Experience vs. Obsolescence: A Vintage-Human-Capital Model *

Matthias Kredler[†]

THIS DRAFT: February, 2007

Abstract

This paper presents a continuous-time model for vintage human capital that draws both on classical human-capital models à la Ben-Porath (1967) and on the vintage-human-capital model of Chari and Hopenhayn (1991). Production occurs in a continuum of vintages. In each of these vintages, the different levels of human capital are inputs to a production function. Agents choose which vintage to enter and at what speed to accumulate human capital. The model replicates the typical hump-shaped age-earnings profiles we know from Mincer regressions. Some agents experience a decline in real wages in later stages of their careers. There is no depreciation of human capital in the model – wages decline because some agents' specific skills become less scarce over the lifecycle of an industry. Furthermore, the model is in line with key stylized facts on the wage distribution across firms.

*I thank my advisor Boyan Jovanovic for the encouragement to pursue this project in the initial phase and the valuable advice he gave to me during the entire course of the project. I also thank Jess Benhabib, Christopher Flinn, Jonathan Halket, Greg Kaplan, Silvana Melitsko and Gianluca Violante for discussion and comments.

†Department of Economics, New York University, mk1168@nyu.edu

1 Introduction

Recently, many models have been constructed in the spirit of Ben-Porath's (1967) seminal work on human-capital accumulation. Examples include Huggett, Ventura, and Yaron (2006), who use a human-capital model to explain the distribution of earnings over the life cycle, and Guvenen and Kuruscu (2006), who use a model in the same spirit to examine the implications of skill-biased technological change on the earnings distribution.

As is done in these two studies, most models that address age-earnings dynamics are single-sector models where the wage structure cannot vary systemically across firms. However, econometric work has established that wage profiles *do* indeed vary significantly across firms: Michelacci and Quadrini (2005) find that the tenure premium is higher in small, young, fast-growing firms and lower in bigger, established, slow-growing ones.[1] Also, they document that starting wages are lower in fast-growing sectors.[2] In line with this, a large number of econometric studies has found that larger firms pay higher wages than smaller ones (CITATION). So it seems that there is systematic connection between wage profiles and firm attributes; if workers stay with their firms for at least some time, then agents in some firms will have systematically steeper wage profiles than others.

A promising avenue of modelling the wage structure of firms and age-earnings profiles jointly is suggested by the model of Chari and Hopenhayn (1991). In their overlapping-generations framework, production occurs in different vintages; workers acquire vintage-specific human capital in the first period of their lives and can use this experience as "managers" (i.e. experienced workers) in the second period. It turns out that in equilibrium the wage profiles of all agents are increasing from the first to the second period, and that wage growth is systematically higher in younger vintages, just as in the data analyzed by Michelacci and Quadrini (2005).

The model presented in this paper builds on the Chari-Hopenhayn framework in adopting the vintage structure. However, it introduces a finer structure for human capital: In each vintage, output is produced using a range of labor inputs that are differentiated by the level of human capital. These

[1] They build a model where firms are financially constrained and "borrow" from their workers in the form of increasing wage profiles; small companies that want to grow are severely constrained and borrow a lot.

[2] The fact that fast-growing firms pay lower wages is independently established in studies by Hanka (1998) and Bronars, Stephen G. and Famulari, Melissa (2001).

human-capital levels can be interpreted as positions in a hierarchy of a firm that the workers climb receiving successive promotions. Agents are infinitely-lived and can climb the hierarchy ladder by learning new tasks, which is penalized by a convex cost functional like in Ben-Porath-type models.

In equilibrium, agents that enter the vintage at different points of will face different wage prospects over their careers, giving rise to heterogeneity in age-earnings profiles. In equilibrium, agents who enter new technologies have steeper profiles and accumulate human capital faster in equilibrium then workers who enter the vintage later on. Figure 1 plots the (log-)age-earnings profile from the model and compares them to a plot of one of Mincer's famous regressions (see figure 2), here carried out by groups defined over years of education.

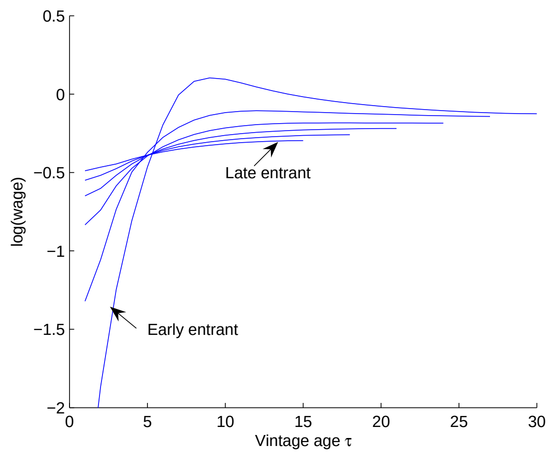
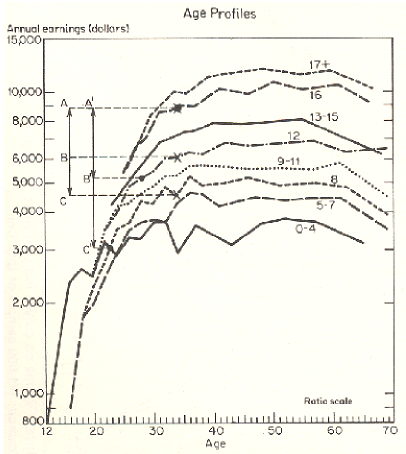


Figure 1: Log wages over career by entry cohort

The model generates age-earnings profiles that are reminiscent of the ones Mincer observed in the data: Agents that invest a lot of time in human-capital accumulation in the beginning of their lives forgo earnings in earlier stages in order to enjoy high wages in later stages of their careers; these agents may be identified with the individuals having 13 years and more of education in figure 2 (Note that in Mincer's plots, the lines of highly-educated agents would cross those of less-educated ones if plotted from an earlier age since earnings during schooling years are very low or even negative). Agents that accumulate less human capital have flatter wage profiles; these agents may be identified with the low-education workers in Mincer's regressions.



Log annual earnings vs. age, grouped by years of education. Source: Mincer (1974)

Figure 2: Age-earnings profiles by education groups (Mincer, 1974)

An especially noteworthy feature of the profiles generated by the model is that some are downward-sloping in the end. A slight downward slope is recognizable in Mincer's data towards the end of the profiles, but it is not very clear. More importantly, real-wage decreases were also found in other data sets that are less susceptible to sample-selection bias versus the end of the work life, e.g. in Baker, Gibbs, and Holmstrom's (1994) study.[3] In the model presented here, real-wage decreases occur despite the fact that human capital does not depreciate. The mechanism at work is a form of obsolescence: Some agents acquire skills that are scarce when a technology is still young. Wages for these individuals are high in the beginning, which lures more workers into this technology who learn the scarce skill. These workers create always larger supply of the once scarce skill and drive down its returns as the technologies[4]

Furthermore, the model generates heterogeneity in the slope of age-earnings

[3]Sample-selection might occur because of the following reason: If higher-wage workers tend to retire earlier, the sample for very old ages contains more workers with relatively low productivity; all workers' wages could be stable and the line in Mincer's graph would be downward-sloping due to this selection. Baker, Gibbs, and Holmstrom (1994), by contrast, follow workers over their work life at a firm and still find a significant fraction of real-wage decreases.

[4]This effect cannot occur in standard Ben-Porath-type models since production is assumed to be linear in all levels of human capital.

profiles between agents that are ex-ante alike. In this way, it introduces a rationale why a phenomenon called "overtaking"[5], as studied in Hause (1980), for example, should occur also between workers of the same skill level and not only between workers with different learning ability.

If vintages are interpreted as sectors or firms, the model's predictions on the wage structure inside and across firms are in line with some key stylized facts from the empirical literature. The youngest and fastest-growing vintages pay low entry wages, but offer high prospective wage growth to entrants. The established vintages (which are bigger in terms of both employment and production) pay higher wages on average, but offer less prospective wage growth. In regard to these first-order effects, the model replicates the results of Chari and Hopenhayn (1991). However, the model constructed here also has predictions for higher-order effects.

