is the familiar result of two-sided assignment models (e.g., Becker, 1981) that positive assortative matching maximizes total output when there is no task reversal (increasing output from $F(H,L)$ to $F(L,H)$ by switching the task of the two workers). With complementarity and perfect substitution between the two tasks, $F(H,L) = F(L,H)$ implies that

$$F(H,H) + F(L,L) > 2F(L,H),$$

so segregated matching dominates mixed matching. On the other hand, with imperfect substitution but no complementarity between the two tasks, the above inequality of the positive assortative matching result becomes equality. Then $F(H,L) < F(L,H)$ implies that

$$F(H,H) + F(L,L) < 2F(L,H).$$

In this example with two-point skill distribution, when both complementarity and imperfect substitution are present, whether segregated or mixed matching dominates depends on the skill difference. Imperfect substitution between the two tasks is more important when skill difference is small, as it is more efficient to distribute the productivity gains from the increase in the skill of two workers by making them managers in two separate firms than to concentrate the gains in a single firm. The opposite is true when the difference is great, because it becomes more efficient to exploit the complementarity between the two tasks by pairing two high skill workers in one firm. To see this in a precise example, suppose that the production function takes the following homogeneous form

$$F(A,B) = A^{\alpha} B^{\beta},$$

where $\beta > \alpha > 0$. Whether optimal matching is mixed or segregated depends on how $2L^{\alpha}H^{\beta}$ compares with $H^{\alpha+\beta} + L^{\alpha+\beta}$. Since the production function is homogeneous, the comparison comes down to $2\delta^{\beta}$ versus $1 + \delta^{\alpha+\beta}$, where $\delta = H/L \geq 1$. We can verify that the difference between the two terms is zero at $\delta = 1$, increases monotonically as δ increases, reaching maximum at $\delta = (2\beta/(\alpha+\beta))^{1/\alpha}$, and then decreases monotonically, becoming negative eventually.[5] Thus, the optimal matching pattern is mixed for small skill differences and segregated when skill difference is large enough.

This example with two-point skill distribution is illustrative of Kremer and Maskin's result that increasing labor market segregation can be explained by the increases in the mean and dispersion of the skill distribution. However, the analysis is limited by the assumption of two skill levels. In a more general setup, segregation by skill cannot be simply described as either segregated or mixed. It turns out that when the skill distribution is more general, the tradeoff between complementarity and imperfect substitution of the two tasks does not depend on dispersion in the skill distribution in a crucial way. We will show that the tradeoff between complementarity and imperfect substitution is instead characterized by the production function. Indeed, the above example with two-point skill distribution offers a glimpse of results to follow. With a constant-elasticity production function $A^{\alpha}B^{\beta}$, for any given skill ratio $\delta = H/L$, the output difference between mixed matching and segregated matching is proportional to the difference between $2\delta^{\beta}$ and $1 + \delta^{\alpha+\beta}$. As the elasticities α and β change, the tradeoff between complementarity and imperfect substitution is altered. For example, if imperfect substitution of the two tasks becomes more important relative to complementarity between the two, in that β increases relative to α while the total elasticities $\alpha + \beta$ remain unchanged, then the output difference between mixed matching and segregated matching increases and mixed matching becomes relatively more advantageous. Since this holds for any two skill levels in the distribution, the tradeoff between complementarity and imperfect substitution as captured by the output elasticities is likely to be invariant to the skill distribution.

[5] The ratio $\delta = (2\beta/(\alpha+\beta))^{1/\alpha}$, at which the output difference is maximized, will be useful in section 3 where we attempt to characterize the optimal matching pattern for continuous skill distributions.

In a more general setup with more than two skill levels, it is useful to use the correspondence between the solution to the social planner's problem of maximizing total output and equilibrium in terms of no pairwise deviation along the lines of assignment game (Koopmans and Beckmann, 1953, Shapley and Shubik, 1972). Consider a model with n skill levels s_i, $i = 1, \ldots, n$. Let $f_{i,j} > 0$ be the output produced by worker i and worker j, with i as the assistant and j as the manager. For simplicity, assume that unmatched workers have zero outside wage.[6] The social planner's problem is:

$$\max_{x_{i,j}} \sum_{i=1}^{n} \sum_{j=i}^{n} x_{i,j} f_{i,j},$$

subject to

$$\sum_{i=1}^{n} \sum_{j=i}^{n} x_{i,j} \leq 1$$

for all i, and $x_{i,j} \geq 0$ for all i and j. The variable $x_{i,j}$ can be interpreted as the probability that worker i is matched, as an assistant, with worker j.[7] Note that $x_{i,i} \leq \frac{1}{2}$, where equality means self-matching for worker i. The dual problem is:

$$\min_{w_i} \sum_{i=1}^{n} w_i$$

subject to

$$w_i + w_j \geq f_{i,j}$$

and $w_i \geq 0$ for all i and j. A solution to the dual problem gives a wage profile such that no pair of workers have incentives to deviate from the optimal matching. The correspondence between optimality in a social planner's problem and equilibrium in a decentralized economy is then established through the duality theorems in linear programming.

