Online appendix to:
Experience vs. Obsolescence:
A Vintage-Human-Capital Model

Matthias Kredler[a]

9 August 2013

Abstract

This online appendix contains proofs and a discussion of the planner's problem for the paper "Experience vs. Obsolescence: A Vintage-Human-Capital Model".

Contents

[a]Departamento de Economía, Universidad Carlos III de Madrid, Calle Madrid 126, 28903 Getafe, Spain, matthias.kredler@uc3m.es. Part of the research was funded by the Spanish Ministerio de Ciencia e Innovación, reference number SEJ2007-62908, and by the Spanish Ministerio de Economía y Competitividad, reference number ECO2012-34581.

1 General setting

This subsection contains the proofs for the statements in Section 2.4 of the main paper, where no additional assumptions are imposed on the production function.

1.1 Proof of Lemma A.1: bounded resources

Proof. Let $\Delta = \{n : \int n = 1\}$ be the unit simplex. By weak concavity of $\tilde{Y}(\cdot)$, the set $B = \{(n, Y) : n \in \Delta, Y \leq \tilde{Y}(n)\}$ is convex and has non-empty interior. Now, fix some interior point $\bar{n} \in \Delta$, say $\bar{n}(\cdot) = 1$. By the separating-hyperplane theorem, there is a bounded linear functional f on Δ such that $Y(n) \leq f(n)$ for all $n \in B$; in other words, all points in B must be in the half-space below the hyperplane $\{(n, Y) : n \in \Delta, f(n) = \tilde{Y}(\bar{n})$. Since f is bounded, we must have $\tilde{Y}(n) \leq M||n|| = M\}$ for all $n \in \Delta$ for some $M < \infty$ (the norm of f), where we use the norm $||n|| = \int |n|$ for the functions n. □

1.2 Proof of Lemma 2.4: finite support of technologies

Proof. Since there exists τ such that $w(\tau, 0) > 0$ (by the assumptions on $\tilde{Y}$ and $\tilde{w}$), there is a strictly positive flow value $\varepsilon > 0$ that a worker can secure by working continuously in $(\tau, 0)$. Now, we will argue that in very old vintages, this value cannot be provided to workers since TFP eventually goes below any positive bound.

Now, fix some old vintage S. Note that in equilibrium, the value of every career segment l' (which may be of finite or infinite length, and where we cut off parts in vintages younger than S) spent in vintages above S must exceed the value of working for ε — if not, the worker should certainly replace the segment by ϵ:

$$\tilde{v}(l') \equiv \int_{l_0}^{l_1} e^{(\gamma-\beta-\delta)t} w(t - s'(t), h'(t)) dt \geq \int_{l_0}^{l_1} e^{(\gamma-\beta-\delta)t} \epsilon dt.$$

The inequality must hold since since l' also includes non-negative human-capital-accumulation costs.

Now, observe that the value of all discounted career segments in vintages older than S has to be lower than total discounted wages and thus production in those vintages. Integrate the above inequality over all career segments of type l' in the economy:

$$\int_{\text{all } l'} \tilde{v}(l') \leq \int_0^\infty e^{(\gamma-\beta-\delta)t} \int_{-\infty}^{t-S} \int_0^1 n(t,s,h) w(t,s,h) dh ds dt \leq$$
$$\leq \bar{y} e^{-\gamma S} \int_0^\infty e^{(\gamma-\beta-\delta)t} \left(\int_{s,h} n(t,s,h) \right) dt,$$

where in the last step I used that there is an upper bound on production for vintages even older than S of $e^{-\gamma S}\bar{y}$ for some $\bar{y} < \infty$ by Lemma A.1.

On the other hand, we know that each agent must weakly prefer working in an old vintage to working for ϵ — again, integrating up over all segments we get:

$$\int_{\text{all } l'} \tilde{v}(l') \geq \epsilon \int e^{(\gamma-\beta-\delta)t} \left(\int_{s,h} n(t,s,h) \right) dt.$$

But combining the above inequalities yields a contradiction: by choosing S large enough, we can make $e^{-\gamma S}\bar{y} < \epsilon$, making it impossible that very old vintages provide enough value to be attractive to workers. □

