We will study the model in Costinot and Vogel (2010) but augmented with capital.

There are a continuum of individuals each endowed with human capital $h \in \mathcal{H}$. Human capital is used to produced task outputs. Without loss of generality, assume that the mass of individuals is 1, with associated distribution function $\mu(h)$. Goods are produced in firms consisting of a continuum of workers working in tasks $j \in \mathcal{J} = [0, J]$. Each task requires both physical and human capital. Integrating over the output of all firms yields total output. Specifically, firm i produces

$$y_i = \left[\int_{j=0}^{J} \nu(j)^{\frac{1}{\sigma}} \tau_i(j)^{\frac{\sigma-1}{\sigma}} dj \right]^{\frac{\sigma}{\sigma-1}}, \tag{1a}$$

$$\tau_i(j) = M(j) k_i(j)^{\alpha} \left[\int_{\mathrm{h}_i(j)} b(h,j) d\mu \right]^{1-\alpha}, \tag{1b}$$

with $\nu(j) \in (0,1)$ for all task j, and $\int_j \nu(j) dj = 1$. The $\tau_i(j)$ is the amount of task j output produced by team i. In team i, task j is allocated $k_i(j)$ units of capital and a set of workers $\mathrm{h}_i(j)$. The function $b(h,j)$ is the contribution of a worker with human capital h assigned to task j.

ASSUMPTION 1: LOG-SUPERMODULARITY *The function $b : \mathcal{H} \times \mathcal{J} \mapsto \mathbb{R}^+$ is strictly positive and twice-differentiable, and is log-supermodular. That is, for all $h' > h$ and $j' > j$:*

$$\log b(h', j') + \log b(h, j) > \log b(h', j) + \log b(h, j'). \tag{2}$$

Assumption 1 ensures that high h-workers sort into high j-tasks in any equilibrium. Integrating $b(h,j)$ over h of the workers in the set $\mathrm{h}_i(j)$ yields the total amount of human capital input used in task j in team i.

The substitutability between tasks is captured by the elasticity parameter σ. The $M(j)$'s capture task-specific TFP's. Total output is simply

$$Y = \int y_i di. \tag{3}$$

1. Planner's Problem

We will assume competitive, complete markets and solve a static planner's problem. A planner allocates aggregate capital K and all individuals into tasks $j \in [0, J]$. Formally, define $l_i(h,j)$ as the number of individuals with human capital h that the planner assigns to task j in firm i. Then the planner's problem is to choose a capital allocation rule $k_i(j)$ and assignment rules $l_i(h,j)$ to maximize current output (3) subject to (1) s.t. for all $j \in \mathcal{J}$,

$$K = \int_i \int_j k_i(j) dj di, \quad H(j) \equiv \int_h b(h,j) l_i(h,j) dh = \int_i \left[\int_{\mathrm{h}_i(j)} b(h,j) d\mu \right] di,$$
$$d\mu = \left[\int_i \int_j l_i(h,j) dj di \right] dh \tag{4}$$

where K is the supply of capital (fixed) and $H(j)$ the total productivity of workers allocated to task j. For existence of a solution, we need to assume that

Assumption 2 *There exists a strictly positive mass of jobs such that $b(0,j) > 0$ and individuals such that $b(h,J) > 0$.*

and for uniqueness:

Assumption 3 *The domain of skills $\mathcal{H} = [0, h_M]$, where $h_M < \infty$ is the upperbound of skill h. The measure $\mu(h)$ is differentiable and $d\mu(h) > 0$ is continuous on $\mathcal{H}$.*

Assumption 3 implies that we can write

$$\mu(\tilde{h}) = \int^{\tilde{h}} dG(h) = \int^{\tilde{h}} g(h) dh,$$

where (G, g), the cumulative and probability density functions of h, are continuous.