The remainder of the paper is organized as follows: Section 2 describes the model setup, shows equivalence of the market equilibrium to the planner's solution and derives properties of the equilibrium. Section 3 presents a discrete approximation technique to compute an equilibrium. Section 4 presents the results for a representative pair of parameters, explains the main effects and compares them to the existing literature. Section 5 concludes.

2 Model

2.1 Technology

Time is continuous. In every instant s, a new production technology (or *vintage*) arrives that is available to the agents in the economy for all $t \geq s$. We will either refer to the vintages by their birth date, s, or – especially in a stationary setting – identify them with their age $\tau \equiv t - s$. All vintages produce the same good.

As inputs, the production technology of age τ uses labor inputs which are arranged on a hierarchy and indexed by $0 \leq x \leq 1$. The different labor inputs on this ladder can be thought of as tasks that are increasing in difficulty and that tasks with a higher index require more vintage-specific human capital. Section 2.2 will specify exactly how this form of human capital is accumulated

[5] "Overtaking" describes the fact that two wage profiles that can have the same present value differ in steepness and hence have to cross each other; it is usually ascribed to different speeds in on-the-job training, as in Hause's paper.

by workers.

The production function is supposed to capture the following notions: (i) Newer vintages are more productive when the same inputs are used; (ii) the production function is complementary in its inputs, i.e. it is optimal to find a good mix between the different inputs of human capital;(iii) tasks that require more experience in a vintage have higher returns than those that require less experience if the same number of workers is employed in both tasks. Specifically, I choose the following constant-elasticity-of-substitution (CES) production function where total factor productivity (TFP) is exponentially increasing in the vintages:

$$Y(t,s) = e^{\gamma s}\left(\int_0^1 [f(x)n(t,s,x)]^\rho dx\right)^{1/\rho}$$

where $0 \leq \rho \leq 1$, $n(t,s,x)$ is the density of workers with experience x in vintage s at time t, and $f(x)$ is a continuously differentiable, non-decreasing, weakly concave function in x that specifies the returns to experience. Note that this production function will induce a trade-off: Newer vintages will be more productive in terms of TFP, but in older vintages experienced labor will be more abundant, which increases productivity.

Total output in the economy at t is

$$Y(t) = \int_{-\infty}^{t} Y(t,s)ds.$$

Firms take the wages for all labor inputs as given in any instant. Since the production technology is constant-returns-to-scale (CRS), profits will be zero for time t and vintage s in equilibrium, of course. Workers will be paid their marginal product in equilibrium:

$$w(s,t,x) = \frac{\partial Y(t,s)}{\partial n(s,t,x)} = e^{\gamma s} f(x)^\rho \left(\frac{\int_0^1 n(t,s,\tilde{x})d\tilde{x}}{n(t,s,x)}\right)^{1-\rho}$$

2.2 Workers

There is a continuum of agents that has mass one. Agents are homogenous in preferences: They have linear utility and discount the future at rate β. Each agent chooses a work life $\{s(t), x(t)\}_{0\leq t<\infty}$, which consists of a function $s(t)$ specifying the vintage the agent works in at time t and a function $x(t)$

specifying the task he performs at time t. It is required that the vintage already exist at time t, i.e. $s(t) \leq t$, and that $s(t)$ be a measurable function in t.

As for human-capital accumulation $x(t)$, we require that a worker start her work life in position $x = 0$ when she enters the vintage; mathematically we impose that $x(\bar{t}) > 0$ only if there is an interval $[a, b]$ around $\bar{t}$ such that $s(t) = \bar{t}$ for all $a < \bar{t} < b$. Also, the worker looses all experience in a vintage once he drops out[6] Apart from this, the function $x(t)$ is required to be continuous and differentiable piecewise, i.e. it is allowed to have discontinuities between smooth intervals. There is no cost of switching between vintages.

To capture the notion that human-capital accumulation inside a vintage is costly, we specify the following cost functional for a career segment $x(t)$ over the interval $[t, t+r]$ in a vintage s:

$$C([t, t+r]) = e^{\gamma s} \lim_{n \to \infty} \sum \frac{c}{2} \left(\frac{\max\{x(t_{n+1}) - x(t_n), 0\}}{\Delta t} \right)^2 \Delta t$$

This says the following: When we chop up time in a fine grid, the cost of climbing the career ladder (per unit of time) is quadratic in the local slope. When one takes steps down, however, this is costless. Taking the limit for any function that takes an upward jump shows that the cost of this is infinite for the agent and hence will not be optimal. Downward, however, jumps can occur. So $x(t)$ will be a differentiable function on its upward-sloping parts and the cost will be $\int_t^{t+r} c\dot{x}(t)/2dt$; convexity implies it is optimal to climb the ladder in a steady fashion rather than to make abrupt leaps, since large slopes are penalized more than proportionally due to the convexity of the quadratic function.

Furthermore, note that the cost of human-capital accumulation is growing at the pace of TFP to ensure stationarity of the system; in economic terms, this means that the costs of human-capital accumulation relative to productivity do not change.

To start off the economy, we also need to specify the initial conditions for agents. Assume that each agent enters the economy with some experience level x for a vintage of age $\tau \geq 0$, and that there is a density $n_0(s, x)$ over these endowments.

[6]This assumption is imposed for tractability and may be relaxed; in equilibrium, workers would not want to return to vintages they have once left.

2.3 Stationary equilibrium

In a stationary environment, I require that the density $n(t, s, x)$ depend only on the age of the vintage $\tau = t - s$ but not on time:

$$n(t, s, x) = n(s + \tau, s, x) = n_{stat}(\tau, x)$$

This means that in any area of the vintage-experience space, the mass of agents stays the same when we index the vintages by age τ instead of their birth date s. Stationarity immediately implies that wages and production grow at rate γ. From now on, we will only work with the stationary distribution; I thus drop the subscript and write simply $n(\tau, x)$.

A *stationary equilibrium* is a distribution $n(\tau, x) = n_0(\tau, x)$, a measure μ on all possible work lives $l(t) = \{\tau(t), x(t)\}$, a wage function $w(\tau, x)$ such that

- The measure μ yields $n(\tau, x)$ for all t.
- The distribution $n(\tau, x)$ is the optimal choice for each firm given wages $w(t, t - \tau, x) = e^{\gamma t} w(\tau, x)$
- Each life $l(t)$ is optimal given the wage profile.

2.4 Properties of equilibrium

Define the value function $V(\tau, x)$ of an agent positioned in the vintage of age τ at level x as the supremum of

$$V(\tau, x) = \sup_{l} \left\{ \int_0^{\infty} e^{\gamma - \beta} w(\tau_t, x_t) + e^{-\beta t} C(x_t) dt \right\},$$

where the supremum is taken over all feasible lives l. Note that by stationarity, the value function at time t is given by $e^{\gamma t} V(\tau, x)$.

The first thing to note is the following: that the value function is weakly increasing in the human-capital (or hierarchy) level x inside a vintage:

Lemma 2.1 (Value function weakly increasing) *The value function $V(\tau, x)$ is weakly increasing in x for all τ.*

Another insight is that at the beginning of the career, all careers pursued in equilibrium have to provide the same value.

Lemma 2.2 Value is equal for all career starters *We have* $V(\tau,0) = W \equiv \max_\tau V(\tau,0)$ *for all* τ, *and* $V(\tau,x) \geq 0$ *for all* τ *and* x.

The economic intuition is of course that no agent would pursue a career with a lower value than another, since she could switch at zero cost.

