Online Appendix for "Learning and Knowledge Diffusion in a Global Economy"

Micro-foundation for the learning technology : Suppose workers have to exert effort in order to learn. If a worker, working for a manager with knowledge y, exerts e, his knowledge next period is $l(e, y)$. For simplicity, I assume that $l(e, y)$ has the following functional form - $\xi(y)e$, with $\xi' > 0$. So all workers, working for a particular manager, can potentially learn the same. Effort, however, is costly. Let us denote this cost by $q_x(e)$ with $q_x' > 0$. Note that the cost of effort depends on the knowledge of the worker. I assume that $q_x'(e)$ is decreasing in x. Let the continuation value of a worker with knowledge $l(e, y)$ in the next period be $C(l(e, y))$. If the Separation Theorem holds, the worker will choose e to maximize $C(l(e, y)) - q_x(e)$. The first-order condition for utility maximization yields $\frac{dC}{dl}\frac{dl}{de} = \frac{dC}{dl}\xi(y) = \frac{dq_x}{de}$. The left-hand side of this equation is constant. Given the assumption about q_x, it then follows that more knowledgeable workers will choose higher e and accordingly have higher knowledge next period.

Proof of Proposition 1 : We prove this proposition in a number of steps. In this economy, every agent with knowledge y, in the role of a manager, offers a "gross" wage schedule $\widetilde{w}_t(x, y)$ such that y *is indifferent between hiring any* x. $\widetilde{w}_t(x, y)$ is the wage offered by y if there were no learning and captures the worker's pay-off from production. The following lemma establishes some properties of $\widetilde{w}_t(x, y)$.

Lemma 1. *$\widetilde{w}_t(x, y)$ is increasing in x for all y, and* $\frac{\partial^2 \widetilde{w}}{\partial x \partial y} > 0$.

Proof. Since y is indifferent along $\widetilde{w}_t(x, y)$, for any x_1 and x_2 we must have $f(y)n(x_1) - \widetilde{w}_t(x_1, y)n(x_1) = f(y)n(x_2) - \widetilde{w}_t(x_2, y)n(x_2)$. Letting $x_2 = x_1 + h$ and re-arranging, we have $\widetilde{w}_t(x_1 + h, y)n(x_1 + h) - \widetilde{w}_t(x_1)n(x_1) = f(y)n(x_1 + h) - f(y)n(x_1)$. Using Taylor series approximation of $n(x_1+h)$ around h (small), we have $n(x_1+h) = n(x_1)+n'(x_1)h+o_2$. Replacing this in the above equation, we have $[\widetilde{w}_t(x_1 + h, y) - \widetilde{w}_t(x_1, y)]n(x_1) + \widetilde{w}_t(x_1 + h, y)n'(x_1)h = f(y)n(x_1+h)-f(y)n(x_1)$. Dividing by h and taking the limit as $h \to 0$, we get $\frac{\partial \widetilde{w}_t(x_1,y)}{\partial x}n(x_1)+ \widetilde{w}_t(x_1, y)n'(x_1) = f(y)n'(x_1)$. Re-arranging, we have $\frac{\partial \widetilde{w}_t(x_1,y)}{\partial x} = \frac{n'(x_1)(f(y)-\widetilde{w}_t(x_1,y))}{n(x_1)}$. Since $f(y) - \widetilde{w}_t(x_1, y) > 0$ (in equilibrium, profits must be positive) and $n'(x_1) > 0$, it follows that $\frac{\partial \widetilde{w}_t(x_1,y)}{\partial x} > 0$. Since x_1 was chosen randomly, the result follows. Furthermore, in equilibrium, $\frac{\partial \widetilde{w}_t(x,y)}{\partial x} = \frac{n'(x)(f(y)-\widetilde{w}_t(x))}{n(x)}$. Differentiating with respect to y, $\frac{\partial^2 \widetilde{w}_t(x,y)}{\partial x \partial y} = \frac{n'(x)f'(y)}{n(x)} > 0$. □

At the same time, each agent, in the role of a worker, offers a rent schedule $\widetilde{r}_t(x,y)$ such that he is indifferent across managers; $\widetilde{r}_t(x,y)$ captures the value of learning. Unlike $\widetilde{w}_t(x,y)$, which is the solution to a static problem, workers compute $\widetilde{r}_t(x,y)$ by taking into account the entire expected earnings profile. The following lemma establishes a key property of the rent function.

