Vintage Human Capital and Learning Curves*

Matthias Kredler†

1 December 2011

Abstract

I study a vintage-human-capital model in which long-lived workers accumulate human capital following an exogenous learning curve. Different skill levels inside a vintage are complementary in production. The predictions are qualitatively in line with those in a comparable model with endogenous human-capital accumulation, see Kredler (2010). The premium on skill shrinks and eventually vanishes as vintages age. Workers experience wage cuts when relocating to a new technology. This shows that it is the scarcity of skill across technologies and not endogenous human-capital accumulation that determines the wage structure when human capital is technology-specific.

Keywords: Vintage human capital, tenure-wage profiles, learning curves
JEL codes: J01, E24

*This paper is part of my dissertation thesis at New York University. I thank Boyan Jovanovic, Jess Benhabib, Greg Kaplan and Silvana Melitsko for discussion and comments.

†Departamento de Economía, Universidad Carlos III de Madrid, matthias.kredler@uc3m.es. Part of the research was funded by the Spanish Ministerio de Ciencia e Innovación, reference number SEJ2007-62908.

1 Introduction

Vintage-human-capital models assume that skill is specific to a technology. A real-life example of technology-specific capital is an architect with drawing skills who sees his skill become obsolete as computer programs take over the design process. Chari & Hopenhayn (1991) are the first to study the theoretical implications of vintage-specific human capital. They do so in a two-period overlapping-generations model in which experienced and inexperienced workers are complementary in production.

Kredler (2010) extends this basic setting. The key difference to Chari & Hopenhayn (1991) is the introduction of endogenous human-capital accumulation à la Ben-Porath (1967). Furthermore, Kredler's (2010) setting allows for long-lived workers and a continuum of skills within a vintage. These novelties gives rise to rich predictions on earnings profiles over workers' careers, which are shown to be qualitatively in line with evidence from employer-employee matched data.

Many of the results in Kredler (2010) are, at least partially, driven by endogenous human-capital accumulation. Entrants into young technologies experience faster wage growth then entrants into old technologies. One important reason for this is that they invest more into human capital. Furthermore, the premium on skill is shown to decline as the technology ages. This occurs as later entrants into the technology catch up with earlier entrants' skills by investing in human capital.

The current paper shows that these and other results continue to hold when replacing endogenous human-capital accumulation by an exogenous learning curve. This tells us that the assumptions made on human-capital accumulation do not matter much. As will become clear, the common force that is present in both models and drives the results is the scarcity of skill across vintages.

Specifically, the current paper shows that the premium on skill shrinks and eventually vanishes entirely as the technology ages, a process that Kredler (2010) refers to as *wage-compression*. This tells us that the following effect is sufficient for wage compression: Skill becomes abundant as a technology ages, so the premium on it decreases.

Second, the current paper shows that endogenous skill accumulation is not necessary to generate what Hause (1981) describes as *overtaking*: Exante identical workers must see their earnings profiles cross at some point if they embark upon different career paths.[1] In the model, early entrants into

[1]In a deterministic setting, this must be the case if the two careers are to deliver the

a technology are paid low wages initially since the technology's productivity is still low. They then experience fast wage growth as the technology's productivity rises and they deliver a relatively scarce input to production. Late entrants into the technology acquire skills that will be obsolete soon; they must be paid higher entry wages to make their career attractive enough. This makes their wage profiles flatter. So again, relative scarcity of skill and not differences in human-capital investment are key for the wage structure.

In one respect, the current paper is able to improve upon the results by Kredler (2010). It presents a formal a proof for increasingness of wages in skill.[2] As for the development of technologies, the model predicts that labor productivity rises as they add more inputs (the different skill levels). However, these gains taper off over time. As the vintage ages, its total factor productivity falls farther and farther behind the frontier. Eventually the vintage is shut down in order to re-allocate workers to more productive technologies. Workers are shown to experience wage losses when this occurs.

The wage structure in Chari & Hopenhayn's (1991) setting shares many properties with the one in this paper, such as the declining skill premium in older vintages and overtaking. However, workers in Chari & Hopenhayn (1991) only live for two periods. This means that the predictions on wage profiles are not as rich as here. Another crucial difference is that workers never suffer wage losses in Chari & Hopenhayn (1991).

I now turn to a brief review of other related papers. Violante (2002) studies a learning-curve specification in a vintage setting, but in his paper workers of different skill levels are not complementary in production. This rules out heterogeneity in earnings profiles as discussed above. Violante (2002) assumes human capital to be specific to machines of different vintages. Another key difference is that the matching of workers to machines is subject to search frictions.