The optimal factor allocation rules across firms i, $[k_i(j), l_i(h,j)]$, are straightforward: Since we assume a constant returns to scale technology, we can assume existence of a representative firm. So we can write the aggregate technology as

$$Y = \left[\int_j \nu(j)^{\frac{1}{\sigma}} T(j)^{\frac{\sigma-1}{\sigma}} dj \right]^{\frac{\sigma}{\sigma-1}}, \qquad T(j) = M(j) K(j)^{\alpha} H(j)^{1-\alpha},$$

where, with some abuse of notation, $K(j)$ is the total amount of capital allocated to task j. We characterize the solution in the following order:

1. Optimal physical capital allocations across tasks;

2. Optimal worker allocations across tasks;

which will also allow us to show existence and uniqueness of the solution.

1.1 Capital allocations

Given aggregate capital K, the planner equalizes marginal product across tasks:

$$\begin{aligned} & MPK(0) = MPK(j) \\ \Rightarrow \quad & \frac{MPT(0)\cdot \alpha T(0)}{K(0)} = \frac{MPT(j)\cdot \alpha T(j)}{K(j)} \\ \Rightarrow \quad & \frac{MPT(j)\cdot T(j)}{MPT(0)\cdot T(0)} = \frac{K(j)}{K(0)} \quad \equiv \quad \pi(j) = \underbrace{\left[\frac{\nu(j)}{\nu(0)}\right]^{\frac{1}{\sigma}}}_{\equiv v(j)} \cdot \left[\frac{T(j)}{T(0)}\right]^{\frac{\sigma-1}{\sigma}}, \end{aligned} \tag{5}$$

where $MPT(j)$ is the marginal product of $T(j)$ and $\pi(j)$ is the capital input ratio between tasks j and 0. Due to the Cobb-Douglas assumption, $\pi(0)$ divided by task output ratios is the marginal rate of technological substitution ($MRTS$) between tasks j and 0; furthermore, $\pi(j)$ divided by either factor input ratios in tasks j and 0 measures the $MRTS$ of that factor between tasks j and 0. (For capital, this is equal to 1.) Given (5) we can write

$$Y = \nu(0)^{\frac{1}{\sigma-1}} \underbrace{\left[\int_j \pi(j)dj\right]^{\frac{\sigma}{\sigma-1}}}_{\equiv\Pi} T(0). \tag{6}$$

1.2 Worker allocations

Since $b(h,j)$ is strictly log-supermodular, Assumptions 1-3 imply that there exists a continuous assignment function $\hat{j} : [0, h_M] \mapsto \mathcal{J}$ s.t. $\hat{j}'(h) > 0$, and $\hat{j}(0) = 0$, $\hat{j}(h_M) = J$.[1] That is, there is positive sorting of workers into tasks, and workers of skill j are assigned to job $\hat{j}(h)$. And since $\hat{j}'(h) > 0$ we can also define its inverse $\hat{h} : \mathcal{J} \mapsto [0, h_M]$.

It should also be clear that $[\hat{j}(h), \hat{h}(j)]$ are identical across firms, and hence are not subscripted by i: Otherwise, the planner would be able to reallocate h across firms and increase output. So we dropped the subscript on the allocation rule $l_i(h,j) = l(h,j)$ already, and can consider the planner's optimal choice. The feasibility constraint (4) and existence of $[\hat{h}(j), \hat{j}(h)]$ implies that the number of people with human capital h assigned to task j is

$$l(h,j)dh = \delta\left[j - \hat{j}(h)\right] \cdot d\mu$$

where $\delta(\cdot)$ is the Dirac delta function. Hence the allocation rule is completely determined by the assignment functions $[\hat{h}(j), \hat{j}(h)]$, and the productivity of all workers assigned to

[1] For a more formal proof, refer to Lemma 1 in Costinot and Vogel (2010).

task $j = \hat{j}(h)$ is

$$H(j) = \int b(h, \hat{j}(h')) \cdot \delta\left[j - \hat{j}(h')\right] dG(h')$$

and we can use the change of variables $j' = \hat{j}(h')$ to instead integrate over j':

$$H(j) = \int b(\hat{h}(j'), j') \cdot \delta\left[j - j'\right] \cdot g(\hat{h}(j')) \cdot \hat{h}'(j')dj' = b(\hat{h}(j), j) \cdot g(\hat{h}(j)) \cdot \hat{h}'(j) \tag{7}$$