The difference between the above one-sided assignment problem and the standard assignment game is that, unlike in the assignment games, there is no guarantee that the

6 Since self-matching is feasible, i.e. workers are "divisible," this assumption implies that in the optimal matching no workers remain unmatched.

7 In the solution to the above linear programming problem, we will necessarily have $x_{i,j} = 0$ for all $i > j$. An alternative way of formulating the optimal matching problem is to make substitutions in the objective function and the constraints so that only those variables $x_{i,j}$ with $j \geq i$ appear.

solution $x_{i,j}$ to the social planner's problem takes the value of 0 or 1. That is, the extreme points of the constraints in the social planner's problems may not take integer values of $x_{i,j} = 0, 1$. The problem has to do with one-sidedness of assignment. That is, the solution may require "dividing" a worker into a few workers. For example, suppose there are 3 workers of the same skill. If workers are indivisible, one of them must remain unmatched and earn zero. Since any of the two matched workers can work with the unmatched and get the total output, there is no stable matching. But the linear programming problem above yields the solution of $x_{1,2} = x_{1,3} = x_{2,3} = \frac{1}{2}$. Thus, under the indivisibility assumption, equilibrium may not exist and the optimal matching may not be decentralized. Note that such problem does not exist in a continuous model—the variable $x_{i,j}$ can be interpreted instead as the proportions of workers with the two skill levels that are matched.

The above linear programing problem can be programed and solved by computer. To get a feeling of the optimal matching pattern, consider the following example. Suppose that production function is $F(A, B) = AB^2$, and skill levels are indexed by integers from 1 to N, with equal number of workers at each level. If $N = 10$ for example, the optimal matching pattern is self-matching for workers at skill level 1 and 2, workers at level 3 matched to those at level 4, 5 to 6, 7 to 9 and 8 to 10. This matching pattern exhibits five non-overlapping "classes," where a class is defined as a group of overlapping matches. If N increases to 11 while number of workers at each level remains equal, the optimal matching pattern becomes self-matching for workers at skill level 1, 2 to 3, 4 to 5, 6 to 7, 8 to 10 and 9 to 11. There are still five classes, but they are reshuffled, with workers at some skill levels joining an adjacent lower class or higher class. What is happening here is that workers at skill level 11 are incorporated in such a way to keep the within-firm ratio of manager's skill level to assistant's skill level roughly constant. Some abrupt adjustments in terms of changes in the ratio are unavoidable due to the constraint of skill distribution under pairwise matching, but they are made at the low end of the skill distribution rather than at the high end. Dispersion in the skill distribution has increased, but there is no reason to say that there is a higher degree of segregation. If N increases to 12, the optimal matching pattern becomes 1 to 1, 2 to 2, 3 to 4, 5 to 7 and 6 to 8, 9 to 11 and 10 to 12. Number of classes remains five, with more reshuffling, but again there is no reason to say segregation has increased. Indeed, the average ratio of manager's skill to assistant's skill within firms has remained roughly constant—it's 1.21 when $N = 10$, 1.24 when $N = 11$ and 1.25 when $N = 12$.[8] Now suppose that the production function becomes $F(A, B) = A^2B^4$. When $N = 10$, the optimal matching pattern is self-matching for workers at skill level 1, 2, 3 and 4, and workers at level 5 matched to those at 6, 7 to 8, and 9 to 10. There are now seven classes, up from five when the production function is AB^2. There is a clear sense that segregation by skill has increased. The average ratio of manager's skill to assistant's skill within firm is now just 1.09, much lower than 1.21 when the production function is AB^2.

We will argue below that changes in segregation of workers by skill are not driven by changes in the dispersion of the skill distribution. Instead, changes in the production technology—in particular, increases in output elasticities to skill levels of both managers and assistants— are behind the trend of increasing segregation.

3. Characterization Results with Continuous Skill Distributions

Let the production function be $F(A, B)$ with positive partial derivatives F_1 and F_2, and positive cross derivatives F_{12}. Suppose that skills are distributed on an interval $[\underline{s}, \overline{s}]$ with a continuous density.

It is useful to imagine that the problem of optimal pair-wise matching is solved in two steps. In the first step, the skill distribution is decomposed into a distribution for assistants and another distribution for managers. Because the manager has higher skill than the assistant in each pair of in the optimal matching, the sub-distribution for managers stochastically dominate that for assistants. In the second step, pair-wise matches are made between the two sub-distributions. The second step is the standard two-sided matching problem of Koopmans and Beckmann (1953), Shapley and Shubik (1972), and Becker (1981). Optimal matching is positive assortative, in that the most skillful manager is

[8] Later in section 4 we derive a manager-assistant skill ratio r_* from parameters of the production function. It turns out that for the production function AB^2, this ratio r_* is about 1.26. We conjecture that in the optimal matching problem with skills at integer levels from 1 to N, the average ratio of manager's skill to assistant's skill within firms converges to 1.26 as N becomes arbitrarily large.

matched with the most skillful assistant, and so on. Let μ be the optimal one-to-one matching function, whose domain is given by the sub-distribution for managers.