Parente (1994) also assumes a learning curve, but embeds it into a ladder model of technology adoption. Unlike in vintage models, all technologies are available to firms at all time. However, firms lose expertise when they adopt new technologies, and these losses are increasing in the technological distance between the abandoned and the adopted technology. This gives rise to a trade-off in the switching decision of firms.

On the technical side, the current paper exploits a continuous-time formulation to study the timing decision of a technology phase-out differen-

same present value of earnings.

[2] Wages are shown to be increasing in skill in Chari & Hopenhayn (1991), but there are only two skill levels.

tially. This leads to a sharp characterization for wages and the distribution of workers in the dying vintage. In the planner's problem, I draw heavily on the Lagrange-multiplier theorem for infinite-dimensional spaces. This allows me to derive properties of the economy's wage structure in a large state space.

The remainder of the paper is organized as follows: Section 2 presents the environment. Section 3 defines and characterizes the competitive equilibrium. Section 4 derives further results from the vantage point of the planner's problem and shows that the solution to the planner's problem constitutes a competitive equilibrium. Section 5 illustrates the results in a numerical example and section 6 concludes.

2 Setup

2.1 Technology

Time t 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 call the vintages by their date of inception, s, or – especially in a stationary setting – identify them with their age $\tau \equiv t - s$.

As inputs, the production technology of age τ uses labor inputs which are differentiated by experience levels $0 \leq h \leq \tau$. For example, workers with experience level $h = 0$ have just joined the technology in this instant; workers with $h = \tau$ possess the maximum possible experience level in the vintage of age τ – they are "founding members" of their technology. In section 2.3, I will exactly specify how experience is accumulated.

It is convenient to introduce another variable which denotes the hierarchy of experience levels within a vintage: Complementary to h, define $z \equiv \tau - h$. This variable describes how far a worker is away from the maximum experience level in his technology. This notation is often more convenient, since z is constant for a worker who makes full use of his experience in his career. Figure 1 illustrates the connection between the different variables.

I assume that the production function for vintage s at time t takes the following functional form:

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

where $\gamma > 0$ and $\rho < 1$. $n(t,s,h)$ denotes the density of workers with experience h in vintage s at time t. $f(\cdot)$ is a learning curve with the usual characteristics. It is assumed to be continuously differentiable, non-decreasing

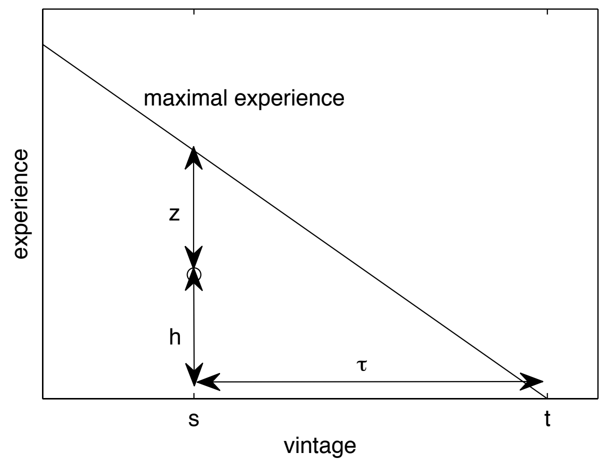

Figure 1: Vintages and experience levels

and weakly log-concave, i.e. $d \ln f(h)/dh \leq h$. Log-concavity is a weaker requirement than concavity; it requires the percentage increments to f to be decreasing in h, whereas concavity requires the absolute increments to be decreasing in h.

Note that both experience creation and the addition of new input factors contribute to productivity growth in a vintage, but that total factor productivity declines exponentially when compared to the frontier technology $t = s$. It is essential for the results in this paper that the gains from learning and factor addition in old vintages are ultimately dominated by productivity growth in new vintages.

Total output in the economy at time t is

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

2.2 Agents and preferences

The economy is populated by a measure 1 of agents. Individuals have linear preferences and live forever:

$$\int_{0}^{\infty} e^{-\beta t} c(t)dt,$$

where $\beta > \gamma$ and $c(t)$ is consumption at time t. A worker enters the model in $t = 0$ with an endowment of experience h_0 in a vintage $s_0 \leq 0$, where

$h_0 \leq -s_0$. The measure of workers ν over the triangle $0 \leq s \leq \infty, 0 \leq h \leq -s$ at $t = 0$ is exogenously given. It tells us about the initial distribution of experience. The distribution of workers over the (s, h)-space is the economy's state.