Lemma 2. $\frac{\partial^2 \widetilde{r}}{\partial x \partial y} \geq 0$.

Proof. The worker's optimization problem is given by

$$V_W(x) = \max_y [\widetilde{w}_t(x,y) - \widetilde{r}_t(x,y) + \int V_W(x') dL(x'|x,y)].$$

The first-order condition is given by $\frac{\partial \widetilde{r}_t(x,y)}{\partial y} = \frac{\partial \widetilde{w}_t(x,y)}{\partial y} + \beta \frac{\partial}{\partial y} \int V_W(x') dL(x'|x,y)$. Differentiating with respect to x, $\frac{\partial^2 \widetilde{r}_t(x,y)}{\partial x \partial y} = \frac{\partial^2 \widetilde{w}_t(x,y)}{\partial x \partial y} + \beta \frac{\partial^2}{\partial x \partial y} \int V_W(x') dL(x'|x,y)$. The first term on the left-hand side is positive (Lemma 1). The second term is also positive because of Assumption 2c and $V_w(x')$ being increasing in x'. □

Lemmas 1 and 2 allow as to prove that the equilibrium exhibits positive assortative matching, as shown in the following lemma.

Lemma 3. *$m_t(x)$ is invertible and strictly increasing in x.*

Proof. I shall prove this lemma by contradiction. The subscript t is dropped for simplicity. To prove that $m(x)$ is not a correspondence (which would rule out invertibility), first note that measure consistency implies that $m(x)$ is not an interval. So, we could have a x_2, y_1 and y_3 ($y_1 < y_3$) such that $m(x_2) = y_1$ and $m(x_2) = y_3$. But then, we can always find a $y \in [y_1, y_3]$ such that $m(x_2) \neq y$. Let us denote this by y_1. Let us also assume, without loss of generality, that $m(x_2) = y_1$ where $x_1 < x_2$. To summarize, there must be $x_1 < x_2$ and $y_1 < y_2$ such that $m(x_1) = y_2$ and $m(x_2) = y_1$. Note that this also implies that we do not have positive assortative matching. For this allocation to be an equilibrium, we must have

$$\begin{aligned} \pi(y_2) &= [f(y_2) - \widetilde{w}(x_1,y_2) + \widetilde{r}(x_1,y_2)]n(x_1) \\ &\geq [f(y_2) - \widetilde{w}(x_2,y_2) + \widetilde{r}(x_2,y_2)]n(x_2) \end{aligned}$$

Similarly, we must have

$$\begin{aligned} \pi(y_1) &= [f(y_1) - \widetilde{w}(x_2,y_1) + \widetilde{r}(x_2,y_1)]n(x_2) \\ &\geq [f(y_1) - \widetilde{w}(x_1,y_1) + \widetilde{r}(x_1,y_1)]n(x_1) \end{aligned}$$

Combining the above two inequalities and using the fact that $n(x_1) < n(x_2)$, we can write

$$[\widetilde{w}(x_2, y_2) - \widetilde{w}(x_1, y_2)] - [\widetilde{w}(x_2, y_1) - \widetilde{w}(x_1, y_1)]$$
$$> [\widetilde{r}(x_2, y_2) - \widetilde{r}(x_1, y_1)] - [\widetilde{w}(x_1, y_2) - \widetilde{w}(x_1, y_1)]$$

Defining $x_2 = x_1 + \Delta$ and $y_2 = y_1 + \Delta$ and taking the limit as $\Delta \to 0$, we have $\frac{\partial^2 \widetilde{w}(w,y)}{\partial x \partial y} > \frac{\partial^2 \widetilde{r}(w,y)}{\partial x \partial y}$. But from Lemma 2, we know that $\frac{\partial^2 \widetilde{w}(w,y)}{\partial x \partial y} < \frac{\partial^2 \widetilde{r}(w,y)}{\partial x \partial y}$. Hence, we get a contradiction. □

To prove that an allocation where agents with $k < k_t^*$ are workers and those with $k > k_t^*$ are managers, we first need to show that such a k_t^*, for which the labor market clears, actually exists. This is shown in the next lemma.