At the optimal allocation, there is no marginal gain from switching any worker to another job. Therefore for any $j' = j + dj$,

$$\begin{aligned}\frac{MPT(j) \cdot T(j)}{H(j)} \cdot b(\hat{h}(j), j) \geq & \frac{MPT(j') \cdot T(j')}{H(j')} \cdot b(\hat{h}(j), j'), \\ \frac{MPT(j') \cdot T(j')}{H(j')} \cdot b(\hat{h}(j'), j') \geq & \frac{MPT(j) \cdot T(j)}{H(j)} \cdot b(\hat{h}(j'), j),\end{aligned}$$

with equality if $|dj| = 0$. Substituting for $H(j)$ using (7), we obtain

$$\frac{b(\hat{h}(j'), j')}{b(\hat{h}(j), j')} \geq \frac{\pi(j')}{\pi(j)} \cdot \frac{g(\hat{h}(j))\hat{h}'(j)}{g(\hat{h}(j'))\hat{h}'(j')} \geq \frac{b(\hat{h}(j'), j)}{b(\hat{h}(j), j)},$$

so we obtain, as $|dj| \to 0$,

$$\left[\partial \log b(\hat{h}(j), j)/\partial h\right] \cdot \hat{h}'(j) = d \log\left\{\pi(j) \Big/ \left[g(\hat{h}(j))\hat{h}'(j)\right]\right\} \Big/ dj.$$

Now using the total derivative of $b(\hat{h}(j), j)$:

$$d \log b(\hat{h}(j), j)/dj = \left[\partial \log b(\hat{h}(j), j)/\partial h\right] \cdot \hat{h}'(j) + \partial \log b(\hat{h}(j), j)/\partial j. \tag{8}$$

Without loss of generality, normalize $b(0, 0) = 1$ and apply $\pi(0) = 1$ to obtain

$$H(j)/\pi(j)H(0) = \exp\left[\int_0^j \frac{\partial \log b(\hat{h}(j'), j')}{\partial j'} dj'\right] \equiv B_j(j; \hat{h}) \tag{9}$$

where B_j is a functional of j and the function $\hat{h}$. Applying this to (5) we obtain

$$\pi(j) = v(j) \Big/ \left[\tilde{M}(j) B_j(j; \hat{h})^{1-\alpha}\right]^{1-\sigma}. \tag{10}$$

where $\tilde{M}(j) \equiv M(j)/M(0)$. Note that (8) also implies that

$$b(h, \hat{j}(h))/B_j(\hat{j}(h); \hat{h}) = B_h(h; \hat{j}) \equiv \exp\left[\int_0^h \frac{\partial \log b(h', \hat{j}(h'))}{\partial h'} dh'\right]. \quad (11)$$

In what follows, we suppress the dependence of (B_j, B_h) on $(\hat{h}, \hat{j})$ unless necessary.

2. Solution and Equilibrium

The optimal allocation is thus described by the function $\hat{h}(j)$ that solves

$$\hat{h}'(j) = H(0) \cdot v(j) \Big/ \left\{ [\tilde{M}(j) B_j(j)^{1-\alpha}]^{1-\sigma} B_h(\hat{h}(j)) g(\hat{h}(j)) \right\} \quad (12)$$

along with the boundary condition $\hat{h}(0) = 0$ and $\hat{h}(J) = h_M$, which implies

$$H(0) = h_M \Big/ \int \left\{ v(j) \Big/ [\tilde{M}(j) B_j(j)^{1-\alpha}]^{1-\sigma} B_h(\hat{h}(j)) g(\hat{h}(j)) \right\} dj$$