Proof It is always an option to start in the career that provides W, so the value everywhere must be at least W. Also, no starting point (τ, x) can provide more value than W by definition. ■

Another result that allows us to make some headway is that we do not have to consider the entire space of vintages $0 \leq \tau < \infty$, but can restrict ourselves to a finite interval $0 \leq \tau \leq T$:

Lemma 2.3 (Finite support of technologies) *In a stationary dynamic equilibrium, there is a bound* T *on the age of the vintages beyond which no production occurs, i.e.:* $Y(\tau) = 0$ *for all* $\tau > T$.

The proof uses the argument that workers can always secure some positive wage in a new vintage without going through training, but that old vintages' productivity goes to zero such that in the end they cannot provide a value higher than this small wage:

Proof There is a small positive value ε in equilibrium that a worker can secure, for example by working continuously as an unskilled worker in the newest vintage: $\varepsilon = f(0)/(\gamma - \beta)$. Now, we will argue that in very old vintages, this value cannot be provided to workers since TFP is so low. To see this, observe that maximal productivity in a vintage cannot exceed some maximal productivity $\bar{y}$ that is achieved when marginal factor productivities are equalized across inputs. However, this bound decreases exponentially with τ since TFP falls.

Now, fix some very old vintage S. Note that in equilibrium, the value of every career segment (i.e. that somebody spends in a vintage older than S must exceed the value of working for ε in the newest vintage. The wage payments that go into the earnings of workers in vintages older than S have to be equal to production in the aggregate. However, note that in all vintages above S productivity is below $e^{-\gamma S}$, which is smaller than ε for S large enough; costs of human-capital accumulation will even lower the value further. So it is impossible that all career segments in careers in vintages older than S provide more value than working for ε always. ■

In the following, we will look for a functions $n(\tau, x)$, $w(\tau, x)$ and $V(\tau, x)$ that are at least once differentiable in both directions. Then lemma 2.1 above implies that $\partial V(\tau, x)/\partial x \geq 0$. Denote the slope of a worker's career at t by $a(t) = \partial x(t)/\partial t$ and by $V(\tau, x)$ the value of being at an interior position (τ, x). Then the slope has to fulfill the following Hamilton-Jacobi-Bellman (HJB) equation:

$$-\frac{\partial V(\tau, x)}{\partial \tau} = w(\tau, x) - (\beta - \gamma)V(\tau, x) + \max_a \left\{ -\frac{c}{2}a^2 + a\frac{\partial V(\tau, x)}{\partial x} \right\} \quad (1)$$

The equation says the following: If we know the value function for a given τ for all experience levels x, we can get the value a tiny bit left of this τ doing the following: Let the agent choose the optimal slope a, which is contingent on the slope of the value function and the cost of learning. Then the change in the value function for some small h to the left (keeping x fixed) is the gain the agent gets from moving up in the hierarchy (the term inside the max-operator) and another term which is the difference between the current wages and the flow value of $V(\tau, x)$ under the "modified" discount factor $\beta - \gamma$ (note that value increases in time since TFP grows).

Since a tractable form was assumed for the cost of human-capital accumulation, we can get the optimal policy in closed form:

$$a^*(\tau, x) = \frac{1}{c}\frac{\partial V(\tau, x)}{\partial x} \quad (2)$$

This says that agents will accumulate human capital faster the greater the value differential in the hierarchy. Also, the slope is inversely related to the marginal cost in accumulating human capital.

Plug this optimal solution back into equation (3) to get the following non-linear first-order partial differential equation (PDE):

$$-\frac{\partial V(\tau, x)}{\partial \tau} = w(\tau, x) - \tilde{\beta}V(\tau, x) + \frac{1}{2c}\left(\frac{\partial V(\tau, x)}{\partial x}\right)^2 \quad (3)$$

$$(4)$$

This equation describes the behavior of $V(\tau, x)$ given a fixed wage profile $w(\tau, x)$. In a partial equilibrium, this equation would give us workers' behavior given market wages.

Now we need rule that tells us how the density $n(\tau, x)$ –and hence wages $w(\tau, x)$– evolve given the optimal local behavior of agents described by the HJB.

Given an entry density $m(\tau)$ which specifies the density of people entering new careers at $x = 0$ we can get the following forward equation governing the evolution of the density. The following gives the relationship between the entry density, the slope and the density of workers and a forward equation for interior points:

$$\frac{\partial n(\tau,x)}{\partial \tau} + a(\tau,x)\frac{\partial n(\tau,x)}{\partial x} = -\frac{\partial a(\tau,x)}{\partial x} n(\tau,x) \tag{5}$$

The equation is the usual mass-transport equation for densities of moving particles.[7] In the appendix it is proven that the equation also holds for points on the upper bound of the experience space (i.e. $x = 1$) using $a(\tau, 1) = 1$ and $a_x(\tau, 1) = \lim_{x \to 1} a_x(\tau, x)$.[8]

Equation (5) says the following: On the left-hand side, we see how the density changes when we follow the path of the worker on a very small interval of time: one unit to the right and $a(\tau, x)$ units upward. The right-hand side says that the density falls at a specific rate on this line – this rate is given by the change of career slopes across the x-dimension. Suppose that this slope did not change and all workers in vintage τ moved upward at the same slope a; then the density n would of course stay the same along each career line. However, if the slopes increase with experience level x, then agents would move apart from each other, which would result in a thinning of n along the career lines.

Also, it is sometimes desirable to talk about the density of entry into a vintage. Define the entry density $m(\tau)$ such that the mass of agents entering vintages between τ and τ' per unit of time is $\int_{\tau}^{\tau'} m(t)dt$. Then we have

$$n(\tau, 0) = \frac{m(\tau)}{a(\tau, 0)}$$

The interpretation is straightforward: the higher the entry density in a vintage, the people there are in the starting job. The faster people move upward from the starting position, the less crowded this position.

[7] It can also be seen as a special case of the Kolmogorov forward equation for Ito processes where the stochastic component is set to zero. Since there is also an HJB for a stochastic environment, a stochastic version of the model with Brownian shocks to human capital, for example, would also amenable to analysis by PDEs.

[8] Let subscripts denote partial derivatives, for example $a_x(\cdot, \cdot) = \partial a(\cdot, \cdot)/\partial x$.

2.5 Planner's problem and uniqueness

When looking at the planner's problem for this economy, we can get some more valuable insights into the nature of this model. The first issue to resolve is how the planner should optimally choose promotion paths for the agents given that she wants to implement some given stationary allocation $n(\tau, x)$. It turns out that it is optimal that the agents' paths never cross – this means that one agent that is above another in a given vintage will stay above the other for the entire lifetime of this vintage.

Lemma 2.4 Promotion along level lines *To implement a given stationary distribution $n(\tau, x)$ it is optimal to let the agents follow the level lines of the function $N(\tau, x) = \int_x^1 n(\tau, x)dx$, the "hierarch-cdf" along the x-dimension seen from the top of the ladder.*

The statement is proven in the appendix. The intuition for the proof is that any other allocation makes the paths of some agents cross. But this can never be optimal, since the agent coming from the bottom has to make more additional effort to cross the worker on the top than the the top worker is saving by dropping down.

The planner's problem is the following in words: Given an initial allocation of experience, choose the function $n(s, t, x)$ such that the discounted integral over all future output minus all future learning costs is maximized; learning costs are computed by following taking the level lines of the hierarchy-cdf in each vintage as promotion paths.

Proposition 2.5 Equivalence of planner's problem to equilibrium *A solution to the planner's problem solves the HJB and the transport equation.*

The statement is proven in the appendix. The best intuition for the result is probably the following: Why *shouldn't* the welfare theorem hold? If we imagine promotions as a technology that produces labor inputs in $t+1$ from those in t using a certain amount of resources, it is not hard to see that this technology and the vintage technologies satisfy the assumptions of the welfare theorems and everything should go through. The mathematical proof, however, is a somewhat tedious application of a Lagrange-multiplier theorem.