Lemma 4. *For a given $\Psi_t(k)$, k_t^* exists and is unique.*

Proof. Equilibrium in the labor market implies that

$$\int_{\underline{k}}^{k_t^*} \psi(s)ds = \int_{k_t^*}^{\overline{k}} n(m_t^{-1}(s))\psi(s)ds$$

where the LHS is the supply of workers while the RHS is the demand for workers. Define

$$\mathcal{L}(k_t^*) = \int_{\underline{k}}^{k_t^*} \psi(s)ds - \int_{k_t^*}^{\overline{k}} n(m_t^{-1}(s))\psi(s)ds$$

Now, $\mathcal{L}(\underline{k}) = -\int_{\underline{k}}^{\overline{k}} n(m_t^{-1}(s))\psi(s)ds < 0$, while $\mathcal{L}(\overline{k}) = \int_{\underline{k}}^{\overline{k}} \psi(s)ds > 0$. Moreover, $\frac{\partial L(k_t^*)}{\partial k_t^*} = [1 + n(m_t^{-1}(k_t^*)]\psi(k_t^*) > 0$. Hence, by the Intermediate Value Theorem, $\exists$ a unique k_t^* such that $\mathcal{L}(k_t^*) = 0$. □

The last step, before we prove the existence of the equilibrium, is to derive some useful properties of the value functions, $V_W(k)$ and $V_M(k)$.

Lemma 5. *$V_W(k)$ and $V_M(k)$ exist, are continuous and increasing in k.*

Proof. The value function for the manager is given by

$$V_M(k, w_t) = \max_x \{f(k)n(x) - w_t(x)n(x)\} + (1 - \delta) \max[V_W(k, w_{t+1}), V_M(k, w_{t+1})]$$

The value function for the worker is given by

$$V_W(k, w_t) = w_t(k) + (1-\delta)\int_k^{m_t(k)} \max[V_W(k', w_{t+1}), V_M(k', w_{t+1})]dL(k'|k, m_t(k))$$

Define the vector function $V = [V_M(k, w_t)\ V_W(k, w_t)]'$. Then $\max\{V_W, V_M\} = \max\{[1\ 0]V, [0\ 1]V\}$. Also, define $\alpha = [\max_x\{f(k)n(x) - w_t(x)n(x)\}\ w_t(k)]'$. Then we have the following equation:

$$\begin{aligned} V &= \alpha + (1-\delta)\begin{bmatrix} \max\{[1\ 0]V, [0\ 1]V\} \\ \int_k^{m_t(k)} \max\{[1\ 0]V, [0\ 1]V\}dL \end{bmatrix} \\ &= T(V) \end{aligned}$$

It can be established, using Blackwell's Sufficiency Conditions, that the operator T is a contraction in the space of continuous vector functions with norm $max[\sup_k |V_M(k)|, \sup_k |V_W(k)|]$. Therefore, a fixed point of V exists and is unique.

To prove that $V_W(k)$ is increasing in k, note that if $V_W'(k) < 0$, a worker will choose not to learn because learning reduces his continuation value. If workers do not learn, they do not pay rent. Consequently, the only payment that is made is wage and we are back to the static framework. But then more knowledgeable agents earn more and $V_W'(k) > 0$ trivially. Thus, we get a contradiction. It can be proved in a similar fashion that $V_M'(k) > 0$. □