The difficulty of this optimization problem lies in the first step of breaking up the skill distribution into sub-distributions for assistants and managers. Instead of attacking the problem directly, in the remainder of this section we provide some rough characterization results by exploiting the trade-off faced by workers who can be both manager of a lower skill worker and an assistant to higher skill worker. This trade-off should in turn be characterized by the technology in terms of the comparison between the degree of complementarity between the two tasks and their differential sensitivities to skill.

Define a "marginal manager" as a worker with skill level s who is indifferent between being a manager to a worker with a lower skill and being an assistant to a worker with a higher skill, such that workers with skill levels just below s weakly prefer to be assistants and those with skill levels just above s weakly prefer to be managers. Define a "marginal assistant" as a worker with skill level s who is indifferent between being a manager and an assistant, such that workers with skill levels just below s weakly prefer to be managers and those with skill levels just above s weakly prefer to be assistants. A marginal firm is one with a marginal manager and/or marginal assistant. Note that a worker can be both a marginal manager and a marginal assistant. A "regular manager" is a worker with skill level s such that all workers with skill levels in a neighborhood of s are managers. "Regular assistants" are defined in a similar way. Firms with regular managers and assistants are called regular firms. We assert without proof that the optimal matching function μ is piece-wise continuous in that each manager is either marginal or regular, and the same is true for each assistant.

For any distribution of skills, optimal matching $\mu(\cdot)$ yields at least one marginal manager, which is $\mu(\underline{s})$. Workers with skill levels on interval $(s, \mu(\underline{s}))$ are assistants. Otherwise, if they were managers, their assistants would have skill levels higher than the assistant of the worker with skill level $\mu(\underline{s})$. Then a switching of partners would increase total output. By piece-wise continuity of the matching function, workers with skill levels just above $\mu(\underline{s})$ are managers. It then follows that $\mu(\underline{s})$ is a marginal manager. If the range of skill distribution is too small, there may not be marginal assistants.[9] Using the definitions of

[9] This is illustrated later in Proposition 3.3.

marginal managers and marginal assistants, we can put some restrictions on matches in marginal firms.

LEMMA 3.1. *Suppose that $\mu(\cdot)$ is optimal. Then, for any marginal manager with skill level s, $F_1(s, \mu(s)) \leq F_2(\mu^{-1}(s), s)$; for any marginal assistant with skill level s, $F_1(s, \mu(s)) \geq F_2(\mu^{-1}(s), s)$.*

PROOF. Let s the skill level of a marginal manager, with assistant $\mu^{-1}(s)$ and manager $\mu(s)$. Let $s' < s$ be a skill level close to s. By assumption s' is an assistant in the optimal assignment, matched with $\mu(s')$. Consider a perturbation of the optimal matching where s' is matched as a manage with $\mu^{-1}(s)$ and s as an assistant with $\mu(s')$. We must have:

$$F(\mu^{-1}(s), s) + F(s', \mu(s')) \geq F(\mu^{-1}(s), s') + F(s, \mu(s')).$$

The stated condition in the lemma is necessary for the above to hold for any s' arbitrarily close to s. The other claim can be shown in a similar way.

Q.E.D.

Lemma 3.1 means that for marginal managers with skill level s, marginal product of as a manager exceeds marginal product as an assistant, and that the opposite is true for marginal assistants. Since in equilibrium, wages are determined by marginal product, Lemma 3.1 implies that wage function is convex at the skill level s of marginal workers (managers or assistants) if there is a kink at s. This is suggestive of how segregation should be measured under continuous skill distributions. Although generally it is unlikely that a single index can capture complexities of optimal matching pattern, the economic significance of segregation has much to do with the perception that skilled workers earn high wages by working each other. In our model, a marginal worker separates regular workers with skills lower than his own from those with higher skills. Since the wage schedule is smooth for skill levels corresponding to regular workers, convexity of the wage schedule at the skill of the marginal worker suggests that it is appropriate to think of the marginal worker as a boundary of skill segregation.