The ODE (12) is separable as $C(\hat{h})d\hat{h} = D(j)dj$, which is easily solvable and where the functions (C, D) are represented by the integrands in

$$\Pi = \int B_h(h) g(h) dh \Big/ H(0) = \int \left\{ v(j) \Big/ \left[\tilde{M}(j) B_j(j)^{1-\alpha} \right]^{1-\sigma} \right\} dj.$$

The functions $(B_j(j), B_h(h))$, which represent relative wages in equilibrium, are defined in (9) and (11). Note that these objects are determined independently of the amount of physical capital, and also the total amount of labor—all that matters are relative masses *across tasks*. To see this more clearly, define $\psi \equiv \nu(0)^{\frac{1}{\sigma-1}}$ and rewrite (6) to obtain

$$Y_i = \psi \cdot \Pi^{\frac{\sigma}{\sigma-1}} M(0) K(0)^{\alpha} H(0)^{1-\alpha}. \quad (13)$$

Using (5), total capital can be written as

$$K = K(0) \cdot \Pi, \quad (14)$$

and using (7) and (9)-(10) we obtain

$$\int \left[H(j) \Big/ b(\hat{h}(j), j) \right] dj = \int g(h) dh = H(0) \cdot \int \left[\pi(j) / B_h(\hat{h}(j)) \right] dj = L$$

where L is the total mass of individuals. Rearranging we obtain

$$L = H(0) \cdot \Pi_l \quad \text{where} \quad \Pi_l \equiv \int \left[\pi(j) / B_h(\hat{h}(j)) \right] dj, \tag{15}$$

and (13) becomes

$$Y = \underbrace{\psi M(0) \cdot \Pi^{\frac{\sigma}{\sigma-1} - \alpha} \Pi_l^{\alpha-1}}_{\Phi\text{: TFP}} K^{\alpha} L^{1-\alpha}. \tag{16}$$

Hence, we recover an aggregate production function in which TFP is completely separated from capital and labor. TFP Φ can be decomposed into 2 parts: $\psi M(0)$, which plays the traditional role of exogenous TFP; and the part determined by (Π, Π_l), which is endogenously determined by the allocation rule $\hat{h}(j)$.

2.1 Equilibrium

Since there are no frictions in this economy, a solution to the planner's problem coincides with an equilibrium, so existence and uniqueness of an equilibrium is equivalent to a unique solution to the planner's problem. The assumptions that we have made so far already satisfy this:

PROPOSITION 1 *Under Assumptions 1-3, the solution to the planner's problem, $[\hat{h}(j)]_{j=0}^{J}$, exists and is unique.*

Proof: Under Assumptions 1-3, existence of a solution to the differential equation (12) is a simple application of Picard-Lindelöf's existence theorem. □

The planner's solution gives all the information needed to derive equilibrium prices (which are unique). Normalize the price of the final good to 1. Let R denote the rental rate of capital and $w_h(h)$ the wage of a worker with skill h. The rental rate R can either be given by the dynamic law of motion for aggregate capital, or fixed in a small open economy. The wage in the lowest job $j = 0$, $w_h(0)$, can be found from

$$w_h(0) = \frac{1-\alpha}{\alpha} \cdot \frac{RK(0)}{H(0) \cdot b(0,0)} = \frac{1-\alpha}{\alpha} \cdot \frac{\Pi_l}{\Pi} \cdot RK,$$

where the second equality follows from (14)-(15) and normalizing $b(0,0)$ and the population to 1.

Similarly, all workers earn their marginal product, so for any two workers with skills

$h \neq h'$ we can write

$$\log w_h(h') - \log w_h(h) = \log \left[\frac{MPT(\hat{j}(h')) \cdot T(\hat{j}(h'))}{MPT(\hat{j}(h)) \cdot T(\hat{j}(h))} \cdot \frac{H(\hat{j}(h))}{H(\hat{j}(h'))} \cdot \frac{b(h', \hat{j}(h'))}{b(h, \hat{j}(h))} \right]$$
$$= \log \left[\frac{\pi(\hat{j}(h'))}{\pi(\hat{j}(h))} \cdot \frac{H(\hat{j}(h))}{H(\hat{j}(h'))} \cdot \frac{b(h', \hat{j}(h'))}{b(h, \hat{j}(h))} \right]$$
$$= \log \left[B_h(\hat{j}(h')) / B_h(\hat{j}(h)) \right],$$