1.3 Proof of Lemma 2.3: all jobs filled in producing vintage

Proof. $Y(\bar{\tau}) > 0$ implies that some open ball $B_\epsilon(\bar{\tau}, \bar{h})$ lies in the support of n for some $\bar{h} \in (0,1)$. If there was some h' such that $(\bar{\tau}, h')$ did not lie in the closure of n's support, then there would be a ball $B_{\epsilon'}(\bar{\tau}, h')$ with $\epsilon' \leq \epsilon$ in which wages must be infinity — if not, firms should optimally choose to employ some workers there. But then, any career segment passing through $B_{\epsilon'}(\bar{\tau}, h')$ would yield infinite wages yet could be reached with a finite cost, implying that $W = \infty$. This is clearly impossible since resources in the economy are bounded, see Lemma A.1. □

1.4 Proof of Lemma 2.7: no holes in vintage space

Proof. Suppose there was some $\tau' \in (\tau_0, \tau_1)$ for which $Y(\tau') = 0$. Then there must be a positive measure of career segments ending on $[\tau_0, \tau')$ and the final wages of these segments must be equalized, which implies that all agents leave the vintage at once for some $\tau_e = \sup\{\tau : Y(\tau > 0)\}$ and that $w(\tau_e, h) = e^{-\gamma\tau_e}\bar{y} = (\beta + \delta - \gamma)W$ for all h. But this contradicts the fact that $w(T^*, h) = e^{-\gamma T^*}\bar{y} = (\beta + \delta - \gamma)W$ since $T^* \geq \tau_1 > \tau'$. □

2 Substitutability in production: $\rho = 1$

This subsection contains the proofs for the statements in Section 2.7 of the paper, throughout which the production function is assumed to be linear.

2.1 Proof of Lemma A.2: always enter newest vintage under substitutability

Proof. Suppose the worker chose a career segment with $s(t_1) > t_1$ on $t \in [t_1, t_2)$. Then this career is strictly dominated by choosing the same career in $s(t_1) = t_1$. Obviously, the same holds true for choosing $s(t) > t$ and $h(t) = 0$ on non-segments of positive measure. □

2.2 Proof of Lemma A.3: paths cross at most once

Proof. Suppose that the paths crossed again and denote by s the first crossing point, i.e. $s = \max_{u<t}\{u : h(u) \leq g(u)\}$. Together with $h(t) \geq g(t)$ this implies

$$h(t) - h(s) \geq g(t) - g(s) \Rightarrow \int_s^t \dot{h}(u)du \geq \int_s^t \dot{g}(u)du, \tag{11}$$

i.e. h must grow by at least as much as g over the interval to end up above g. By the assumption on the wage function, $w_h(h)$ is a decreasing function in h. Using the FOC (3), this implies that for all $u \in (s,t)$, we have

$$\begin{aligned} c'\big(\dot{h}(u)\big) = \int_u^t e^{-\beta v} w_h[h(v)]dv + e^{-\tilde{\beta}(t-u)} c'\big(\dot{h}(t)\big) < \\ < \int_u^t e^{-\tilde{\beta} v} w_h[g(v)]dv + e^{-\tilde{\beta}(t-u)} c'\big(\dot{g}(t)\big) = c'\big(\dot{g}(u)\big) \end{aligned}$$

since by assumption $\dot{h}(t) < \dot{g}(t)$ and $w_h[h(v)] \leq w_h[g(v)]$ point-wise; the inequality follows from c' being increasing. This again implies $\dot{h}(u) < \dot{g}(u)$ for all u, which in turn contradicts (11). □

2.3 Proof of Proposition 2.8: shorter horizon lowers tenure premia

Proof. Without loss of generality, take two optimal career segments $\tilde{h}$ and h in vintage $s = 0$ starting with $\tilde{h}(0) = h(0) = 0$ and $\tilde{T}^* > T^*$. Now suppose that $\tilde{h}(T^*) \leq h(T^*)$. First, note that $\dot{h}(T^*) = 0$ but $\dot{\tilde{h}}(T^*) > 0$ by equation (7) and the fact that $c'(\dot{h}) = V_h$. By Lemma A.3, the two paths cannot cross again for any $0 \leq t > T^*$. But this is a contradiction to $h(0) = \tilde{h}(0) = 0$. By the same argument, the two paths cannot intersect at any other point $0 < t < T^*(\gamma)$. So we must have $\tilde{h}(t) \geq h(t)$ and so $w(\tilde{h}(t)) \geq w(h(t))$, which implies the desired result. □