2.3 Experience accumulation

Workers can choose both the vintage in which they want to work and the promotion paths they want to follow within this vintage. It is useful to first define the concept of a *career* within one vintage and then turn to the sequencing of careers, which will constitute the economic *life* of an agent. A life will consist of a sequence of careers indexed by $n = 0, 1, \dots$. An agent who makes full use of her experience climbs the h-latter at normalized speed 1, i.e. his z-level stays constant. However, agents are also allowed to remain at tasks they already master and climb the task latter slower or even drop down in the task latter.

Definition 2.1. A *career* (indexed by $n \in \{0, 1, \dots\}$) is a collection of

- a vintage $s_n \in (-\infty, \infty)$
- an entry time $t_n \geq s_n$
- an exit time $t_{n+1} > t_n$, where t_{n+1} may be infinity
- a promotion path, which is a measurable function $z_n(t) : [t_n, t_{n+1}) \to \mathbb{R}$ satisfying:
  - start without experience: $z_n(t_n) = t_n - s_n$
  - promotion constraint: $z_n(\cdot)$ is a non-increasing function

This definition has to be modified in an obvious way for careers starting at $t_0 = 0$, where the initial endowment of the agent comes into play: The starting point for a worker endowed with experience h_0 in vintage s_0 may be any experience level $h \leq h_0$ in vintage s_0, but has to be $h = 0$ for any vintage $s \neq s_0$.

Figure 2.3 shows possible careers. The entry points into the careers are fixed to time $t_0 = 0$. A worker chooses a vintage s and enters with experience $h = 0$ (or, in hierarchy notation, $z = -s$ away from the top level of experience in this vintage) in $t = 0$. The graph only shows careers for workers who do not possess any initial endowment (i.e. $h_0 = 0$) in order to avoid confusion about the careers' starting points. As time progresses and the vintage ages, the workers move rightward in the graph at a speed equal to unity. They may or may not choose to use the entire experience available to them at any point of the career.

Career A starts in a rather young vintage. The worker pursuing career A stays at the same z for all t, which means that he is constantly being "promoted" – he uses all the experience available to him at all points in time. Career B differs from A by the end point t_1: The worker pursuing B stays in his vintage until $t = 100 - 20$, whereas worker A already quits his vintage at an earlier stage. Careers C and D display two possible forms of "demotion careers": C drops at once by various rungs in the career ladder at $t = 75 - 30$, whereas D is constantly underachieving with respect to the full-promotion career (which would be a horizontal line as in B).

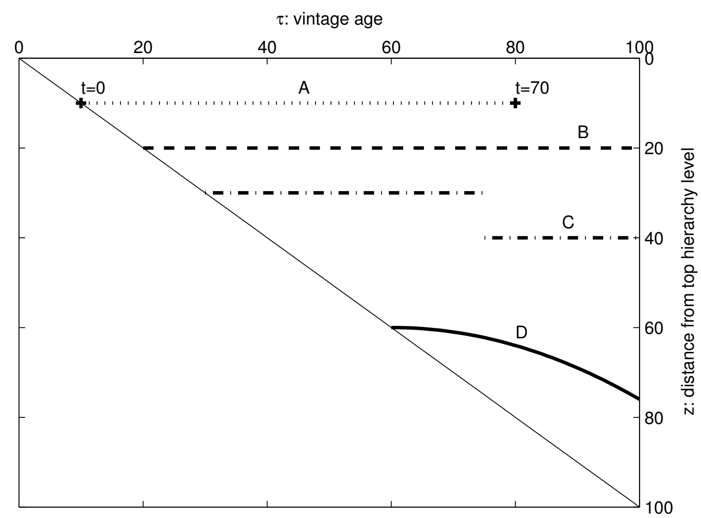

Figure 2: Some possible careers for workers entering at $t = 0$

A worker can concatenate careers in an almost arbitrary fashion over the course of his life. The only requirement made comes in form of the following definition:

Definition 2.2. The *life* of an agent is a sequence of subsequent careers. Precisely, a life is a function $l(t) = [s(t), z(t)]$, $l : [0, \infty) \to \mathbb{R}^2$, mapping any instant in time t to a vintage $s \leq t$ and an experience level z that is feasible given the worker's past. The function $s(\cdot)$ is a step function that jumps at the points $\{t_n\}$. $\{t_n\}$ is a (possibly infinite) sequence of switching points fulfilling $0 = t_0 < t_1 < t_2 < \dots$.

A worker is not allowed to be idle at any point during his life. Since workers do not value leisure, this would not occur in equilibrium anyway.

3 Competitive equilibrium

In this section, I define a competitive equilibrium for the economy described in section 2. First, I analyze the firm's problem, which is static. I then turn to the worker's problem, which incorporates the dynamic aspect in this economy. The section concludes with the definition of competitive equilibrium and a characterization for the support of vintages.