where the second equality follows from (5) and the third from (9) and (11). So letting $h' = h + dh$ and sending $|dh| \to 0$, we obtain

$$w_h(h) = w_h(0) \cdot B_h(h), \tag{17}$$

so if we assume that $\partial b(h, j)/\partial h > 0$, $w(h)$ is strictly increasing in h. Note that this is *not* implied by log-supermodularity alone.

2.2 Wage and Job Polarization

Now consider growth an increase in $M(j)$ for all $j \in \mathcal{J}_1 \equiv [\underline{j}, \bar{j}]$, where $0 < \underline{j} < \bar{j} < J$; this are the middle-skill jobs, due the existence of $\hat{h}(j)$ (positive sorting). Since $M(j)$'s are separate from workers' human capital h, this can be interpreted as a rise in capital-augmented TFP for middle-skill jobs, or *routinization*. This captures the notion that machines have become the most productive in such jobs, which used to fall in the middle of the job wage distribution. This is the same exercise as Lemma 6 in (Costinot and Vogel, 2010). It can also be viewed as the theory behind (Goos et al., 2014), who do model capital: They have a discrete number of tasks, assume that labor is task-specific, and vary wages exogenously according to the data.

Proposition 2: Routinization and Polarization *Let $\mathcal{J}^1 \equiv (\underline{j}, \bar{j}) \subset \mathcal{J}$, where $0 < \underline{j} < \bar{j} < J$. Let $M(j)$ uniformly grow to $M^1(j) = M(j)e^m$ for all $j \in \mathcal{J}^1$, where $m > 0$. Then under Assumptions 1-3, and if $\sigma < 1$, there exists $j^* \in \mathcal{J}^1$ such that $\hat{h}^1(j) > \hat{h}(j)$ for all $j \in (0, j^*)$ and $\hat{h}^1(j) < \hat{h}(j)$ for all $j \in (j^*, J)$.*

To prove the proposition, we first prove the following Lemma:

Lemma 1 *Suppose $[\hat{h}(j), \hat{h}^1(j)]$ are both an equilibrium. For any connected subset $\mathcal{J}^1 \subseteq \mathcal{J}$, $\hat{h}$ and $\hat{h}^1$ can never coincide more than once on $\mathcal{J}^1$.*

Proof: Let **(i)** $\hat{h}(j_a) = \hat{h}^1(j_a)$ and $\hat{h}(j_b) = \hat{h}^1(j_b)$ such that both $(j_a, j_b) \in \mathcal{J}^1$. Without loss of generality, we assume that $j_a < j_b$ are two adjacent crossing points. Then, since $[\hat{h}, \hat{h}^1]$ are Lipschitz continuous and strictly monotone in j, it must be the case that

1. (ii) $\hat{h}^{1\prime}(j_a) \geq \hat{h}'(j_a)$ and $\hat{h}^{1\prime}(j_b) \leq \hat{h}'(j_b)$; and **(iii)** $\hat{h}^1(j) > \hat{h}(j)$ for all $j \in (j_a, j_b)$; or

2. (ii) $\hat{h}^{1\prime}(j_a) \leq \hat{h}'(j_a)$ and $\hat{h}^{1\prime}(j_b) \geq \hat{h}'(j_b)$; and **(iii)** $\hat{h}^1(j) < \hat{h}(j)$ for all $j \in (j_a, j_b)$.