2.4 Proof of Proposition 2.9: faster growth shortens careers and lowers tenure premia

Proof. Let $Z(\gamma)$ be the value of being an inexperienced worker at time 0 given vintage productivity growth γ. The worker's problem is then to choose the switching time T when to leave the vintage to maximize

$$\tilde{Z}(\gamma, T) = K(T) + e^{-(\beta-\gamma)T} Z(\gamma),$$

where $K(\cdot)$ is given in (8). Invoking the assumption that $K(\cdot)$ is twice differentiable, the derivatives are computed as

$$\tilde{Z}_T(\gamma, T) = K'(T) - (\beta - \gamma) e^{-(\beta-\gamma)T} Z(\gamma), \tag{12}$$

$$\tilde{Z}_{TT}(\gamma, T) = K''(T) + (\beta - \gamma)^2 e^{-(\beta-\gamma)T} Z(\gamma). \tag{13}$$

The FOC for the optimal career length $T^*(\gamma)$ is $\tilde{Z}_T(\gamma, T^*(\gamma)) = 0$, the second-order condition is $\tilde{Z}_{TT}(\gamma, T^*(\gamma)) < 0$.

I will now state the problem in slightly different terms, which will enable us to derive how $Z^*(\gamma) \equiv Z(\gamma, T^*(\gamma))$ changes as γ changes. Note that since the worker's problem is recursive, we can write his value as $\tilde{Z}(\gamma, T) = K(T)/(1 - e^{-(\beta-\gamma)T})$. $T^*(\gamma)$ maximizes the function $\tilde{Z}(\gamma, T)$ for a given γ — indeed, the first-order conditions yield just the same result as in the problem above when maximizing $\tilde{Z}(\gamma, \cdot)$. But the formulation here is much more handy to see what happens to the agent's value when we change γ:

$$\frac{\partial Z^*(\gamma)}{\partial \gamma} = \left.\frac{dZ(\gamma, T^*(\gamma))}{d\gamma}\right|_{\gamma, T^*(\gamma)} = \frac{e^{-(\beta+\gamma)T^*}}{1 - e^{-(\beta-\gamma)T^*}} T^* Z(\gamma, T^*),$$

where the envelope condition $\tilde{Z}_T(\gamma, T^*(\gamma)) = 0$ is used.

Now, re-state the first-order condition for $T^*(\gamma)$ from (12):

$$K'\big(T^*(\gamma)\big) = (\beta - \gamma) Z^*(\gamma) e^{-(\beta-\gamma)T^*(\gamma)}.$$

Take the total derivative of this equation with respect to γ and use (13) to obtain

$$\frac{dT^*}{d\gamma} \underbrace{\tilde{Z}_{TT}\big(\gamma, T^*(\gamma)\big)}_{<0 \text{ by SOC (13)}} = \underbrace{(\beta - \gamma) e^{-(\beta-\gamma)T^*(\gamma)} Z^*(\gamma)}_{>0} \Bigg[\underbrace{\frac{T^*(\gamma)}{1 - e^{-(\beta-\gamma)T^*(\gamma)}} - \frac{1}{\beta - \gamma}}_{\equiv \Phi_\gamma(T^*)}\Bigg]. \tag{14}$$

We see that if T^* is large, also Φ_γ grows large, implying that also the effect $dT^*/d\gamma$ is negative and large in absolute value. When taking $T^* \to 0$ and using L'Hopital's rule, one finds that $\Phi_\gamma \to 0$, implying that the effects on T^* become very small.

The derivative of Φ_γ in T^* is

$$\Phi'_\gamma(T^*) = \frac{1 - \frac{1+(\beta-\gamma)T^*}{e^{(\beta-\gamma)T^*}}}{(1 - e^{-(\beta-\gamma)T^*})^2}.$$

Note that in the numerator, $1 + (\beta - \gamma)T^*$ is nothing but the first-order Taylor expansion of the function $e^{(\beta-\gamma)T^*}$ in T^* around 0, which always stays below the function itself since the exponential function is convex. This implies that the fraction in the numerator is always smaller than one, implying that Φ_γ is globally increasing. This in turn implies $\Phi_\gamma > 0$ (recall that $\lim_{T^* \to 0} \Phi(T^*) = 0$), which tells us we have $dT^*/d\gamma < 0$ for all $\gamma > 0$.[1]

There may exist values of γ where $T^*(\gamma) = 0$; in this case, the statements in the proposition are trivial.