Furthermore, the planner's problem has a unique solution. This can be established as follows: Suppose there were two stationary allocations $n_1(\tau, x)$ and $n_2(\tau, x)$ that achieve the maximum. Then a convex combination of these two is also feasible, since the total-population constraint is linear. In terms of human-capital accumulation cost, the combination must yield the same total cost as the two input distributions, since the costs of the two are just averaged. However, in terms of production there is a gain, since in every vintage we have a production function with convex level sets. Consider production in a vintage τ:

$$Y(\tau) = e^{-\gamma\tau}\left(\int_0^1 \Big[\lambda n_1(\tau,x) + (1-\lambda)n_2(\tau,x)\Big]^\rho dx\right)^{1/\rho}$$

for any $0 < \lambda < 1$.

Under the integral sign, for $0 < \rho < 1$, we have for every x that

$$\Big[\lambda n_1(\tau,x) + (1-\lambda)n_2(\tau,x)\Big]^\rho > \lambda n_1(\tau,x)^\rho + (1-\lambda)n_2(\tau,x)^\rho$$

since $g(z) = z^\rho$ is strictly convex for $0 < \rho < 1$. Then, integrating over all x and noting that the transform $h(z) = z^{1/\rho}$ is an increasing transform, we get the desired result.

TO ADD:

- ENTRY WAGES INCREASING IN VINTAGE AGE
- TOP WAGES DECREASING IN VINTAGE AGE
- WAGES INCREASING IN ALL VINTAGES
- LEARNING EFFORT DECREASING ALONG ALL CAREERS (PROOF OR COUNTEREXAMPLE)
- LEARNING EFFORT DECREASING IN POSITION X INSIDE VINTAGE (PROOF OR COUNTEREXAMPLE)

3 A discrete approximation technique

The following method discretizes the model into a finite number of vintages and a finite number of ladder rungs by introducing a random element to promotion. This setting is not only useful to compute an approximation to the equilibrium, but also get form some intuition about the value function, agent's paths and other objects of the continuous version.

Given an allocation with density $n(\tau, x)$ on the rectangle $(0 \leq \tau \leq T, 0 \leq x \leq 1$, construct a discrete grid as follows: Divide the vintages into S sub-intervals and the experience levels into X sub-intervals. Denote the size of these intervals by $\Delta\tau$ and by Δx.

To approximate careers of any slope and smooth things out, we make promotions for the agent stochastic: Take the value of $a(\tau_i, x_j)$ at a certain grid box (τ_i, x_j) to be the value of the function a in the center of respective box. Set the probability p_{ij} that the agent moves one box up (to x_{j+1} in vintage $\tau_i + 1$, that is) such that the expected slope of his career equals $a(\tau_i, x_j)$, but that it does not exceed one:

$$a(\tau_i, x_j) = \min\left\{p_{ij}\frac{\Delta x}{\Delta\tau}, 1\right\}$$

This means that in order to be able to replicate very steep slopes in this fashion, we need to make the slope $\Delta x/\Delta\tau$ become successively greater as k grows. I will make the following limiting argument: If we have an infinite sequence of discrete approximations as described above, choose the number of grid points as follows: $S_k = kS_0$ and $X_k = k^{3/2}X_0$ (the reason for this choice will become clear later). Now, since the number of grid points for the hierarchy grows faster than the number of grid points for vintages, the maximal possible slope $\Delta x_k/\Delta\tau_k$ will grow to infinity, so any slope at any point of the grid will be replicable from some k on.

$$n(\tau+1, x) = \big[1 - p(\tau, x)\big]n(\tau, x) + p(\tau, x-1)n(\tau, x-1)$$

Now, introduce the difference operators $\Delta_x f(\tau, x) = f(\tau, x+1) - f(\tau, x)$ and $\Delta_\tau f(\tau, x) = f(\tau+1, x) - f(\tau, x)$. Then we can re-write the above as

$$\begin{aligned}\Delta_\tau n(\tau, x) = &- \Delta_x\Big[n(\tau, x-1)p(\tau, x-1)\Big] = -n(\tau, x-1)\Delta_x p(\tau, x-1)\\ &- p(\tau, x-1)\Delta_x n(\tau, x-1) - \Delta_x n(\tau, x-1)\Delta_x p(\tau, x-1)\end{aligned}$$

Note that the last term in the second line will become very small compared to the others when we make the grid very small. So in the limit, the equation becomes equivalent to the mass-transport PDE (5).

Now, we want to find an equivalent to the value function. First, calculate production in vintage τ_i as an approximation to the continuous case; note that this variable is calculated *without* adjusting for TFP in the vintage, i.e.

$Y(\tau_i) = e^{-\gamma\tau_i}\tilde{Y}$:

$$\tilde{Y}(\tau_i) = \left[\sum \left(f(\tau_j)n(\tau_i, x_j)\right)^\rho \Delta x\right]^{1/\rho}$$

where again the function f is evaluated in the middle of the corresponding box (τ_i, x_j). The discrete counterpart for wages is

$$w(\tau_i, x_j) = \exp[-\gamma\tau_i] f_j^\rho \left(\frac{\tilde{Y}(\tau_i)}{n(\tau_i, x_j)}\right)^{1-\rho}$$

Note that this gives the wage rate *per unit of time*. If we want to calculate the counterpart to wage payments over the width of the box, we have to multiply this wage rate by $\Delta\tau$. The value function is

$$\begin{aligned} V(\tau_i, x_j) &= w(\tau_i, x_j)\Delta\tau + e^{-(\beta-\gamma)\Delta\tau} V(\tau_{i+1}, x_j) + \\ &= \max_a \left\{ \frac{c}{2} a^2 \Delta\tau + \underbrace{-a\frac{\Delta\tau}{\Delta x}}_{=p} e^{-(\beta-\gamma)\Delta\tau} \Delta_x V(\tau_{i+1}, x_j) \right\} \end{aligned}$$

Solving for the optimal policy gives us

$$a^*(\tau_i, x_j) = \frac{e^{-(\beta-\gamma)\Delta\tau}}{c} \frac{\Delta_x V(\tau_{i+1}, x_j)}{\Delta x},$$

which converges to the optimal policy (2) in the continuous case. Plugging back in, we get the Bellman equation

$$V(\tau_i, x_j) = w(\tau_i, x_j)\Delta\tau + e^{-(\beta-\gamma)\Delta\tau} V(\tau_{i+1}, x_j) + e^{-2(\beta-\gamma)\Delta\tau} \frac{1}{c} \left(\frac{\Delta_x V(\tau_{i+1}, x_j)}{\Delta x}\right)^2 \Delta\tau$$

When dividing this equation by $\Delta\tau$ and taking the limit as the boxes get very small, we get the continuous HJB (3).

I solve the system for a given rectangle with length T and height 1 as follows by how a real economy might converge to a steady state under adaptive expectations with some inertia in the agents' actions. Given a distribution of agents over all jobs and the promotion flows leading to it, we can calculate the value at all points of the grid. Given the value function, we can find the optimal human-capital-accumulation strategy in each cell and the value of entering vintages in the cells on the bottom of the career ladder.

As for the promotion efforts a, we know mix some of the optimal policies into the existing ones. As for the entry decisions, I send more mass into the starting points with higher value and less mass into those with higher value. Since wages are inversely related to the density, this algorithm drives the system towards an equilibrium if the tuning parameters are chosen right. In further work, one could try to prove that this algorithm is indeed a contraction mapping.

4 An example and preliminary results

Figure 4 shows a summary of the results for the approximation technique described in section 3 using the parameter values $\beta = 0.2$, $\gamma = 0.1$, $c = 0.5$ and $f(x) = 0.2 + \sqrt{x}$. The size of the grid is chosen to be $X = 7$ for the x-axis and $S = 30$ for the τ-axis. T –the optimal age of the oldest vintage– was chosen such that the value W of a worker just entering any of the careers was maximized.

4.1 Wage profiles across firms

As described before, wages are increasing in human capital x within each vintage. The steepness of these wage profiles decreases with vintage age. Specifically, wages for entrants ($x = 0$) are decreasing in vintage age, but wages on the top of the ladder ($x = 1$) are decreasing in vintage age. Note that this is in line with the evidence from Michelacci and Quadrini (2005) who report that fast-growing firms pay higher tenure premia than slow-growing ones and that entry wages in fast-growing firms are lower on average.