Given the above lemmas, let us derive the equilibrium conditions for a threshold equilibrium. Since k_t^* is indifferent between being a worker and a manager, we must have $V_W(k_t^*, w_t) = V_M(k_t^*, w_t)$. Furthermore, for k_t^* to be the threshold, it must be the case that $\overline{k}$ cannot hire $k_t^* + \epsilon$ and be strictly better-off. If $k_t^* + \epsilon$ is a manager, he earns $V_M(k_t^* + \epsilon)$. In order to hire $k_t^* + \epsilon$, the manager has to pay him a wage such that he is just indifferent between being a manager and a worker. Let this wage be ω. ω should satisfy

$$\omega + (1-\delta)\int V_M(k)dL(k|, k_t^* + \epsilon, \overline{k}) = V_M(k_t^* + \epsilon)$$

Therefore, period profit of $\overline{k}$ if he hires $k_t^* + \epsilon$ is given by

$$\begin{aligned} \pi_{k_t^*+\epsilon}(\overline{k}) &= (f(\overline{k}) - \omega)n(k_t^* + \epsilon) \\ &= f(\overline{k})n(k_t^* + \epsilon) - n(k_t^* + \epsilon)(V_M(k_t^* + \epsilon) - (1-\delta)\int V_M(k)dL(k|, k_t^* + \epsilon, \overline{k})) \end{aligned}$$

For k_t^* to be a threshold equilibrium, it must be the case that $\underset{\epsilon\to 0}{lim}\frac{\partial \pi_{k_t^*+\epsilon}(\overline{k})}{\partial \epsilon} \leq 0$ Now,

$$\lim_{\epsilon\to 0}\frac{\partial \pi_{k_t^*+\epsilon}(\overline{k})}{\partial \epsilon} = f(\overline{k})n'(k_t^*) - n(k_t^*)(V_M'(k_t^*) - (1-\delta)\frac{\partial}{\partial k_t^*}\int V_M(k)dL(k|, k_t^*, \overline{k}))$$
$$- n'(k_t^*)V_M(k_t^*) - (1-\delta)\int V_M(k)dL(k|, k^*, \overline{k})$$

From the manager's profit-maximizing problem, we have

$$f(\overline{k})n'(k_t^*) = w_t'(k_t^*)n(k_t^*) + w_t(k_t^*)n'(k_t^*)$$

Also, for a worker with knowledge k_t^*,

$$V_W(k_t^*) = w_t(k_t^*) + (1-\delta)\int V_M(k)dL(k|k_t^*, \overline{k})$$

Since k_t^* is the threshold, $\max[V_W(k), V_M(k)] = V_M(k)\ \forall k \geq k_t^*$. Differentiating w.r.t. k_t^*,

$$V_W'(k_t^*) = w_t'(k_t^*) + (1-\delta)\frac{\partial}{\partial k_t^*}\int V_M(k)dL(k|k_t^*, \overline{k})$$

Replacing in the expression for $\lim_{\epsilon\to 0}\frac{\partial \pi_{k_t^*+\epsilon}(\overline{k})}{\partial \epsilon}$ and using the fact that $V_W(k_t^*) = V_M(k_t^*)$, we have

$$\lim_{\epsilon\to 0}\frac{\partial \pi_{k_t^*+\epsilon}(\overline{k})}{\partial \epsilon} = [w_t'(k_t^*) - V_M'(k_t^*) - V_W'(k_t^*) + w_t'(k_t^*)]n(k_t^*)$$
$$= V_W'(k_t^*) - V_M'(k_t^*)$$

where we use the fact that $V_W(k_t*, w_t) = V_M(k_t^*, w_t)$. Hence, $\underset{\epsilon\to 0}{lim}\frac{\partial \pi_{k_t^*+\epsilon}(\overline{k})}{\partial \epsilon} < 0$ implies that

$$V_W'(k_t^*) < V_M'(k_t^*)$$

The above condition needs to be satisfied for k_t^* to be the equilibrium threshold. We shall prove this proposition in a slightly different way. First, we shall prove the existence of the threshold equilibrium, assuming that the equilibrium is unique. Then we shall show that the sufficient condition for existence is also sufficient for uniqueness.