Finally, since $T^*(\gamma)$ is a decreasing function it follows from Proposition 2.8 that $p_\gamma(t)$ is decreasing in γ for any fixed $t > 0$. □

3 Planner's problem

3.1 Stating the problem

Consider a social planner who weighs the utility of an agent born at t with $e^{-\beta t}$. Since it costs the planner $e^{-\delta(u-t)}$ units of time-u output to supply one unit to each surviving member of a cohort born at t and since utility is linear for all agents, it is easy to see that the planner's criterion is then to choose a function $n(t, s, h)$, which we require again to be C^1 on a given support S_n, to maximize

$$J(n) = \int_0^\infty e^{-\beta t} \left(Y\left(n(t;\cdot)\right) - C(t) \right) dt,$$

where $C(t)$ denotes the aggregate cost of human-capital accumulation at t. First, I will derive an expression for $C(t)$ given the optimal strategy to implement a given density n.

3.2 Proof of Lemma A.4: no-crossing measure is optimal

Lemma A.4, which is proven in the following, claims that the planner's optimal promotion strategy is such that agents' career paths inside a vintage never cross. Intuitively, if a positive measure of agents crossed each other's way, then one could improve upon the strategy by maintaining the ordering inside the vintage, making agents go shorter paths and hence lowering total cost for the planner.

Proof. I will proceed constructively to engineer the optimal measure on life paths by a discrete approximation procedure. Cut time and vintages into intervals of length $2^{-k}T^*$ for $k = 1, 2, \ldots$ to obtain grids $\{t_i^{(k)}\}_{i=1}^\infty$ and $\{s_i^{(k)}\}_{i=1}^{N_s}$. For human capital, slice such that the points $\{h_i^{(k)}\}_{i=1}^{N_s}$ yield intervals of length 2^{-k}. Approximate every path by connecting the middle of the interval $[h_i, h_{i+1}]$ it passes through at t_i for $t = 0, 2^{-k}, \ldots$ with straight lines. For every given measure μ on lives, summing up the costs over all possible promotion paths weighted by the densities induced by the measure μ gives us an approximation $C_k(\mu)$ for the total cost of human-capital accumulation for this μ.

Now, we will construct a lower bound C_k^* on this cost for a fixed iteration k in the algorithm. Note that it is enough to consider the task of moving workers

[1] Note that these calculations fail to provide us with any upper bound on $dT^*/d\gamma$, so in principle this change can be arbitrarily large.

between t_i and t_{i+1} for each point in time. It does not matter how we combine these path segments sequentially later, any such combination must obviously yield the same value.

Without loss of generality, consider the case $k = 1$ for $t_1 = 0$ and $t_2 = 1$ for the vintage $s = 1$ (note that the case $s = 0$ is trivial). The claim is that it cannot be optimal to choose a promotion scheme under which the paths of a positive measure of agents cross. Suppose we chose a promotion scheme under which a positive measure of agents crossed, i.e. a measure $\bar{\epsilon}$ went from $\bar{h}_0$ to h_1 and a measure $\epsilon > 0$ from h_0 to $\bar{h}_1$, where all the mentioned h-levels are center points of the approximation grid, and where $\bar{h}_j > h_j$. Now, set $\epsilon' = \min\{\epsilon, \bar{\epsilon}\}$ and consider the alternative of moving ϵ' agents from $\bar{h}_0$ to $\bar{h}_1$ and the measure ϵ' from h_0 to h_1. This would dominate the original allocation because of the following argument: take z to be the intersection of the lines $\bar{h}_0$ to h_1 and h_0 to $\bar{h}_1$. Then, clearly the process of sending everybody to z but then exchanging the flows to keep workers positions in the hierarchy fixed is just as cheap as the original policy. However, notice that this new policy must be weakly inferior to sending workers on the direct line $\bar{h}_0$ to $\bar{h}_1$ and h_0 to h_1, since this is the cost-minimizing strategy by Jensen's inequality.

Also, notice that there always exists a policy which does not make any worker flows cross: first, fill the uppermost interval at $t = 1$ with the uppermost workers from $t = 0$; proceed by filling the second interval with the uppermost workers left at $t = 0$ after the first step, and so forth. It is also clear that any process that does not follow these rules must make some workers cross and that any such process can be rendered into the proposed no-crossing algorithm by a finite number of improving operations; this shows that the no-crossing mechanism is optimal for a fixed k.