3.1 The firm's problem

Firms take the wages for all labor inputs as given in any instant. Hence, maximization of profits over any time horizon is equivalent to maximizing profits at each instant. The profits for a firm operating in vintage s at time t are

$$\pi(t,s) = \max_{n(t,s,\cdot)} \left\{ e^{\gamma s} \left(\int_0^{t-s} [f(h)n(t,s,h)]^\rho dh \right)^{1/\rho} - \right.$$
$$\left. - \int_0^{t-s} w(t,s,h)n(t,s,h)dh \right\},$$

where

- $n(t,s,h)$ is experience-h labor demanded by a vintage-s firm at t and
- $w(t,s,h)$ is the market wage for experience-h labor in vintage s at t.

Since the production technology is constant-returns-to-scale, profits $\pi(t,s)$ will be zero for any pair (t,s) in equilibrium. Since there are no externalities across vintages or across time, the structure of firms is indeterminate: Firms could operate in a single vintage s or spread their operations across more than one vintage.

Given the distribution of workers $n(t,s,h)$, wages will be given by the marginal productivity of workers in equilibrium:

$$w(t,s,h) = e^{\gamma s} f(h)^\rho \left(\frac{\left[\int_0^{t-s} f(\tilde{h})n(t,s,\tilde{h})^\rho d\tilde{h} \right]^{1/\rho}}{n(t,s,h)} \right)^{1-\rho}.$$

3.2 The worker's problem

Given a wage function $w(t,s,h)$ and the initial experience h_0 in vintage s_0, the worker has to choose a life l that maximizes her utility. Denote by $l(s_0,h_0)$ a feasible life for a worker with initial endowment (s_0,h_0) and by $L(s_0,h_0)$ the space of all possible $l(s_0,h_0)$ for this agent. Define the value of a life l for an agent as

$$v\big(l(s_0,h_0)\big) = \int_0^\infty e^{-\beta t} w\big(t, s(t), h(t)\big) ds.$$

The least upper bound for this value gives us the value function and is denoted by

$$V_0(s_0,h_0) = \sup_{l \in L(s_0,h_0)} v\big(l(s_0,h_0)\big). \tag{1}$$

3.3 Dynamic competitive equilibrium

Before introducing the equilibrium concept, it is useful to first define the space of lives for all agents: Let $L = \bigcup_{(s_0,h_0)} L(s_0,h_0)$. We will need a measure on the space L for the definition of an equilibrium. Define the following σ-algebra on L: Let $\mathcal{B}$ be the σ-algebra generated by sets of the form $\{l : l(t_1) \in B_1, \ldots, l(t_n) \in B_n\}$, where $0 \leq t_1 < t_2 < \ldots < t_n$ and $\{B_k\}_{k=1}^n$ are rectangles in $\mathbb{R}^2$.

Definition 3.1. A *dynamic competitive equilibrium* is a collection of

- a measure μ on the space $(L, \mathcal{B})$ of all possible lives
- a worker density $n(t,s,h)$
- a wage function $w(t,s,h)$

which fulfills the following conditions:

- Optimality in production: $n(t,s,\cdot)$ maximizes

  $$\pi(t,s) = e^{\gamma s} \left(\int_0^{t-s} [f(h) n(t,s,h)]^\rho dh \right)^{1/\rho} - \int_0^{t-s} n(t,s,h) w(t,s,h) dh$$

  given wages $w(t,s,h)$ for all pairs (t,s).
- Only optimal careers in μ: Any set $A \subset L$ fulfilling $l(s_0,x_0) \in A \Rightarrow v(l(s_0,x_0)) < V_0(s_0,n_0)$ has measure zero.

- Normalization: $\int_L d\mu(l) = 1$.

- μ yields $n(t, s, h)$: For every rectangle $B \subset \mathbb{R}^2$, for all $0 \leq t < \infty$, we have
$$\int_{\{l:l(t)\in B\}} d\mu(l) = \int_B n(t, s, h) dh ds.$$

- Consistency with initial conditions: For every rectangle $B \subset \mathbb{R}^2$ of the form $s \in [s_l, s_u]$, $h \in [h_l, -s]$, we have
$$\mu(\{l : l(0) \in B\}) \leq \nu(B).$$

The last condition ensures that we never use more experience then prescribed by the initial measure ν at $t = 0$. Note that workers can use less than their initial experience level h_0 in vintage s_0, but never more.

Note the this equilibrium definition automatically precludes any measure μ giving rise to mass points at any point (t, s, h) since we posit the existence of a density function $n(\cdot)$. Since I assumed complementarity of inputs in the production technology this is unproblematic. Concentrating too many workers in one area is automatically discouraged since it drives marginal returns in this region to zero by the Inada condition in the production function.