By assuming uniqueness, we are basically assuming that the set of workers and managers has to be connected in equilibrium. Given that there exists a unique market-clearing threshold k_t^*, we check whether the threshold satisfies the equilibrium condition $V_W'(k_t^*) < V_M'(k_t^*)$.

Dropping the time subscript, we have

$$V_M(k^*) = \frac{1}{\delta}(f(k^*) - w(\underline{k}))n(\underline{k})$$

Using the Envelope Theorem,

$$V'_M(k^*) = \frac{1}{\delta}f'(k^*)n(\underline{k})$$

Also,

$$\begin{aligned} V'_W(k^*) &= w'(k^*) + (1-\delta)\frac{\partial}{\partial k^*}(\int V_M(k)dL(k|k^*, \overline{k})) \\ &= \frac{(f(\overline{k}) - w(k^*))n'(k^*)}{n(k^*)} + (1-\delta)\frac{\partial}{\partial k^*}(\int V_M(k)dL(k|k^*, \overline{k})) \end{aligned}$$

where the second line follows from the manager's profit-maximization condition. Therefore, for k^* to be an equilibrium, it must be the case that

$$\frac{(f(\overline{k}) - w(k^*))n'(k^*)}{n(k^*)} + (1-\delta)\frac{\partial}{\partial k^*}(\int V_M(k)dL(k|k^*, \overline{k})) \leq \frac{1}{\delta}f'(k^*)n(\underline{k})$$

If $\delta = 1$, this condition reduces to

$$\frac{(f(\overline{k}) - w(k^*))n'(k^*)}{n(k^*)} \leq f'(k^*)n(\underline{k})$$

Since $w(k^*) > 0$, for the above inequality to hold, we need to find the conditions under which $\frac{f(\overline{k})n'(k^*)}{n(k^*)} \leq f'(k^*)n(\underline{k})$, or $f(\overline{k})n'(k^*) \leq f'(k^*)n(\underline{k})$, since $n(k^*) \geq 1$.

But $f(\overline{k})n'(k^*) \leq f(\overline{k})n'(\underline{k})$ ($\because n''(.) \leq 0$) and $f'(k^*)n(\underline{k}) \geq f'(\underline{k})n(\underline{k})$ (because $n''(.) \geq 0$). Hence, it follows that

$$f(\overline{k})n'(k^*) \leq f(\overline{k})n'(\underline{k}) \leq f'(\underline{k})n(\underline{k}) \leq f'(k^*)n(\underline{k})$$

where the inequality in the middle follows from Assumption 1b. Thus for $\delta = 1$, the condition on technology is sufficient for an equilibrium. But when $\delta \neq 1$, we need to determine the magnitude of $\frac{\partial}{\partial k^*}(\int V_M(k)dL(k|k^*, \overline{k}))$, since this term is positive by assumption on the learning technology. This term is endogenous and it depends on the invariant distribution, which in turn is determined by the learning distribution. This term is bounded above, since the domain is compact. Hence by the Least Upper Bound Property, the supremum exists. Let

$$\zeta = \sup \{ \frac{\partial}{\partial k^*} (\int V_M(k) dL(k|k^*, \overline{k})) \}$$

Define δ^* as the value of δ that satisfies

$$f(\overline{k})n'(\underline{k}) + (1-\delta^*)\zeta = f'(\underline{k})n(\underline{k})$$

This can be re-written as

$$\frac{n'(\underline{k})}{n(\underline{k})} + (1-\delta^*)\frac{\zeta}{f(\overline{k})n(\underline{k})} = \frac{f'(\underline{k})}{f(\overline{k})}$$

The fact that $\frac{n'(\underline{k})}{n(\underline{k})} < \frac{f'(\underline{k})}{f(\overline{k})}$ implies that $\delta^* < 1$. Hence $\forall \delta \in [\delta^*, 1]$, we have

$$f(\overline{k})n'(\underline{k}) + (1-\delta)\zeta \leq f'(\underline{k})n(\underline{k})$$