Obviously, the values C_k^* converge to the value of implementing the no-crossing measure μ_{nc}. Now, observe that no other measure μ' can yield a cost strictly lower than this: if we approximate μ' by the above scheme, by the above argument it must be that $C_k(\mu') \geq C_k(\mu_{nc})$. This precludes $C(\mu') = \lim_{k\to\infty} C_k(\mu') < \lim_{k\to\infty} C_k^* = C^*$.

It remains to prove that the lines of the no-crossing measure follow the proposed law. By the algorithm above, it is clear that an agent who at t had $N(t, s, h)$ workers above himself (position h) in vintage s and survives until $t + u$ will have $\exp(-\delta u)N(t, s, h)$ workers above himself at $t+u$ if none of the other workers crosses his path. This proves the second claim of the statement. □

3.3 The aggregate promotion cost $C(t)$

In the following, it will prove useful to work with the anti-cdf $N(t, s, h) \equiv \int_h^1 n(t, s, \tilde{h})d\tilde{h}$. Since workers' paths do not cross according to Lemma A.4, this function must decrease at the death rate δ when we evaluate it along an agent's path staying in a fixed vintage s. A first-order approximation following a career line $\{h(t), \tau(t)\}$ yields:

$$N_t(t, s, h) + \dot{h}(t, s, h)N_h(t, s, h) = -\delta N(t, s, h), \tag{15}$$

where we note that $N_h = -n$. Taking the h-derivative of the above and imposing stationarity yields the PDE for the evolution of n, which we already know from competitive equilibrium, see equation (6).

Re-arranging Equation (15) gives us an expression for the career slope $\dot{h}$ that the planner should choose given that she wants to implement a given n:

$$\dot{h}(t,s,h) = \frac{N_t(t,s,h) + \delta N(t,s,h)}{n(t,s,h)}. \tag{16}$$

In order to aggregate costs over all agents, we have to weigh the cost of $\dot{h}$ by the mass of agents across the (t,s,h)-space and obtain $C(t) = \int_{s,h} n(t,s,h)c[\dot{h}(t,s,h)]$.

3.4 First-order conditions

The strategy to obtain the first-order conditions (FOCs) for the planner's problem is as follows: I will first allow the planner to to choose any – possibly time-varying – density $n(t,s,h)$. I then look for a stationary distribution which solves this unrestricted problem. This ensures that the planner would not want to deviate from the stationary density $n(\tau,h)$ although she could do so. I will first restrict S_n to the entire rectangle below a maximal vintage age T and then let T vary to find the optimal support T^*.

It turns out that it is useful to introduce the variable $u(t,s,h) \equiv n_t(t,s,h)$ and connect it to the functions n, N and N_t with equality constraints. The Lagrangian is then[2]

$$\begin{aligned}
\mathcal{L} = &\int_0^\infty e^{-\beta t}\left[\int_{t-T}^t Y(t,s) - e^{\gamma s}\left(\int_0^1 c\big[\dot{h}(t,s,h)\big]n(t,s,h)dh\right)ds\right]dt+ \\
&+ \int_{t,s,h} e^{-(\beta-\gamma)t}\left[\nu(t,s,h)\left(\dot{h} - \frac{\dot{N}+\delta N}{n}\right)+\right. \\
&\quad + \lambda(t,s,h)\left(n_0(s,h) + \int_0^t u(\tilde{t},s,h)d\tilde{t} - n(t,s,h)\right)+ \\
&\quad + \eta(t,s,h)\left(\dot{N}(t,s,h) - \int_h^1 u(t,s,\tilde{h})d\tilde{h}\right)+ \\
&\quad + \xi(t,s,h)\left(N(t,s,h) - \int_h^1 n(t,s,\tilde{h})d\tilde{h}\right)+ \\
&\quad \left. + \mu(t)\left(1 - \int_{t-T}^t\int_0^1 n(t,s,h)dhds\right)dt\right],
\end{aligned}$$

where the Lagrange multipliers are scaled by $e^{-(\beta-\gamma)t}$ to render them stationary. The set of constraints linked to the multipliers ν is taken from equation (16). The

[2]See Luenberger (1973) for necessary conditions of constrained-optimization problems in infinite-dimensional spaces.

constraints connected to μ enforce that total population not exceed the bound 1. The rest of the constraints link the various variables related to the density n.