The intuition why this must be the case is the following: Agents who enter a technology at a very early stage do not have many experienced agents above them, as can be seen in the plot of the density n. Since the production function is complementary in the labor inputs, their marginal product is low – in terms of the motivation of the model, one could say that they are lacking experienced people above them who tell them what to do.

On the other hand, this means that very experienced agents are so scarce at the early stages of a technology that they command very high returns, hence the steep wage profile in the early stages of the technology. This steep wage profile induces agents to accumulate human capital very fast, as can be seen in the plot of the career lines in the top-left panel of figure 4. However,

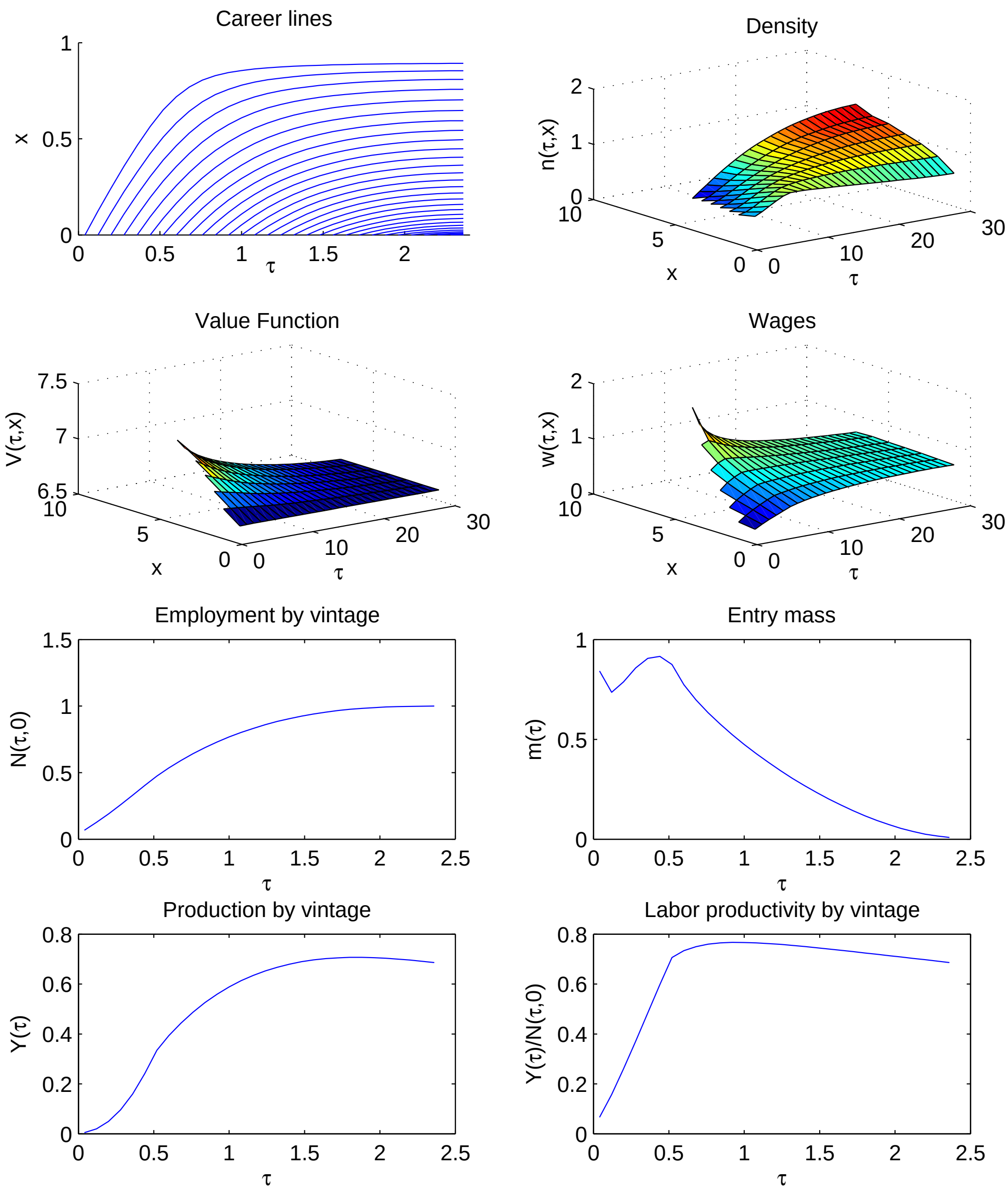


Figure 3: Equilibrium

note that this is not the only effect at work: There is also a horizon effect as is present in Ben-Porath-type models with finitely-lived agents. Early in the career incentives to accumulate human capital are especially large since the gains accrue over a longer time horizon. However, note that in this model the horizon effect is not caused by finite-livedness of *agents* but by the (endogenous) finite-livedness of *technologies*.

As vintages age, entering agents' marginal productivity rises since there are many old hands around who tell them what to do, so entry wages rise. On the other hand, more and more people press into the higher echelons of the human-capital hierarchy and depress returns there, so the tenure premium falls over the lifetime of the vintage. In fact, as stated before in theorem (?), this process continues until marginal returns are completely equalized and the wage profile is flat when the vintage dies.

4.2 Firm size and organization capital

Another very robust empirical fact is that larger firms pay higher wages [CITATION HERE!!!]. Notice that the model also predicts that larger firms pay higher average wages, if we associate a vintage with a firm or assume that the number of firms per vintage does not change over time. Average wages are identical to average labor productivity in this framework, which is plotted in the lower left corner of figure 4. The economic mechanism behind this is the trade-off between experience accumulation and obsolescence. Younger vintages have higher TFP, but experienced labor is still very scarce in these technologies. Under complementarity, this means that marginal productivities across human-capital inputs have a large spread (as is apparent from the steep wage profiles), which means that these firms are far away from the optimal input structure without experience constraints. Later on, this gap closes – marginal productivities converge and average labor productivity increases manyfold.

In fact, in the final stages of the process the gains from further experience accumulation become smaller, and the obsolescence effect takes over. The vintage looses ground compared to slightly younger vintages, as is apparent from the graph for labor productivity.

the vintage falls farther behind the frontier technology in terms of TFP, which in the end leads to it being shut down and the workers being allocated to newer technologies.

Average productivity profiles as well as the wage profiles of the early entrants display some similarity to the hump-shaped profiles that Atkeson and Kehoe (2005) measure for rents from organization capital. In their model, there is an exogenous hump-shaped productivity process which is inherited by the returns to the firm owner because costs of labor and capital inputs are time-invariant.

The vintage-human-capital model developed in this paper may be interpreted as one that endogenizes Atkeson and Kehoe (2005) exogenous process for learning: Increases in average labor productivity arise over the lifetime of a vintage which are brought about by the individual learning of workers and gains due to the successively better assignment of tasks inside the vintage. Atkeson and Kehoe (2005) say that "measuring the return to workers from organization capital is an importan task, but not one that we attempt". In the model presented in this paper, things are reversed: *All* proceeds from organization capital go to workers, but there is no plant/firm owner who reaps returns from organization capital.[9]

As for the allocation of the rents from organization capital to workers, the model says that workers entering at different points in the firm's life have the same per-period value from their careers, so the proceeds are evenly distributed in this sense. Differences occur only in the time structure of these proceeds over careers, which will be discribed in the following subsection.

4.3 Age-earnings profiles

Figure 4.3 shows the profiles of agents over their careers; the left panel shows a cross section through the economy at $t = 0$, i.e. the wages are depicted at a fixed point in time. The right panel shows the wages following an agent from $t = 0$ onward, i.e. the left-most point in each profile shows the wage at $t = 0$ and along the profile time progresses.