Lemma 3.1 is shown with a perturbation argument. The interpretation becomes clear if we consider a production function that takes the following homogeneous form

$$F(A, B) = A^{\alpha} B^{\beta},$$

with $\beta > \alpha > 0$. Homogeneity makes it sensible for us to focus on the ratio of manager's skill to assistant's skill in a typical firm. The inequality in Lemma 3.1 for a marginal manager s becomes:

$$\left(\frac{s}{\mu^{-1}(s)}\right)^{\alpha} \left(\frac{\mu(s)}{s}\right)^{\beta} \leq \frac{\beta}{\alpha}.$$

Thus, with a homogeneous production function, this inequality can be rewritten so that it involves ratios of manager's skill to assistant's skill alone. Lemma 3.1 then implies an upper bound on the product of the two ratios of manager's skill to assistant's skill in the two marginal firms associated with a marginal manager. For a marginal assistant, Lemma 3.1 gives a lower bound on the product of the two ratios in the two marginal firms:

$$\left(\frac{s}{\mu^{-1}(s)}\right)^{\alpha} \left(\frac{\mu(s)}{s}\right)^{\beta} \geq \frac{\beta}{\alpha}.$$

We have argued that it is appropriate to think of the marginal worker as a boundary of skill segregation. With homogeneous production functions, the degree of segregation can be gauged by looking at the manager-assistant skill ratios in marginal firms. Lemma 3.1 therefore imposes upper bound and lower bound on how segregated the optimal matching is.

Lemma 3.1 can be generalized to matches in regular firms. For any s, define

$$r_f = \text{argmax}_r s^{\alpha}(rs)^{\beta} - \frac{1}{2}(s^{\alpha+\beta} + (rs)^{\alpha+\beta});$$

$$r_b = \text{argmax}_r (s/r)^{\alpha} s^{\beta} - \frac{1}{2}(s^{\alpha+\beta} + (s/r)^{\alpha+\beta}).$$

In words, for a fixed skill level s of an assistant, the ratio r_f maximizes the output difference between matching a manager of skill level $r_f s$ with the assistant and self-matching. Similarly, for a fixed skill level s of a manager, the ratio r_b maximizes the output difference between matching an assistant of skill level s/r_b with the manager and self-matching. Due the homogeneity of the production function, the above definitions do not depend on s. We can show that

$$r_f = \left(\frac{2\beta}{\alpha+\beta}\right)^{1/\alpha};$$

$$r_b = \left(\frac{\alpha+\beta}{2\alpha}\right)^{1/\beta}.$$

The above follows because $s^\alpha(rs)^\beta - (s^{\alpha+\beta} + (rs)^{\alpha+\beta})/2$ is concave at $r = (2\beta/(\alpha+\beta))^{1/\alpha}$, and monotone on both sides of the point. Similarly, the function $(s/r)^\alpha s^\beta - (s^{\alpha+\beta} + (s/r)^{\alpha+\beta})/2$ is concave at $r = (2\alpha/(\alpha+\beta))^{1/\beta}$, and monotone on both sides of the point. Define

$$r_* = (\beta/\alpha)^{1/(\alpha+\beta)}.$$

Then, we have $r_b < r_*$ and $r_* < r_f$, so that

$$r_b < r_* < r_f.$$

LEMMA 3.2. *Suppose that the production function is given by $F(A,B) = A^\alpha B^\beta$ with $\beta > \alpha > 0$. If a matching function $\mu(\cdot)$ is optimal, then there do not exist s_1 and s_2 with $s_1 < s_2 < \mu(s_1) < \mu(s_2)$ such that $s_2/s_1 \geq r_f$ and $\mu(s_2)/\mu(s_1) \geq r_b$; there do not exist s_1 and s_2 with $s_1 < \mu(s_1) < s_2 < \mu(s_2)$ such that $s_2/s_1 \leq r_f$ and $\mu(s_2)/\mu(s_1) \leq r_b$.*

PROOF. Suppose $s_1 < s_2 < \mu(s_1) < \mu(s_2)$. If $s_2/s_1 \geq r_f$, then $\mu(s_1)/s_1 > r_f$. It follows from the definition of r_f that

$$s_1^\alpha s_2^\beta - \frac{1}{2}(s_1^{\alpha+\beta} + s_2^{\alpha+\beta}) > s_1^\alpha \mu(s_1)^\beta - \frac{1}{2}(s_1^{\alpha+\beta} + \mu(s_1)^{\alpha+\beta}).$$

Similarly, if $\mu(s_2)/\mu(s_1) \geq r_b$, then $\mu(s_2)/s_2 > r_b$, and the definition of r_b implies that

$$\mu(s_1)^\alpha \mu(s_2)^\beta - \frac{1}{2}(\mu(s_1)^{\alpha+\beta} + \mu(s_2)^{\alpha+\beta}) > s_2^\alpha \mu(s_2)^\beta - \frac{1}{2}(s_2^{\alpha+\beta} + \mu(s_2)^{\alpha+\beta}).$$

Summing the above two inequalities, we have

$$s_1^\alpha s_2^\beta + \mu(s_1)^\alpha \mu(s_2)^\beta > s_1^\alpha \mu(s_1)^\beta + s_2^\alpha \mu(s_2)^\beta,$$

a contradiction to the optimality of matching s_1 to $\mu(s_1)$ and s_2 to $\mu(s_2)$. The other statement is similar.