The FOC with respect to $\dot{N}(t,s,h)$, $\dot{h}(t,s,h)$ and $N(t,s,h)$ immediately tell us that η is the marginal cost of human-capital accumulation, and that ν and ξ are closely linked to η:

$$\begin{aligned}\eta(t,s,h) &= e^{-\gamma\tau}c'\big(\dot{h}(t,s,h)\big),\\ \nu(t,s,h) &= e^{-\gamma\tau}c'\big(\dot{h}(t,s,h)\big)n(t,s,h),\\ \xi(t,s,h) &= \delta e^{-\gamma\tau}c'\big(\dot{h}(t,s,h)\big).\end{aligned} \tag{17}$$

Using these equalities, the FOC with respect to $n(t,s,h)$ becomes

$$\begin{aligned}\lambda(t,s,h) =& w(t,s,h) - e^{-\gamma\tau}c\big(\dot{h}(t,s,h)\big) + e^{-\gamma\tau}\dot{h}(t,s,h)c'\big(\dot{h}(t,s,h)\big) - \\ &- \mu(t) - \delta\int_0^h \eta(t,s,\tilde{h})d\tilde{h},\end{aligned} \tag{18}$$

where we recognize in the terms involving $c(\cdot)$ the Hamiltonian from the value function (2) in the worker's problem. The last remaining derivative is the one with respect to to $u(t,s,h)$, which will prove crucial to obtain the PDE that is equivalent to the HJB (2):

$$\int_\tau^T e^{-(\beta-\gamma)(\tilde{\tau}-\tau)}\lambda(\tilde{\tau},h)d\tilde{\tau} = \int_0^h \eta(\tau,\tilde{h})d\tilde{h}. \tag{19}$$

At a stationary solution, we require that the density fulfill $n(t,s,h) = \bar{n}(\tau,h)$. As a consequence wages grow at rate γ: $w(t,s,h) = e^{\gamma t}\bar{w}(\tau,h)$. The Lagrange multipliers must also be time-independent, i.e. $\nu(t,s,h) = \bar{\nu}(\tau,h)$, $\mu(t) = \bar{\mu}$ and so forth. Again, I drop the bar-notation in the following.

When substituting the expressions for the Lagrange multipliers (17) and (18) into (19) and imposing stationarity, one obtains

$$\begin{aligned}\int_\tau^T e^{-(\beta-\gamma)(\tilde{\tau}-\tau)}\Big[& w(\tilde{\tau},h) - e^{-\gamma\tilde{\tau}}c\big(\dot{h}(\tilde{\tau},h)\big) + \dot{h}(\tilde{\tau},h)c'\big(\dot{h}(\tilde{\tau},h)\big) - \mu - \\ &-\delta\int_0^h \eta(\tilde{\tau},\tilde{h})d\tilde{h}\Big]d\tilde{\tau} = \int_0^h e^{-\gamma\tau}c'\big(\dot{h}(\tau,\tilde{h})\big)d\tilde{h} \equiv \Lambda(\tau,h).\end{aligned} \tag{20}$$

We will now see that $\Lambda(\tau,h)$ is an "excess-value function": it tells us what the value of an agent to the planner in position (τ,h) is *in excess* of the unconditional value μ of an additional unskilled agent.

Directly from (20), we can get the following insights that are analogous those from the discussion of competitive equilibrium. First, when $\tau \to T$, the left-hand side and with it the marginal cost of human-capital accumulation $c'(\dot{h})$, and hence $\dot{h}$ itself, go to zero. This says that one should not accumulate human capital anymore just before the vintage shuts down, which also implies that $w(T,h)$ must be weakly increasing in h by non-negativity of the multipliers η. Second,when we let $h \to 0$, the right-hand side of (20) goes to zero and we see that $\lambda(\tau,0)=0$ for all τ. This says that for all entry jobs the value function must be equalized. Third, when we let both $\tau\to T$ and $h\to 0$ and use the insights from above, we obtain $w(T,0)=\mu$. This says that $w(T,0)$ is the reference wage of the economy: it does not provide any valuable experience, so it has to be just as attractive per se as any other career (in flow terms).