It is worthwhile to point out that the equilibrium definition requires wages to be specified also for regions of the (t, s, h)-space where n is zero. In such regions, the wage schedule must be such that firms optimally set labor demand to zero and workers' optimal decisions make labor supply zero. This will be relevant for old vintages that have been phased out.

3.4 Stationary competitive equilibrium

I now turn to the definition of a stationary equilibrium. The natural requirement for stationarity in this economy is the following: For the vintage of age τ, there is always the same number of workers occupying any fixed rung z in the skill ladder. In the context of a stationary economy, it turns out that the variables $\tau = t - s$ and $z = \tau - h$ have advantages over their counterparts s and h, so – with a slight abuse of notation – $n(t, \tau, z)$ will be used in the following to denote the density of workers at (t, τ, z).

Definition 3.2. A *stationary competitive equilibrium* is a dynamic competitive equilibrium that satisfies:
$$n(t, \tau, z) = n(0, \tau, z) \quad \text{for all } t > 0 \text{ given any } \tau \geq 0,\ z \in [0, \tau].$$

The definition immediately implies that wages grow at rate γ for a fixed location (τ, z), that production in the vintage of age τ grows at rate γ and finally that total output in the economy grows at rate γ:

$$w(t,\tau,z) = e^{\gamma t} w(0,\tau,z),$$
$$Y(t,\tau) = e^{\gamma t} Y(0,\tau),$$
$$\bar{Y}(t) = e^{\gamma t} \bar{Y}(0).$$

3.5 The relationship between μ and n

It is clear by the definition of competitive equilibrium that any measure $\mu(l)$ over lives maps to a unique distribution function $n(t,\tau,z)$. However, it is not quite clear what the opposite relationship looks like. Given an arbitrary distribution $n(\cdot)$, does there always exist a measure μ over feasible lives yielding $n(\cdot)$? The answer to this question is no: The density might require that some workers accumulate experience faster than they are allowed to. This leads to the next question: What are the requirements on $n(\cdot)$ such that there exists a permissible μ inducing n? Is this measure μ unique? It will be important to know the answers to these questions once we move on to the planner's problem for this economy since it is easier to let the planner choose a density than a measure over lives.

To clarify the relationship between n and μ, notice first that it is in general possible that two different measures yield the same distribution. This is not surprising if one regards the lives of workers as a stochastic process on $\mathbb{R}^2$ for $0 \leq t < \infty$. In the language of stochastic processes, there exists more than one stochastic process that yields the same marginal distributions across $t \in [0,\infty)$. An example may be in order to illustrate this point.

Example 3.1. Define the measures μ_1 and μ_2 as follows: For $0 \leq t < 1$, let both measures spread the unit mass uniformly over the square $0 \leq \tau \leq 1$ and $0 \leq z \leq 1$ (assume that the initial distribution of experience is such that this is feasible), and uniformly over the square $1 \leq \tau \leq 2$, $1 \leq z \leq 2$ for all $s > 1$. Within the time intervals $[0,1)$ and $[1,\infty)$, let every worker be stuck at one position z in the skill hierarchy of his vintage. For the economy, this distribution means that production in $t \in [0,2]$ only takes place in the vintages that are from zero to one year old at $t = 0$ and that the experience created in $t < 1$ is thrown overboard at $t = 1$ in order to re-locate everybody to some lower skill level.

To engineer the set of demotions at $s = 1$, we will let all workers stay in the vintage they worked in, but re-shuffle their positions within their vintage in different ways: μ_1 specifies that the hierarchy of workers within the company stays the same, i.e. $z(1) = z(0) + 1$ for all careers. In contrast, μ_2 engineers a revolution at $s = 1$: Bosses drop to the lowest level of the firm hierarchy, and former handymen take the highest positions after the revolution. Mathematically, specify $z(1) = 2 - z(0)$ for all careers.

Finally, it is possible to create a "stochastic" measure μ_3: Instead of making the position in the firm at $t = 1$ dependent on the position before $t = 1$, re-shuffle positions randomly: make $z(0)$ a random variable independent of $z(1)$. In this measure, knowing the position of a worker before $t = 1$ does not confer any information about his position after $t = 1$, whereas μ_1 and μ_2 make the entire future career predictable when knowing the position at $t = 0$.

To conclude this example, observe that all three measures yield the same marginal distributions $n(t, \cdot)$ for all t, but that the careers constituting the measures are by no means the same.

The distribution function $n(\cdot)$ in the preceding example was of an especially simple form. But what happens if we take more elaborate functions $n(\cdot)$? It turns out that the consistency requirement is that the number of workers above a certain z-level in a given vintage s does not increase over time.