To understand the forces at work in these profiles, it is useful to decompose the growth of log-wages into its different components. Parameterize a career by time t, i.e. take two functions $x(t)$ and $\tau(t)$ (s.t. $d\tau/dt = 1$) and consider infinitesimal changes in log-wage along the career – note that this corresponds to the profiles plotted in the right panel of figure 4.3, since time

[9] If one interpreted the first entrants into a vintage as the firm owners, then one would obtain a similar pattern of returns for them as Atkeson and Kehoe (2005) do. However, to create a world that is really equivalent to theirs in spirit, one would have to introduce a form of vintage capital in the model, which has to be established at some cost and then reaps proceeds as the vintage ages.

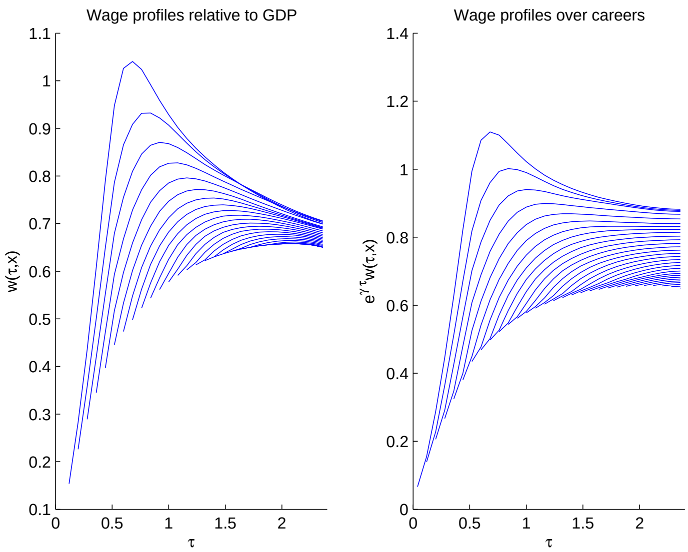


Figure 4: Equilibrium

progresses as the agent ages:

$$\left.\frac{d\ln w}{dt}\right|_{\tau(t),x(t)} = a(\tau,x)\frac{f'(x)}{f(x)} + (1-\rho)\frac{\partial \ln \tilde{Y}(\tau)}{\partial \tau} - (1-\rho)\left.\frac{d\ln n}{dt}\right|_{\tau(t),x(t)} \tag{6}$$

The first term $a(\tau,x)f'(x)$ captures returns from learning: The higher the agent's career slope $a(\tau,x)$, the higher the gains from learning. Also, the lower the agent stands in the hierarchy, the higher are the gains for a fixed effort: Since $f(x)$ was assumed to be concave, the growth rate $f'(x)/f(x)$ is decreasing in x. So since effort is seen to be decreasing over all careers in the example in figure 4, we can conclude that wage growth from this component is always positive but decreasing in this economy.

The second term involving adjusted production $\tilde{Y}(\tau)$ is increasing; it represents the gains from complementarity between labor inputs, which grow as the vintage ages.

Finally, the last term involving the density n (which has to be understood as going along the specified career path $x(t)$) is key to understanding why the wage profiles are decreasing for some careers even when following an agent over his lifetime – in the cross section, this would be less surprising since older vintages have lower productivity and a term $-\gamma$ would show up in equation (6).

Recall from equation (5) that the density along a career line decreases if and only if the career lines around it are drawing closer together (as happens in the example). Although it could not be proven yet, it seems to be true that the career lines are indeed contracting towards the right in the entire (τ, x)-space, so that the effect from encroachment on the log wage is always negative.

The economic intuition is the following: As a vintage ages and more agents enter it, human capital that was once scarce becomes now more abundant – over time, the cost of learning the skill effectively becomes lower because it is cheaper to learn gradually than to learn very quickly when a technology is new.

An interesting feature of the wage profiles generated by the model is that they have heterogenous slopes and curvature. In Ben-Porath-type models, this heterogeneity in shape is usually attained by assuming heterogenous learning ability, as done in Guvenen and Kuruscu (2006) and Huggett, Ventura, and Yaron (2006), for example. In contrast to these models, there is no underlying heterogeneity in the model presented here – the heterogeneity in the profiles is induced by an endogenous process. Workers are needed in all positions, but productivities vary widely across vintages and human-capital levels. To make workers indifferent between different careers, spells of high and low productivity have to be such that agents are indifferent between entering any career in equilibrium and do not have an incentive to leave a career. This also means that the weights on these profiles, i.e. the mass of agents choosing each profile, is determined through an endogenous mechanism ; in Ben-Porath type models, one usually has to introduce an additional parameter, the skill distribution across agents, to study heterogeneity.

Another topic from the labor literature, see for example Hause (1981), addressed by the model is "overtaking". Hause (1981) defines overtaking as the fact that two wage profiles with different slope but the same present value have to intersect at a certain point. The model has precise predictions on when this overtaking point occurs for different pairs of agents in the economy. As can be seen in figure 1 in the introduction, overtaking takes place rather early in the career for the chosen parameters. A more careful estimation of the model would be necessary to make more qualified statements on the point of overtaking, though.

4.4 Careers and human-capital accumulation

The upper-left panel in figure 4 shows that early entrants into technologies climb the occupational ladder fastest. Later entrants make less effort to climb the ladder fast. Also along the career lines the learning effort diminishes for each agent as her tenure increases. This effect was found for all parameter values considered so far in simulations but could not be proven yet. Economically, the flagging effort is due to both the decrease in the tenure premium and the horizon effect.

If one takes the model at face value, the different rungs in the firm ladder can be interpreted as the typical stations in a career, say from an executive over lower management up to the CEO. This would be in the spirit of models like the one by Gibbons and Waldman (1999), whose objective is to explain the joint dynamics of promotions and wages. The model presented in this paper is similar to theirs insofar as it models the hierarchy in the firm as a one-dimensional job ladder. Gibbons and Waldman (1999) argue that this is in line with the patterns observed in data by Baker, Gibbs, and Holmstrom (1994), who analyze wage and promotion data on management employees in one large U.S. firm.

On the positive side, the model gets the following right: There are some decreases in real wages, but they are not the norm. Also, there are no demotions, which are very rare in the data analyzed by Baker, Gibbs, and Holmstrom (1994). However, there are some serious problems with the promotion-interpretation when looking at the mass of people in the different positions. As the graph for the density n shows in figure 4, the ratio of high-x to low-x workers is low for young vintages but then increases sharply as the vintage ages. To say the least, it is hard to believe that new firms display a more pyramidal structure than older ones in the data.

Another problem with taking the model seriously in terms of career stages is that there would be more CEOs in the end of a vintage's life than there are common executives. In order for this to change, a production function closer to a Leontief-type would be necessary.

5 Conclusions

This paper has presented a continuous-time model for vintage human capital where agents endogenously decide on human-capital accumulation. In equilbrium, the wage profiles of agents over their careers vary in shape but deliver identical present value when the costs of human-capital accumulation are taken into account. Agents who enter new technology have very steep wage profiles in the beginning of their careers and experience real wage decreases later on. This happens because more agents press into their positions and make their once scarce skill more commonplace. Agents who enter older technologies have flatter wage profiles but higher entry wages. This is because they are complemented by many experienced workers who boost their marginal productivity.

For the wage structure inside firms this means that young firms pay a high premium for experience, whereas older firms have a comparatively flat wage structure. This prediction is in line with data analyzed by Michelacci and Quadrini (2005) who report that workers in young, fast-growing firms are paid lower entrance salaries and experience larger wage increases than workers in older, slow-growing firms.

In the model, average labor productivity over the life cycle has a hump-shaped pattern: It increases strongly when a technology is very young but then flattens out. This feature is reminiscent of the concept of organization capital[10]. Indeed, the model presented here may be construed as a micro-foundation of this concept. The rise in average labor productivity is driven both by the separate learning efforts of the workers in a vintage and by the rising gains coming from complementarity. The latter occur since the proportions of the different labor inputs become more favorable over time. Specifically, the shortage of highly qualified agents diminishes over time as the technology ages.