Lemma 3.2 extends Lemma 3.1 to regular firms, but is in fact implied Lemma 3.1. To see this, suppose that there are two skill levels s_1 and s_2 with $s_1 < s_2 < \mu(s_1) < \mu(s_2)$ such that $s_2/s_1 > r_f$ and $\mu(s_2)/\mu(s_1) > r_b$, where $\mu(\cdot)$ is optimal. Since workers with skill level s_2 are assistants and workers with skill level $\mu(s_1)$ are managers, piece-wise continuity of the matching function $\mu(\cdot)$ implies that there exists a skill level $s' \in [s_2, \mu(s_1)]$ that workers with skill level s' are marginal managers. Then, by $\mu(s') \geq \mu(s_2)$ and $\mu^{-1}(s') \leq s_1$. It follows that

$$\left(\frac{\mu(s')}{s'}\right)^{\beta}\left(\frac{s'}{\mu^{-1}(s')}\right)^{\alpha} \geq \left(\frac{\mu(s_2)}{\mu(s_1)}\right)^{\beta}\left(\frac{s_2}{s_1}\right)^{\alpha}.$$

By assumptions, the above is greater than $r_b^{\beta} r_f^{\alpha}$, which equals β/α, a contradiction to Lemma 3.1.

For any pair of overlapping matches, with assistants' skill levels $s_1 < s_2$ satisfying $s_2 < \mu(s_1)$, Lemma 3.2 says that the manager-assistant skill ratios in the two matches cannot be too great. Similarly, for any pair of non-overlapping matches, with assistants' skill levels $s_1 < s_2$ satisfying $\mu(s_1) < s_2$, the manager-assistant skill ratios cannot be too small.

The rough characterizations of Lemmas 3.1 and 3.2 can be applied to study the case of tight skill distributions considered by Kremer and Maskin (1996).[10] In one version of their model, the production function is given by AB^2, and the skill distribution is uniform on $[\underline{s}, \overline{s}]$. They prove that when $\overline{s}$ is sufficiently close to $\underline{s}$, the optimal matching such that each worker with skill $s > (\underline{s} + \overline{s})/2$ is a manager matched with a worker with skill $s - (\overline{s} - \underline{s})/2$. In fact, this result can be generalized to any skill distribution and any production function. The idea is that when skill distributions are tight enough, the optimal matching pattern is "median matching," where workers with skill levels below the median are matched as assistants, in a positive assortative manner, to workers with skill levels above the median. The following proposition shows how the results in the previous section can be used to establish this generalization.

[10] The case has also been considered by Legros and Newman (1998).

PROPOSITION 3.3. *Suppose that the production function is* $F(A,B) = A^\alpha B^\beta$ *with* $\beta > \alpha > 0$, *and skills are continuously distributed on* $[\underline{s}, \overline{s}]$, *with* s_m *being the median skill level. If* $\overline{s}/\underline{s} \leq r_b$, *then median matching is optimal; if* $s_m/\underline{s} > r_f$ *and* $\overline{s}/s_m > r_b$, *then median matching is not optimal.*

PROOF. Consider the first claim. Let $\mu(s)$ the optimal matching function. That is, any worker with skill s who is an assistant is optimally assigned to a worker with skill $\mu(s)$. Let s_m be the median level. Clearly, $\mu(\underline{s}) \leq s_m$. Otherwise, since by positive assortative matching all workers with skill levels between $\underline{s}$ and $\mu(\underline{s})$ are assistants, there are not enough workers who can be managers. Similarly, $\mu^{-1}(\overline{s}) \geq s_m$. Suppose that the above two inequalities are strict. Then, we have $\underline{s} < \mu(\underline{s}) < s_m < \mu^{-1}(\overline{s}) < \overline{s}$. From the assumptions of the proposition we have $\mu^{-1}(\overline{s})/\underline{s} < r_b < r_f$ and $\overline{s}/\mu(\underline{s}) < r_b$, which by Lemma 3.2 contradicts the optimality of μ. Thus, $\mu(\underline{s}) = \mu^{-1}(\overline{s}) = s_m$. It follows that median matching is optimal.

Consider the second claim. As above, if μ is the optimal matching function, we have $\mu(\underline{s}) \leq s_m$ and $\mu^{-1}(\overline{s}) \geq s_m$. If median matching is optimal, then $\mu(\underline{s}) = \mu^{-1}(\overline{s}) = s_m$. Since $\overline{s}/s_m > r_b$, by piecewise continuity of μ, there exists $s < s_m$ such that $\mu(s)/s > r_b$. Then, we have $s_m/\underline{s} > r_f$ and $\mu(s)/s > r_b$, which by Lemma 3.2 contradicts the optimality of μ.