Definition 3.3 (Consistency of n)**.** We call a density $n(\cdot)$ *consistent* if the function $N(t, s, z) \equiv \int_0^z n(t, s, \tilde{z}) d\tilde{z}$ is non-increasing in t for all $t \geq 0$, $s \leq t$, $z \in [0, t - s]$.

The following result tells us that consistency is necessary and sufficient to create a measure μ that yields n. The proof, which is given in appendix A.1, shows that μ may be constructed by requiring that workers' career paths never cross.

Proposition 3.1. (Construction of the no-crossing measure) *Consider an arbitrary density function n for workers satisfying $\int_{s,z} n(t, s, z) = 1$ for all t. There exists a measure μ on $(L, \mathcal{B})$ such that μ yields $n(t, \tau, z)$ in the sense of definition 3.1 if and only if n is consistent. This measure may be chosen such that workers' careers follow the level lines of $N(t, s, z)$ in definition 3.3, which prevents workers' career paths from crossing each other.*

Notice that the no-crossing measure is non-stochastic in the sense of example 3.1 – knowing the position of a worker at one point in time confers complete information about his future and past work life. It is also worth noting that there would have been a potential degree of freedom in the construction of the measure in regions where the level lines of N are strictly decreasing in t. At these points, it would have been possible to let the careers of some agents cross in the style of example 3.1. However, in regions where N is constant in t, no crossing can be allowed for sets of lives with positive measure.

We can say that a competitive equilibrium characterized by a no-crossing measure μ_{nc} is equivalent to another competitive equilibrium characterized by μ' if they yield same distribution n. This equivalence is in the following sense: Wages are the same for all locations (t, s, z) in both economies since the distribution n is the same. Then, every worker with the same initial state (s_0, z_0) obtains the same value $V_0(s_0, z_0)$ under μ_{nc} and under μ', since both must reach the supremum $V_0(s_0, n_0)$ by the definition of equilibrium. Note that V_0 only depends on wages and not on other agents' lives. Hence none of the workers would mind to be transferred from the μ'- to the μ_{nc}-universe at $t = 0$. Of course there are no differences between the two universes for firms either since wages are identical. So from now on, we will focus on the no-crossing measure when considering a given density n.

3.6 Characterization of the competitive equilibrium

I will restrict attention to stationary equilibria. The first question that arises is if vintages will die at some point. Unlike in the economies considered by Chari & Hopenhayn (1991) and Kredler (2010), it is not obvious that old vintages must fall back in productivity with respect to the frontier vintage in this framework. After all, the number of inputs to production increases without bound as the vintage ages, which contributes to productivity growth. Also, returns to learning are allowed to be unbounded. However, the following result shows that these forces cannot keep up with TFP increases of the frontier vintage and that vintages are phased out eventually. This is true at least for the CES production function that I assumed.

Lemma 3.2. (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.: $Y(t, \tau) = 0$ for all $\tau > T$.*

In a nutshell, the proof (given in appendix A.2) uses the following idea: When we integrate up the value for all careers beyond some old vintage T, then this value has to be delivered by wages paid in technologies $\tau > T$. However, the maximum attainable output in a vintage becomes infinitely small compared to the output levels in the newest technologies for large τ.

This is because the exponential growth in total factor productivity dominates the returns to experience and factor accumulation eventually. It turns out that old vintages can never deliver the value that agents would obtain when entering the newest vintage. Interestingly, the log-concavity assumption on the learning curve does not have to be invoked to prove the result.

Other properties of the equilibrium are better understood from the vantage point of the planner's problem, which is introduced in the following section.

4 The planner's problem

4.1 Statement of the problem

By the construction of the no-crossing measure (proposition 3.1), we know that any consistent stationary distribution $n(\cdot)$ can be implemented by some measure μ on $(L, \mathcal{B})$, and that any measure yielding the same distribution $n(\cdot)$ gives the same utility to each agent as μ does. We can therefore let the social planner choose the function $n(t, s, z)$ instead of a measure μ. It will turn out that for the planner's problem, it is most convenient to denote vintages by the date of inception s and the hierarchy positions top-down by z, thus the repeated abuse of notation which the reader is asked to forgive.

It is important to leave the planner the possibility of deviating from the initial density $n_0(s, z)$ when maximizing. I thus return to writing the density in all three arguments here: $n(t, s, z)$. Stationarity will only be imposed after deriving the first-order conditions (FOCs) of the planner's problem. If we can find a stationary density that obeys the FOCs, then this gives us initial conditions for the economy that the planner does not want to deviate from even when given the possibility to do so.