Consider case **1**. Condition **(ii)** implies

$$\hat{h}^{1\prime}(j_b)/\hat{h}^{1\prime}(j_a) \leq \hat{h}'(j_b)/\hat{h}'(j_a)$$

and using (8)-(9) and (12), and applying $\hat{h}^1(j) = \hat{h}(j)$ for $j \in \{j_a, j_b\}$ we obtain

$$[\alpha + \sigma(1-\alpha)] \cdot \left[\int_{j_a}^{j_b} \frac{\partial \log b(\hat{h}^1(j'), j')}{\partial j'} dj' - \int_{j_a}^{j_b} \frac{\partial \log b(\hat{h}(j'), j')}{\partial j'} dj' \right] \leq 0$$

but under **(iii)**, the log-supermodularity of b in (2) implies

$$\partial \log b(h^1, j)/\partial j > \partial \log b(h, j)/\partial j \qquad \forall h^1 > h, \tag{18}$$

a contradiction. Case **2** is symmetric. □

Proof of Proposition 2 By Lemma 1, we know that no crossing can occur on $(0, \underline{j})$ or $(\bar{j}, J)$, since $\hat{h}$ and $\hat{h}^1$ already coincide at the boundaries 0 and J. It also implies that it can never be the case that there is no crossing ($\hat{h}^1(j) > \hat{h}(j)$ or $\hat{h}^1(j) < \hat{h}(j)$ for all $j \in \mathcal{J} \setminus \{0, J\}$). Hence, there must be a single crossing in $\mathcal{J}^1$.

At this point, the only possibility for j^* not to exist is if instead, there exists a single crossing j^{**} such that **(i)** $\hat{h}^1(j) < \hat{h}(j)$ for all $j \in (0, j^{**})$ and **(ii)** $\hat{h}^1(j) > \hat{h}(j)$ for all $j \in (j^{**}, J)$. If so, since $[\hat{h}, \hat{h}^1]$ are Lipschitz continuous and strictly monotone in j, it must be the case that $\hat{h}^{1\prime}(0) < \hat{h}'(0)$, $\hat{h}^{1\prime}(j^{**}) > \hat{h}'(j^{**})$ and $\hat{h}^{1\prime}(J) < \hat{h}'(J)$. This implies

$$\hat{h}^{1\prime}(j^{**})/\hat{h}^{1\prime}(0) \geq \hat{h}'(j^{**})/\hat{h}'(0), \qquad \hat{h}^{1\prime}(J)/\hat{h}^{1\prime}(j^{**}) \leq \hat{h}'(J)/\hat{h}'(j^{**}). \tag{19}$$

Let us focus on the first inequality. Using (9) and (12) we obtain

$$0 > [\alpha + \sigma(1-\alpha)] \cdot \left[\int_0^{j^{**}} \frac{\partial \log b(\hat{h}^1(j), j)}{\partial j} dj - \int_0^{j^{**}} \frac{\partial \log b(\hat{h}(j), j)}{\partial j} dj \right] \geq \ (1-\sigma)m.$$

where the first inequality follows from (18), and applying **(i)**. Since $m > 0$, if $\sigma \in (0,1)$ we have a contradiction. The case for the second inequality in (19) is symmetric. □

Hence, capital and labor flow out from middle jobs into the extremes (job polarization), and relative wages also decline for middle jobs (wage polarization), since from (17)

$$\log\left[\frac{w^1(h)}{w^1(h^*)}\right] - \log\left[\frac{w(h)}{w(h^*)}\right] = \int_{h^*}^{h}\left[\frac{\partial \log b(h, \hat{j}^1(h))}{\partial h} - \frac{\partial \log b(h, \hat{j}(h))}{\partial h}\right] dh$$

is negative for all $j \neq j^* \in \mathcal{J}^1$.

When $\sigma < 1$, the exogenous rise in productivity causes factors to flow out to other tasks since tasks are complementary, and we get employment polarization. This also leads to wage polarization in the presence of positive sorting.

Changes in the skill distribution can have different effects as shown in Costinot and Vogel (2010), which I leave up to you.

References

Costinot, A. and J. Vogel (2010, 08). Matching and Inequality in the World Economy. *Journal of Political Economy 118*(4), 747–786.

Goos, M., A. Manning, and A. Salomons (2014). Explaining job polarization: Routine-biased technological change and offshoring. *American Economic Review 104*(8), 2509–26.