Q.E.D.

The above argument is essentially that if the distribution is tight enough, median matching must be optimal because it is not possible to have a marginal assistant. Conversely, non-optimality of median matching follows when the distribution is wide enough so that it is not possible for the median skill level to be a marginal manager. In the latter case, we have

$$\mu(\underline{s}) < s_m < \mu^{-1}(\overline{s}),$$

so that both $\mu(\underline{s})$ and $\mu^{-1}(\overline{s})$ are marginal managers. By piece-wise continuity of μ, there is at least one marginal assistant s exists between $\mu(\underline{s})$ and $\mu^{-1}(\overline{s})$. In general, when the skill distribution is wide enough, we can imagine that the optimal matching is characterized by a sequence of skill levels corresponding to workers who are alternatively marginal managers and marginal assistants.

4. Optimal Constant Skill Ratio Matching

The problem of solving for the optimal matching pattern is difficult without making assumptions on the skill distribution. However, Lemma 3.1 and 3.2 can be used to get some rough idea about segregation. We have argued that with a homogeneous production function, we may measure the degree of segregation by manager-assistant skill ratios in marginal firms. When the skill distribution is wide enough, there can be multiple marginal managers and multiple assistants. Imagine that the optimal matching is characterized by a sequence of skills $\{s_1, s_2, \ldots, s_n, \ldots\}$ such that $s_{n+1} = \mu(s_n)$ for each n and s_n is alternately marginal manager and marginal assistant. Then, using Lemma 3.1 and 3.2 we can show that if the range of the skill distribution is sufficiently wide, the alternate bounds on the ratio of manager's skill to assistant's skill in marginal firms are arbitrarily close to r_*, a ratio between r_b and r_f.

The above claim can be established as follows. Define a function

$$C(r) = \left(\frac{\beta}{\alpha}\right)^{1/\beta} r^{-\alpha/\beta}.$$

This function is decreasing for all $r \geq 1$. Let r_n be the ratio of manager's skill to assistant's skill in the n-th marginal firm with workers of skill levels s_n and s_{n+1}. We know that $r_1 \geq 1$. By Lemma 3.1, since s_1 is a marginal manager, we have

$$r_2 \leq C(r_1) \leq C(1).$$

Applying the two inequalities in Lemma 3.1 in turn, we find that r_n is alternately bounded from above (if s_n is a marginal manager) or from below (if s_n is marginal assistant) by $C^{n-1}(1)$. It is straightforward to verify that $\lim_{n\to\infty} C^{n-1}(1) = (\beta/\alpha)^{1/(\alpha+\beta)}$.

Thus, due to the convergence property of the two inequalities in Lemma 3.1, if there is a sufficiently long sequence of marginal managers and assistants, then bounds on the manager-assistant skill ratio are arbitrarily close to r_*. This leads us to consider matching with a constant manager-assistant skill ratio. With a homogeneous production function, from Lemma 5.1, if optimal matching is characterized by constant-ratio matching for all workers including marginal managers and marginal assistants, then the ratio r must satisfy both $r^{\alpha+\beta} \leq \beta/\alpha$ and $r^{\alpha+\beta} \geq \beta/\alpha$. It follows that the ratio must be equal to $r_* = (\beta/\alpha)^{1/(\alpha+\beta)}$. Let's construct an example where such matching pattern is optimal and study the implications to segregation and wage inequality.

Proposition 4.1. *Suppose that the production function is $F(A,B) = A^\alpha B^\beta$ with $\beta > \alpha > 0$. If pair-wise matching with constant manager-assistant skill ratio r_* is feasible, then it is optimal.*

Proof. We will construct an equilibrium of pair-wise matching with constant ratio r_*. Let $w(s)$ be the wage schedule. For simplicity, we assume that $\underline{s} = 1$.[11] If the matching pattern is optimal we must have:

$$w'(s) = \begin{cases} \alpha(sr_*)^\beta s^{\alpha-1}, & \text{if the worker with skill } s \text{ is an assistant;} \\ \beta s^{\beta-1}(s/r_*)^\alpha, & \text{if the worker with skill } s \text{ is a manager.} \end{cases}$$

Since $r_* = (\beta/\alpha)^{1/(\alpha+\beta)}$, we can write $w'(s) = ks^{\alpha+\beta-1}$, where k is a constant defined by $k = \alpha r_*^\beta = \beta r_*^{-\alpha}$. Integrating $w'(s)$, we have:

$$w(s) = \frac{k}{\alpha+\beta}s^{\alpha+\beta} - \frac{k}{\alpha+\beta} + w(1).$$

Since $w(1) + w(r_*) = r_*^\beta$, we have $w(1) = k/(\alpha+\beta)$, and therefore

$$w(s) = \frac{k}{\alpha+\beta}s^{\alpha+\beta}.$$

It is straightforward to verify that: (i) for any s, $w(s) + w(sr_*) = s^\alpha(sr_*)^\beta$, and $w(s) + w(s/r_*) = (s/r_*)^\alpha s^\beta$; and (ii) no pair of workers can obtain total income greater the sum of their equilibrium income. The second property derives from the definition of the wage function, and the fact that the problem of choosing one's partner is concave so that the second-order condition is also satisfied.