Now, take the derivatives of Λ in (20) in both dimensions to see how this excess-value function behaves on the interior:

$$\Lambda_h(\tau,h)=e^{-\gamma\tau}c'\big(\dot h(\tau,h)\big), \tag{21}$$

$$-\Lambda_\tau(\tau,h)=w(\tau,h)-e^{-\gamma\tau}c(\dot h)+e^{-\gamma\tau}\dot h c'(\dot h)-\mu-(\beta+\delta-\gamma)\Lambda(\tau,h). \tag{22}$$

When adding an agent's value at the start of a career segment $W=\mu/(\beta+\delta-\gamma)$ to Λ by defining $V=\Lambda+W$, we obtain

$$-V_\tau(\tau,h)=w(\tau,h)-e^{-\gamma\tau}c(\dot h)+\dot h V_h-(\beta+\delta-\gamma)V(\tau,h), \tag{23}$$

where we use $V_h=\Lambda_h=c'(\dot h)$. When imposing the boundary conditions $V(\tau,0)=V(T,h)=0$ for all τ and for all h, this system is the same as the agent's HJB (2) and its boundary conditions in the decentralized problem. The proof for Proposition 2.14 in Subsection 3.7 of this online appendix will discuss equivalence of the planner's problem to the competitive equilibrium more carefully; before, it is useful to analyze the effects of variations in T.

3.5 Discussion: uniqueness and existence of solution

Proposition A.5 tells us that the solution to the planner's problem is unique. It is worthwhile to note that the argument of the proof does not hinge on the assumption of n being continuous or differentiable, nor on any restriction on S_n.

If Y is not strictly concave, uniqueness is a slightly more complicated issue. Take the example with a linear production function: uniqueness of the planner's problem depends on uniqueness of the partial-equilibrium solution for the agent. If the agent's problem has a unique solution for any starting value of h, then the solution to the planner's problem is unique.

Existence of equilibrium is not a problem computationally, but cannot be established formally without making an equicontinuity assumption on the function space for n. The following discussion will show why.

In order to reap the benefits of compactness, we may restrict ourselves to seek a maximand n in the planner's problem that satisfies the following conditions: we reparameterize the density from $n(t,s,h)$ to $n(t,\tau,h)$, which ensures that the partial derivative $n_t\to 0$ everywhere as $t\to\infty$ for any n that converges to a stationary distribution. Then, compactify the t-dimension using an increasing concave transform that maps $[0,\infty)\to[0,1)$ and define $\lim_{t\to\infty} n(t,\tau,h)$ as $\tilde n(1,\cdot)$. We then impose a Lipschitz condition uniformly on the entire family of $\tilde n$ in which we look for a maximand (This essentially means that the modulus of continuity for the original

n becomes always stricter in the t-direction as t increases; the "wiggling" in n has to become smaller as t grows).

If we further assume that n is point-wise bounded – which is unproblematic – , equicontinuity allows us to employ the Arzela-Ascoli theorem which tells that such a family of functions $\tilde{n}$ is a compact set; see Rudin (1973) for a statement of the theorem. The computational exercises indicate that indeed the optimizer n^* satisfies a Lipschitz condition; decreasing the grid size to allow for always steeper functions n does not significantly alter the solution after some point. However, it is hard to prove that the solution really satisfies such a Lipschitz condition.

3.6 Varying T

So far, we had fixed the maximal vintage age T and imposed it on the planner; we will now be concerned with varying T and finding the optimal T^* under the assumption that $\tilde{Y}$ is strictly concave. By the concavity argument in Lemma A.5, there is at most one T^* for which the planner's criterion is maximized. An argument analogous to the proof for Lemma 2.4 shows that $T^* < \infty$. However, it is very hard to further characterize T^*. Computationally, it may be found by finding the optimal n for each fixed T and then pick the value T^* that yields the highest value to the planner. The following discussion describes regularities and problems that arose in this process.

First, for $T < T^*$, the simulations usually yield the wage structure is not flat in the last vintage yet. In this case, an argument along the lines of Corollary 2.6 shows that it is preferable for the planner to extend the vintage horizon T marginally; marginal productivities for different h-levels are not aligned yet and there is room for further gains through human-capital accumulation.

Second, for $T > T^*$ computational problems may arise because of the following issue: the problem of finding the optimal n given T will usually not have a maximand in the space of continuous differentiable functions. To see this, suppose there was such a maximand $n^*(T)$. Since $J(n^*(T^*)) > J(n^*(T))$, by concavity also $J(n_\lambda) > J(n^*(T))$ where we define $n_\lambda = \lambda n^*(T^*) + (1-\lambda)n^*(T)$ for any $\lambda \in (0,1)$. In turn, any n_λ may be approximated arbitrarily well by any continuous, differentiable n with support until T. So there is a sequence of densities for which J converges to the global optimum, but the global optimum is not in the space we are considering since its support only extends to $T^* < T$ and is discontinuous at this point.