As the construction of the no-crossing measure shows, the consistency requirement for $n(t, s, z)$ to be implementable is that the function $N(t, s, z) = \int_0^z n(t, s, \tilde{z}) d\tilde{z}$ be non-increasing in t at any point (s, z). This constraint can be enforced in an elegant way if we require $n(t, s, z)$ to be differentiable in t everywhere. Also, it turns out that the stronger requirement that $n(\cdot)$ be continuously differentiable in *all* arguments allows us to make full use of the thrust provided by the differential characterizations yielded by the Lagrange-multiplier theorem. Thus, assume from now on that $n(t, s, z)$ is continuously differentiable in s, t and z. Furthermore, denote the time increase in workers on ladder rung z in vintage s at time t by

$$u(t,s,z) \equiv \frac{\partial n(t,s,z)}{\partial t}.$$

We can then re-write the density $n(t,s,z)$ as the sum of workers who joined the ladder rung z at its inception and the additions over time:

$$n(t,s,z) = n(s+z,s,z) + \int_{s+z}^{t} u(\tilde{t},s,z)d\tilde{t}.$$

For ease of exposition, we denote the initial number of workers on a given rung z as

$$n_0(t,z) = n(s+z,s,z).$$

We now have everything in place to state the planner's problem. In this economy, welfare optimization amounts to the maximization of total discounted production in the economy since the agents' preferences are identical and linear. Arguing along the lines of lemma 3.2, it can be shown that it is never optimal for the planner to keep a vintage in production forever. In a stationary equilibrium, obviously, the upper bound T^* on the age of a technology must be the same for all vintages. Therefore, the problem will be restricted to choosing a function $n(t,s,z)$ over the triangle $0 \leq \tau \leq T, 0 \leq z \leq \tau$ for a fixed T. We will later turn to the problem of finding the T^* that is optimal for the planner.

Given T, the planner's problem is:

$$\max_{n_0 \in \mathcal{N}_0, u \in \mathcal{U}} = \int_0^\infty e^{-\beta t} \int_{t-T}^{t} e^{\gamma s} \left[\int_0^{t-s} f(t-s-z)^\rho n(t,s,z)^\rho dz \right]^{1/\rho} ds dt \tag{2}$$

$$\text{s.t.} \quad \int_{t-T}^{T} \int_0^{t-s} n(t,s,z) dz ds \leq 1 \qquad \text{for all } t \tag{3}$$

$$\int_0^z u(t,s,\tilde{z})d\tilde{z} \leq 0 \qquad \text{for all } s,t,z \tag{4}$$

$$\text{given} \quad n(0,s,z) \qquad \text{for } s \in [-T,0] \text{ and } z \in [0,-s]$$

$$\text{where} \quad n(t,s,z) = \begin{cases} n(0,s,z) + \int_0^s u(\tilde{t},s,z)d\tilde{t} & \text{if still dependent on } n(0,\cdot) \\ n_0(s,z) + \int_{s+z}^t u(\tilde{t},s,z)d\tilde{t} & \text{otherwise} \end{cases}$$

where $\mathcal{N}_0$ denotes the space of continuously differentiable functions on $[-T,\infty)\times[0,T]$ and $\mathcal{U}$ is the space of continuously differentiable functions on the set

$\{(t,s,z) : t \in [0,\infty), s \in [t-T, t], z \in [0, t-s]\}$. We may call (3) the *total-population constraint* and (4) the *experience constraint* for the problem. The experience constraint is nothing but a re-statement of the consistency requirement for densities from proposition 3.1 in terms of the function $u(\cdot)$.

Note that the problem presented in (2) to (4) is particularly well-behaved: The constraints are linear in n_0 and u, and it can easily be shown that the objective function is concave in n_0 and u. This means that the FOCs are necessary and sufficient.

The Lagrangian for this problem is

$$\begin{aligned}\mathcal{L} = &\int_0^\infty e^{(\gamma-\beta)t} \int_{t-T}^{t} \left[\int_0^{t-s} f(t-s-z)^\rho n(t,s,z)^\rho dz\right]^{1/\rho} dsdt - \\ &- \int_0^\infty \mu(t) \left[\int_{t-T}^{t}\int_0^{t-s} n(t,s,z) dsdz\right] - \\ &- \int_0^\infty \int_{t-T}^{t}\int_0^{t-s} \lambda(t,s,z)\left[\int_0^z u(t,s,\tilde{z})d\tilde{z}\right] dzdsdt.\end{aligned} \tag{5}$$

4.2 Characterization of the planner's optimum

We now turn to taking first-order conditions and characterizing the solution of the planner's problem for a given T.