Q.E.D.

[11] The value of $\underline{s}$ does not matter as long as it is strictly positive. If $\underline{s} = 0$, the optimal matching problem is not well-defined.

In the above equilibrium each worker with skill s (not just marginal workers with skills r_*, r_*^2 and so on) is indifferent between being a manager of a worker with skill s/r_* and being an assistant to a worker with skill sr_*. This follows from the homogeneity property of the production function and hence the wage function. Conditions of Lemma 3.1 are satisfied with equalities, not only for marginal firms (so that there are no kinks in the wage function), but for the entire range of skill distribution.

Feasibility of pairwise matching with constant ratio r_* places restrictions on the skill distribution. Consider the interval $[1, r_*^2]$. Workers with skills from 1 to r_* are assistants; those with skills from r_* to r_*^2 are managers. Pairwise matching with constant ratio r_* implies that the density of skill distribution over $[r_*, r_*^2]$ is completely determined given the density over $[1, r_*]$, and vice versa. In general, pairwise matching with constant ratio r_* places serious restrictions on the skill distribution, but still leaves substantial degree of freedom. Moreover, since in equilibrium each worker with skill s is indifferent between being a manager of a worker with skill s/r_* and being an assistant to a worker with skill sr_*, any overlapping of intervals managers and assistants is possible as long as constant ratio matching is maintained. This provides further degree of freedom for distribution specification.

In the equilibrium of Proposition 4.1, optimal degree of segregation in the economy depends on the changes in skill distribution only because constant ratio matching at s_* is not feasible. Given the difficulty of characterizing optimal matching pattern for general skill distributions, and given the arbitrariness of segregation with totally unrestricted skill distribution, in studying the recent trends in segregation by skill and wage inequality, it makes more sense to consider what happens when skill distribution is not a crucial factor than to impose unexplained restrictions on the skill distribution.

Proposition 4.1 implies that when skill distribution is not a factor in determining optimal matching pattern, segregation depends on the trade-off faced by workers who can be both manager of a lower skill worker and an assistant to higher skill worker. This trade-off is in turn characterized by the technology in terms of the comparison between the degree of complementarity between the two tasks and their differential sensitivities to skill.

Now we consider some simple comparative statics, assuming that skill distribution "adapts" to technological changes. The degree of segregation of this economy depends only on the technology parameters α and β. It is increasing in the overall elasticities $\alpha+\beta$, and decreasing in the ratio of two elasticities β/α. The first type of technological change is more reasonable to explain the recent increase in the degree of segregation. As technological changes make output more elastic to skills of both the manager and the assistant, the gains from complementing high skill managers as assistants become more important for high skill workers than the need to place these high skill workers at managerial positions in other firms. Workers with similar skills match together, and segregation increases.[12]

Increasing overall elasticities can also explain the accompanying trend of increasing wage inequality. Consider technological changes such that $\alpha+\beta$ increases but β/α remains constant. In the proof of Proposition 4.1, when $\alpha+\beta$ increases while β/α remains constant, $w(1)$ will not change. This means that the whole wage schedule moves up. Although all workers benefit from an increase in the overall elasticities, higher skill workers benefit more. Economy-wide wage inequality increases in the sense that the ratio of wage rates for any two skills rises.

Although increasing overall elasticities explain both more segregated labor market and greater wage inequality, greater inequality is in fact derived from more segregation rather than a parallel consequence of increasing overall elasticities. This can be seen from the proof of Proposition 4.1: in each firm the manager's share of total output is $\beta/(\alpha+\beta)$, and the assistant's wage is $\alpha/(\alpha+\beta)$. If the optimal matching pattern does not change with $\alpha+\beta$, then the wages of assistants and managers would increase at the same rate as their joint output, and wage inequality would remain unchanged.

Thus, when $\alpha+\beta$ increases but β/α remains constant, wage inequality (in terms of ratio of wages) can increase across firms, but within-firm wage inequality does not change. This is consistent with the findings of Dunne, Foster, Haltiwanger and Troske (2000), cited in the introduction, that most of the recent increases in wage inequality take the form of increasing inequality across firms. In our model, within-firm wage inequality are

[12] The type of technological change is alluded to in Kremer's (1993) O-ring theory of economic development, but he does not consider the issue of segregation.

across-firm inequality are independently determined by changes in relative output elasticity and changes in overall output elasticities, while degree of segregation is affected by both changes. As a result, correlation of within-firm wage rates is likely to under-estimate the degree of segregation. Other indicators of ability such as education level or job experience should be used instead.