3.7 Proof of Proposition 2.14: equivalence of CE and planner's problem

Proof. I will first show that the global solution to the planner's problem constitutes a CE. Set wages $w(\tau,h) = \partial Y(\tau)/\partial n(\tau,h)$ for $\tau \leq T^*$ and $w(\tau,h) = w(T^*,0) = \mu$ for all $\tau > T^*$, all h. This implies that firms optimally choose not to produce for $\tau > T^*$ since even the cost-minimizing input combination leads to losses. For $\tau \leq T^*$, $n(\tau,h)$ is an optimal input choice and profits are zero. For agents, the

HJB (23) and its boundary conditions imply that any career segment which fulfills $\dot{h} = V_h$ everywhere is an optimal strategy with starting value μ. This weakly dominates any career segment in vintages (T^*, ∞). One may then insert agents into careers to engineer the entry density $n(\tau, 0)$ since agents are indifferent between all careers. Equation (6) ensures that the density n reproduces itself given the optimal decisions of agents.

Second, I prove that any CE is a solution to the planner's problem for some $T \leq T^*$. To start, note that the worker's HJB (2) and the corresponding optimal policy (3) in competitive equilibrium have their exact counterparts in Equations (23) and (22) for the planner's problem. Equation (20) follows by integrating from the boundary over τ and h, which in turn is equivalent to (19). Since the first-order conditions (17) and (18) can be used to define the Lagrange multipliers, equation (19) already ensures that all first-order conditions for the Lagrangian hold for any competitive equilibrium.

This means that any competitive equilibrium is a stationary point of the Lagrangian.[3] However, there can be at most one stationary point for a given T since J is a concave function and the set of permissible n is convex. Hence this stationary point must be the global maximum of the planner's problem corresponding to the T induced by the respective CE. As the discussion in the previous Subsection 3.6 showed, no such maximizer exists for $T > T^*$, which means that there cannot be any CE with $T > T^*$. □

It is hard to formally rule out competitive equilibria with $T < T^*$. If there is such a CE, then it must be that $\mu_T > \mu_{T^*}$ since these multipliers equal wages in the last vintage. This seems to suggest that $J_T > J_{T^*}$, which would be a contradiction to T^* being associated with a global maximizer. However, J also includes the excess value for agents already born at $t = 0$ starting with $h(0) > 0$, which is not comprised in the multiplier μ.[4]

3.8 Proof of Proposition A.6: upper bound on aggregate learning cost

Proof. The planner's criterion can be decomposed as follows by integrating over the single agents' values:

$$J = \frac{w(T,0)}{\beta + \delta - \gamma} + \underbrace{\int_{\tau,h} \Lambda(\tau, h)}_{\equiv \bar{\Lambda}} + \int_0^\infty e^{-(\beta-\gamma)t} \delta \frac{w(T,0)}{\beta + \delta - \gamma} dt = \frac{w(T,0)}{\beta - \gamma} + \bar{\Lambda},$$

where the first equality decomposes the value for the measure one of agents alive at $t = 0$ according to $V = W + \Lambda$ and uses the fact that $\Lambda = 0$ for all agents born

[3] A stationary point is defined as a point where the Frechet-derivative is zero in all directions, see Luenberger (1973) — this is the equivalent to the gradient being zero in $\mathbb{R}^n$.

[4] In the numerical exercises, however, enforcing $T < T^*$ always led to an increasing wage structure at T which is not compatible with a CE according to Corollary 2.6.

later. We can now juxtapose this decomposition and the decomposition of J into production and promotion costs:

$$Y(0) - C(0) = (\beta - \gamma)J = w(T, 0) + (\beta - \gamma)\bar{\Lambda}.$$

$\Lambda(\tau, h) \geq 0$ for all (τ, h) implies $\bar{\Lambda} \geq 0$, and thus the claim in the proposition follows. □

References

Luenberger, D. G. (1973), *Optimization Using Vector Space Methods*, AddisonWesley, New York, NY, USA.

Rudin, W. (1973), *Functional Analysis*, McGraw-Hill, New York, NY, USA.