An extension that is planned in the near future is the introduction of exponential death for agents. This would not change the degree of difficulty of the equations dramatically so the fundamental results should still go through. More importantly, this feature would allow one to calibrate the model to the typical length of a work life and make the assumption of infinitely-lived agents more palatable.

Another possible extension that is beyond the reach of this paper is to introduce riskiness in human capital.[11] This addition would add the possi-

[10]as analyzed recently by Atkeson and Kehoe (2005), for example

[11]If standard Brownian processes are used, the central equations of the model –the mass-transport equation, the HJB and the wage equation– are still well-understood partial differential equations. However, second-order derivatives would show up, making the analysis somewhat more cumbersome than here.

bility of demotions to the model and possibly induce unlucky agents to quit their career before reaching higher levels in the hierarchy. Also, estimation of the model would be more credible with a true element of randomness in the model.

Finally, the model could be used to analyze a range of economic questions that are beyond the original motivation of the paper. For example, one could assess the effects of changes in the demographic structure on wage profiles and labor productivity. Also, the model has predictions on the change of wage profiles when long-term TFP growth changes due to a slower arrival of blueprints. This was arguably the case in the Japanese economy in the transition from the high-growth decades in the catch-up phase after the war to the moderate-growth decades recently. Interestingly, Mincer and Higuchi (1988) find that wage profiles were systematically steeper in Japanese firms than in U.S. firms over the post-war years; they argue that the rapid adoption of new technologies made it necessary for Japanese firms to create strong incentives for on-the-job training in this way.

References

Atkeson, A., and P. J. Kehoe (2005): "Modeling and Measuring Organization Capital," *Journal of Political Economy*, 113, 1026–1053.

Baker, G., M. Gibbs, and B. Holmstrom (1994): "The Wage Policy of a Firm," *The Quarterly Journal of Economics*, 109(4), 921–955.

Ben-Porath, Y. (1967): "The Production of Human Capital and the Life Cycle of Earnings," *The Journal of Political Economy*, (4), 352–365.

Bronars, Stephen G., and Famulari, Melissa (2001): "Shareholder Wealth and Wages: Evidence for White-Collar Workers," *The Journal of Political Economy*, (2), 328–354.

Chari, V. V., and H. Hopenhayn (1991): "Vintage Human Capital, Growth, and the Diffusion of New Technology," *Journal of Political Economy*, 99(6), 1142–65.

Gibbons, R., and M. Waldman (1999): "A Theory of Wage and Promotion Dynamics inside Firms," *The Quarterly Journal of Economics*, (4), 1321–1358.

Guvenen, F., and B. Kuruscu (2006): "Understanding Wage Inequality: Ben-Porath Meets Skill-Biased Technical Change," *Working Paper*.

Hanka, G. (1998): "Debt and the terms of employment," *Journal of Financial Economics*, 48(3), 245–282.

Hause, J. C. (1981): "The Fine Structure of Earnings and the On-the-Job Training Hypothesis," *Econometrica*, (1), 277.

Huggett, M., G. Ventura, and A. Yaron (2006): "Human capital and earnings distribution dynamics," *Journal of Monetary Economics*, 53, 265–290.

Michelacci, C., and V. Quadrini (2005): "Financial Markets and Wages," *Working Paper*.

Mincer, J. (1974): *Schooling, Experience and Earnings*. Columbia University Press, New York.

MINCER, J., AND Y. HIGUCHI (1988): "Wage structures and labor turnover in the United States and Japan," *Journal of the Japanese and International Economies*, 2(2), 97–133, available at http://ideas.repec.org/a/eee/jjieco/v2y1988i2p97-133.html.

A Proof for forward equation at $x = 1$

Also, we need an equation at the upper limit $x = 1$. Derive this by the following mass-preservation equation:

$$\int_{1-h}^{1} n(t,x)dx = \int_{1-h+a(t,x)h+o(h)}^{1} n(t+h,x)dx$$

Now, take in differentiability of $n(\cdot,\cdot)$ and $a(\cdot,\cdot)$ as well as the fact $a(t,1) = 0$ to get

$$\int_{1-h}^{1} n(t,1) - n_x(t,1)(1-x)dx + o(h)dx =$$
$$\int_{1-h-a_x(t,x)h^2+o(h^2)}^{1} n(t,1) + n_t(t,1)h - n_x(t,1)(1-x) + o(h)dx$$

Now, we can get rid of the term $hn(t,1)$ on both sides. Also, the term $\int_{1-h} n_x(t,1)(1-x)$ goes on both sides. Since $(1-x) \leq h$, the term $\int_{-a_x(t,x)h^2}^{1} n_x(1-x)dx$ is $o(h^3)$ and hence goes. Keep only $o(h^2)$-terms to get

$$n_t(t,1) = -a_x(t,1)n(t,1)$$

which makes perfect sense. Note that since $a(t,1) = 0$ implies $V_x(t,1) = 0$, and we must have $V_x x(t,1) \leq 0$ (since $V_x(t,x) \geq 0$ everywhere), which implies $a_x(t,1) \leq 0$. So the density on the top must be decreasing over time, and wages must be decreasing, even when we leave out TFP:

$$n_t(t,1) \geq 0, \qquad w_t(t,1) \leq 0$$

B Proof for planner's cost-minimizing promotion strategy

TO BE ADDED HERE!

C Derivation of first-order conditions in the planner's problem

Introduce the following notation:

$$u(s,t,x) = \frac{\partial n(t,s,x)}{\partial t}$$
$$N(t,s,x) = \int_{x}^{1} n(t,s,\tilde{x})d\tilde{x}$$

If the optimal career paths don't cross, then agents follow the quantile lines as careers. Then the local career slope $a(\tau, x)$ has to be such that the agent's career follows the iso-$\frac{\partial N(\tau,x)}{\partial \tau}(\tau, x)$-lines. To express the slope $a(\tau, x)$ in terms of the functions n and N, take a first-order approximation of N following an iso-N line: [PUT GRAPH HERE]

$$\underbrace{\frac{\partial N(\tau, x)}{\partial \tau}}_{=\dot{N}(\tau,x)} h - a(\tau, x) \underbrace{\frac{\partial N(\tau, x)}{\partial x}}_{=n(\tau,x)} h + o(h) = 0,$$

where $o(h)$ are terms with the property $\lim_{h\to 0} o(h)/h = 0$. Taking limits with respect to h we get

$$a(\tau, x) = \frac{\dot{N}(\tau, x)}{n(\tau, x)}.$$

Now, we want to aggregate the costs for the planner to move population through the grid. We want the total cost over all time to be equal to

$$C_{PL} = \int_l \int_0^\infty e^{-\beta t} C\big[a(t, s_l, x_l)\big] dt dM(l),$$

where $(s_l(t), x_l(t))$ are the policies for individual l and $M(l)$ is the measure over these individuals. Note that definitely, we can chop up the whole thing by time (i.e. change the order of integration) and then approximate:

$$\begin{aligned} C_{PL} &\sim \sum_{t_i} e^{-\beta t_i} C_{PL}(t) \\ C_{PL}(t) &\sim \sum_{\tau_j} e^{\gamma(t-\tau_j)} \sum_{x_k} \frac{c}{2} [a(t_i, \tau_j, x_k)]^2 n(t_i, \tau_j, x_k) \\ n(t_i, \tau_j, x_k) &= \int_l I(l : \tau_{j-1} \le \tau_l(t) \le \tau_j, x_{k-1} \le x_l(t) \le x_k) dM(l) \end{aligned}$$

Note that multiplying the cost of human-capital accumulation by the local mass of agents $n(\cdot)$ is necessary since we have to account for how many agents have to incur the learning cost in each position (τ, x). The above converges to

$$C_{PL} = \int e^{-\beta t} \left[\int e^{\gamma s} \left(\int_0^1 \frac{c}{2} \frac{\dot{N}(t, s, x)^2}{n(t, s, x)} dx \right) ds \right] dt$$

where the second line invokes our particular specification for the cost function.