Technological changes in the relative elasticities of output to skills of manager and assistants may have occurred too. One possibility is that the recent technological advances in communications have made and "flattened" organization structures. As a result, jobs for bosses and for assistants have become more similar. In our model, this translates into smaller differences between the elasticities α and β, and reinforces segregation by skill due to increasing complementarity. On the other hand, in the spirit of the argument advanced by some economists based on skill-biased technological changes, some may argue that output has become relatively more sensitive to the skill of managers, while the total elasticity has not changed. If so, the present analysis predicts decreasing segregation instead of increasing segregation!

There is some evidence, cited by Kremer and Maskin (1996), which suggests that decreasing wage inequality within firm as well as increasing segregation across firms. To accommodate both evidence, changes in relative output elasticity must be considered together with changes in overall output elasticities. Our model is parsimonious in that two underlying variables, relative output elasticity to the skill of manager and overall output elasticities of skills of manager and assistant, determine three potentially quantifiable variables, across-firm wage inequality, within-firm wage inequality, and degree of segregation of workers by skill. Empirical testing of our model can be conducted by constructing indices of across-firm wage inequality, within-firm inequality and degree of segregation, and checking the trends of the indices against the predictions of the model.

5. Conclusion

This paper uses a one-sided assignment model, with manager and assistant tasks, to explain observed trends of increasing segregation of workers by skill in the workplace and increasing wage inequality. We are not able to completely characterize the equilibrium of the assignment model, but we show how it depends critically on the trade-off faced by marginal workers who are indifferent between being a manager of a lower skill worker and an assistant to higher skill worker. Our analysis suggests that the driving force behind the two trends mentioned above is technological changes that make output of a firm more sensitive to skill of both manager and assistant, and therefore increase the relative importance of complementarity between the managerial and assisting tasks.

Our model has the virtue of parsimony. There are only two underlying variables: relative output elasticity to the skill of manager and overall output elasticities of skills of manager and assistant. Across-firm wage inequality is affected only by changes in overall output elasticities while within-firm wage inequality is affected only by changes in relative elasticity. The degree of segregation of workers by skill is positively affected by the relative output elasticity and negatively by the overall output elasticity. With two underlying variables, empirical testing of our model can be conducted by constructing indices of across-firm wage inequality, within-firm inequality and degree of segregation, and checking the trends of the indices against the predictions of the model. As an illustration of potential applications of our framework to empirical investigations, consider Kremer and Maskin's observation that "economic activity has shifted from firms such as General Motors, which use both high-and low-skill workers, to firms such as Microsoft and McDonald's, whose workforces are much more homogeneous." While such stylized observation is useful in focusing the issue, changes in the degree of segregation by skill are not uniform across industries. One testable implication of our explanation of increasing segregation is that industries with quickest pace of technological change in terms of increasing output elasticity to skill experience most dramatic increase in the degree of labor market segregation. We plan to investigate this implication in future works.

References

Gary Becker (1981): *A Treatise on Family*, Harvard University Press.

Chinhui Juhn, Kevin Murphy, and Brooks Pierce (1993): "Wage Inequality and the Rise in Returns to Skill," *Journal of Political Economy* 101(3), 410-442.

Timothy Dunne, Lucia Foster, John Haltiwanger and Kenneth Troske (2000): "Wage and Productivity Dispersion in U.S. Manufacturing: The Role of Computer Investment," NBER Working Paper Series 7465.

James Kahn and Jong-Soo Lim (1994): "On the Contribution of Human Capital to Growth," University of Rochester Working Paper.

Lawrence F. Katz and Kevin M. Murphy (1992): "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, 35-78.

Tjalling Koopmans and Martin Beckmann (1953): "Assignment Problems and the Location of Economic Activities," *Econometrica*, 53-76.

Francis Kramarz, Stephen Lollivier, and Louis-Paul Pele (1996): "Wage Inequalities and Firm-Specific Compensation Policies in France," *Annales D'Economie et de Statistique, Institute National de la Statistique et des Etudes Economiques*, No. 41/42, 369-386.

Michael Kremer (1993): "The O-Ring Theory of Economic Development," *The Quarterly Journal of Economics*, 551-575.

Michael Kremer (1997): "How Much Does Sorting Increase Inequality?" *Quarterly Journal of Economics* 112(1), 115-139.

Michael Kremer and Eric Maskin (1996): "Wage Inequality and Segregation by Skill," NBER Working Paper Series W5718.

Patrick Legros and Andrew Newman (1998): "Monotone Matching in Perfect and Imperfect Worlds," University of Columbia Working Paper.

Lloyd Shapley and Martin Shubik (1972): "The Assignment Game I: The Core," *International Journal of Game Theory*, 111-130.