Thus,

$$\begin{aligned} f(\overline{k})n'(k^*) + (1-\delta)\frac{\partial}{\partial k^*}(\int V_M(k) dL(k|k^*, \overline{k})) &\leq f(\overline{k})n'(\underline{k}) + (1-\delta)\zeta \\ &\leq f'(\underline{k})n(\underline{k}) \\ &\leq f'(k^*)n(\underline{k}) \end{aligned}$$

$\delta \leq 1$ implies that $f'(k^*)n(\underline{k}) \leq \frac{1}{\delta} f'(k^*)n(\underline{k})$. Therefore,

$$f(\overline{k})n'(k^*) + (1-\delta)\frac{\partial}{\partial k^*}(\int V_M(k) dL(k|k^*, \overline{k})) \leq \frac{1}{\delta} f'(k^*)n(\underline{k})$$

This completes our proof about the existence of equilibrium. As mentioned before, showing uniqueness entails showing that the set of workers and managers is connected. Suppose not. In particular, let us assume that the knowledge distribution has the following partition - $([\underline{k}, k_1], [k_1, k_2], [k_2, k_3], [k_3, k_4])$. Workers in $[\underline{k}, k_1]$ work for managers in $[k_1, k_2]$ while workers in $[k_2, k_3]$ work for managers in $[k_3, k_4]$. For this to be an equilibrium, it must be the case that k_2 must be indifferent between being a worker and a manager. In other words, a deviation involving k_3 hiring $k_2 - \epsilon$ should not make both k_3 and $k_2 - \epsilon$ better off. Using a similar logic as developed above, one can show that that the condition for equilibrium is $V'_W(k_2) > V'_M(k_2)$. One can then show that if $\frac{n'(\underline{k})}{n(\underline{k})} < \frac{f'(\underline{k})}{f(\overline{k})}$, then for δ high enough, this condition will always be violated. Therefore, an allocation with disconnected sets of workers and managers can never be sustained as an equilibrium implying that the only equilibrium is the threshold equilibrium.

To prove the uniqueness and existence of the invariant distribution $\Phi(k)$, note there is an alternative way of looking at the evolution of knowledge. Let A be any Borel set of $[\underline{k}, \overline{k}]$. Then the transition function for the knowledge distribution satisfies, for every $k \in [\underline{k}, \overline{k}]$,

$$P_t(k, A) = \begin{cases} (1-\delta) \int\limits_A dL(s|k, m_t(k)) + \delta \int\limits_A d\Phi(s) & \text{if } k \in \overline{W} \\ \delta \int\limits_A d\Phi(s) & \text{if } k \in \overline{M} \end{cases}$$

Suppose P is monotone, has the Feller property and satisfies a mixing condition. Then P has a unique, invariant probability measure Ψ^* (Stokey, Lucas with Prescott, 1989). Define the operator T as

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

where $f : [\underline{k}, \overline{k}] \to \mathbb{R}$ is a bounded function. If f is non-decreasing, then the first-order stochastic dominance property of the learning distribution implies that Tf is also non-decreasing. (Monotone Property) It is straight-forward to verify that if f is bounded and continuous, then the same holds for Tf, i.e., $T : C(k) \to C(k)$ (Feller Property). The mixing condition requires that $\exists c \in [\underline{k}, \overline{k}]$, $\epsilon > 0$ and $N \geq 1$ such that $P^N(\underline{k}, [c, \overline{k}]) \geq \epsilon$ and $P^N([\underline{k}, c], \overline{k}]) \geq \epsilon$. Choose $k' \in [\underline{k}, \overline{k}]$. Define $\epsilon_1 = \int_{[k', \overline{k}]} d\Psi_N(s)$ and $\epsilon_2 = \int_{[\underline{k}, k']} d\Psi_N(s)$. By the assumption on $\Psi_N(.)$, we know that both these objects are greater than 0. Choose $\epsilon = \delta \min\{\epsilon_1, \epsilon_2\}$ and $N = 1$. Then $P(\underline{k}, [k', \overline{k}]) \geq \epsilon$ and $P([\underline{k}, k'], \overline{k}]) \geq \epsilon$. Therefore all the conditions for the existence and uniqueness of the invariant distribution are satisfied.