C.0.1 Setting up the Lagrangian

Denote the initial conditions by $n(0,s,x) = n_0(s,x)$. Then the planner maximizes the following:

$$
\begin{aligned}
\max_{u(t,s,x)} \int_0^\infty & e^{-\beta t}[Y(t) - C(t)]dt = \\
& \int_0^\infty e^{-\beta t}\left[\int_0^T Y(t,s) - e^{\gamma s}\frac{c}{2}\int_0^1 \frac{\dot{N}(t,s,x)^2}{n(t,s,x)}dx\right]dt \\
\text{s.t.} \quad & n(t,s,x) = n_0(s,x) + \int_0^t u(\tilde{t},s,x)d\tilde{t} && \text{XC: } \lambda(t,s,x) \\
& \dot{N}(s,t,x) = \int_x^1 u(t,s,\tilde{x})d\tilde{x} && \text{HCC: } \eta(t,s,x) \\
& \int_0^T \int_0^1 n(t,s,x)dxds = 1 && \text{TPC: } \mu(t)
\end{aligned}
$$

where XC is for *experience*, HCC is for *human capital* and TPC is for *total-population constraint.* The Lagrange multipliers are given, too. So the Lagrangian is the following handy object:

$$
\begin{aligned}
\mathcal{L} = & \int_0^\infty e^{-\beta t}\left[\int_{t-T}^t Y(t,s) - e^{\gamma s}\frac{c}{2}\int_0^1 \frac{\dot{N}(t,s,x)^2}{n(t,s,x)}ds\right]dt- \\
& - \int_{t,s,x} e^{-\tilde{\beta}t}\lambda(t,s,x)\left[n(t,s,x) - n_0(s,x) - \int_0^t u(\tilde{t},s,x)d\tilde{t}\right]- \\
& - \int_{t,s,x} e^{-\tilde{\beta}t}\eta(t,s,x)\left[\dot{N}(t,s,x) - \int_x^1 u(t,s,\tilde{x})d\tilde{x}\right]- \\
& - \int_0^\infty e^{-\tilde{\beta}t}\mu(t)\left[\int_{t-T}^t \int_0^1 n(t,s,x)dxds - 1\right]dt,
\end{aligned}
$$

where $\tilde{\beta} = \beta - \gamma$ and we scale the Lagrange multipliers to obtain a stationary solution. The FOC are then:

$$
\frac{\partial \mathcal{L}}{\partial n(t,s,x)} = e^{-\beta t}w(t,s,x) + e^{\tilde{\beta}t}\left[\frac{c}{2}\left(\frac{\dot{N}(t,s,x)}{n(t,s,x)}\right)^2 - \lambda(t,s,x) - \mu\right] = 0 \tag{7}
$$

$$
\frac{\partial \mathcal{L}}{\partial \dot{N}(t,s,x)} = e^{-\tilde{\beta}t}\left[-c\frac{\dot{N}(t,s,x)}{n(t,s,x)} - \eta(t,s,x)\right] = 0 \tag{8}
$$

$$
\frac{\partial \mathcal{L}}{\partial u(t,s,x)} = \int_t^{s+T} e^{-\tilde{\beta}\tilde{t}}\lambda(\tilde{t},s,x)d\tilde{t} + e^{-\tilde{\beta}t}\int_0^x \eta(t,s,\tilde{x})d\tilde{x} = 0 \tag{9}
$$

At a stationary solution, we require the following:

$$
\begin{aligned}
n(t,s,x) &= n^*(\underbrace{t-s}_{=\tau},x) \\
w(t,s,x) &= e^{\gamma t} w(\tau,x) \\
\dot N(t,s,x) &= \dot N^*(\tau,x) \\
\lambda(t,s,x) &= \lambda^*(\tau,x) \\
\eta(t,s,x) &= \eta^*(\tau,x) \\
\mu(t) &= \mu^*
\end{aligned}
$$

where the last five equations follow from the first.

We first observe from (8) that $\eta(\tau,x)$ stands for the marginal cost of human-capital accumulation for the guys $n(\tau,x)$ (omit the star-superscripts for convenience):

$$\eta(\tau,x) = -c\frac{\dot N(\tau,x)}{n(\tau,x)} = -ca(\tau,x) \tag{10}$$

Second, we get from (7) a formula for $\lambda(\tau,x)$:

$$w(\tau,x) + \frac{c}{2}\left(\frac{\dot N(\tau,x)}{n(\tau,x)}\right)^2 - \mu = \lambda(\tau,x) \tag{11}$$

As for the relationship between $\lambda(\tau,x)$ and $\eta(\tau,x)$, we can re-formulate (9):

$$\int_\tau^T e^{-\tilde\beta(\tilde\tau-\tau)}\lambda(\tilde\tau,x)d\tilde\tau = -\int_0^x \eta(\tau,\tilde x)d\tilde x \tag{12}$$

When plugging the expressions for the Lagrange multipliers (10) and (11) into (12), we get

$$\int_\tau^T e^{-\tilde\beta(\tilde\tau-\tau)}\Big[w(\tilde\tau,x) + \frac{c}{2}a(\tau,x)^2 - \mu\Big]d\tilde\tau = \int_0^x ca(\tau,x)d\tilde x \tag{13}$$

Directly from this equation, we can get the following insights:

- When $\tau \to T$, the left-hand side and with it the marginal cost of education $ca(\tau,x)$ and hence education itself goes to zero. This says that one shouldn't accumulate human capital anymore just before the vintage shuts down.
- When we let $x \to 0$, the right-hand side goes to zero and we see that $\lambda(\tau,0) = 0$ for all τ. This says that at all entry jobs, the value function must be equalized, see later.
- When we combine the two, we see from equation (7) that $w(T,0) = \mu$. This says that $w(T,0)$ is the reference wage of the economy: It doesn't give any valuable experience, so it has to be just as attractive per se as any other career.

Add the value of an agent after dropping out of the career τ to the above equation to get something that resembles the HJB in the private economy more:

$$\underbrace{\int_{\tau}^{T} e^{-\tilde{\beta}(\tilde{\tau}-\tau)}[w(\tilde{\tau},x)+ca(\tilde{\tau},x)^2/2]d\tilde{\tau}+\int_{T}^{\infty} e^{-\tilde{\beta}t}\mu dt}_{\equiv\Lambda(\tau,x)}=-\int_0^x \eta(\tau,\tilde{x})d\tilde{x}+\frac{\mu}{\tilde{\beta}} \tag{14}$$

First, take the τ-derivative of $\Lambda(\tau,x)$ to get something that looks like a Bellman equation:

$$\begin{aligned} -\frac{\partial\Lambda(\tau,x)}{\partial\tau} &= w(\tau,x)+\frac{c}{2}a(\tau,x)^2-\tilde{\beta}\Lambda(\tau,x) \\ \Lambda(T,x) &= \mu/\tilde{\beta}=W \end{aligned} \tag{15}$$

Note that we still need to substitute in the optimal policy $a(\tau,x)$, which is related to the slope of the value function. Equation (14) tells us that η is linked to the x-derivative of $\Lambda(\tau,x)$ – differentiate (14) with respect to x to get

$$\frac{\partial\Lambda(\tau,x)}{\partial x}=-\eta(\tau,x)$$

Optimal learning, in turn, is given by equation (10). So we get

$$a(\tau,x)=\frac{-\eta(\tau,x)}{c}=\frac{1}{c}\frac{\partial\Lambda(\tau,x)}{\partial x}$$

So now we plug this back into (15) to get

$$-\frac{\partial\Lambda(\tau,x)}{\partial\tau}=w(\tau,x)+\frac{1}{2c}\left(\frac{\partial\Lambda(\tau,x)}{\partial x}\right)^2+\tilde{\beta}\Lambda(\tau,x)$$

So this is equivalence to the planner's problem.