As for $n_0(s,z)$, notice that a change in the starting level of workers in a career (s,z) affects production at all periods of this career until the demise of the vintage at $s+T$. On the other hand, it also weighs in on all total-population constraints over this time interval, so that the first-order condition is (see appendix A.3 for a formal derivation)

$$\int_{s+z}^{s+T} e^{-\gamma t} w(t,s,z)dt = \int_{s+z}^{s+T} \mu(t)dt \qquad \text{for all } s \in [-T,\infty], z \in [0,T], \tag{6}$$

where $w(t,s,z)$ is the marginal product of $n(t,s,z)$ in the production at time t in vintage s. Equation (6) says that the discounted wage payments over any full career without demotions must equal the value of a marginal worker $\mu(t)$ to the planner integrated over the duration of the career. Furthermore, it implies that concatenating any combination of no-demotion careers from time t to infinity must yield the same value $\int_t^\infty \mu(\tilde{t})d\tilde{t}$, whatever this concatenation looks like. This relates to the fact that in competitive equilibrium workers must be indifferent between the different possible careers at t. If the value of careers differed, there would be no entry into careers offering inferior values. This also implies that all agents are alike once the initial vintages have disappeared since agents have by then lost any advantage stemming from initial endowments.

More properties of the solution can be inferred when we observe what happens to equation (6) towards the end of a vintage's life. When driving $z \to T$ in (6), i.e. when looking at the workers who enter the vintage shortly before its death, we obtain $w(s+T, s, T) = \mu(t)$. Since workers who enter a technology just before it dies do not relax any experience constraints, their marginal productivity must be exactly equal to $\mu(t)$. This tells us that the wage of an inexperienced worker in the dying vintage is an important benchmark for the economy: It reveals the marginal value of an additional untrained worker to the social planner.

We now turn to the FOC for $u(t, s, z)$. Just as $n_0(s, z)$, an increase in $u(t, s, z)$ at t will increase production in vintage s until the vintage dies at $s + T$. As for the constraints, a higher $u(t, s, z)$ increases the left-hand side of all experience constraints for $\tilde{z} \in [s - T, z]$. Furthermore, it weighs in on the total-population constraints until the demise of vintage s at $s + T$. Mathematically, this amounts to

$$\int_t^{s+T} e^{-\beta\tilde{t}} w(\tilde{t}, s, z) d\tilde{t} = \int_t^{s+T} \mu(\tilde{t}) d\tilde{t} + \int_z^{t-s} \lambda(t, s, \tilde{z}) d\tilde{z} \qquad (7)$$

$$\text{for all } t \in [0, \infty), s \in [t - T, t], z \in [0, t - s].$$

By non-negativity of the Lagrange multipliers $\lambda(\cdot)$, equation (7) implies that the value of discounted wages at any point of a career is decreasing in z; so workers higher up in the skill hierarchy of a given vintage are better off. It also means that "promotion weakly dominates demotion". Workers will never strictly prefer to move down in the hierarchy (they might be indifferent, however).

When adding wages of the worker after the demise of the vintage to both sides of equation (7), we can construct the value function $V(t, s, z)$ for workers in the competitive framework (note that we do not have to worry about potential demotions by the argument above about promotion dominating demotion). For this, we use the insight from the FOC (6) that any concatenation of full careers must yield discounted wages equal to the sum over the multipliers $\mu(t)$'s until infinity:

$$\underbrace{\int_t^{s+T} e^{-\beta\tilde{t}} w(\tilde{t}, s, z) d\tilde{t} + \int_{s+T}^{\infty} \mu(\tilde{t}) d\tilde{t}}_{\equiv V(t,s,z)} = \int_t^{\infty} \mu(\tilde{t}) d\tilde{t} + \int_z^{t-s} \lambda(t, s, \tilde{z}) d\tilde{z} \qquad (8)$$

We can now clearly see, again by non-negativity of the Lagrange multipliers $\lambda(\cdot)$, that workers who possess experience in an active vintage are weakly better off than career starters (whose value function is given by $\int_t^\infty \mu(\tilde{t})$). So workers always weakly prefer to stay in a career in an active vintage over starting a new career.

Combining equations (7) and (8) can tell us something about the wages of career starters. Consider two agents: The first embarks upon a long career at t. The second starts a very short career and enters a second career at $t+\epsilon$. During the short career, the second worker gets a wage of roughly $\mu(t)$, which she can attain by working in the oldest vintage $s = t+T$. After the ϵ-interval has passed, the worker in the long career must be better off than the second agent, as we saw just saw. But since the two must have been indifferent between the careers they chose at t, this implies that the entry wage for the agent with the long career must have been lower than $\mu(t)$. The following proposition, which is formally proven in appendix A.4, summarizes:

Proposition 4.1 (Entry wages highest in oldest vintage)**.** *In any solution to the planner's problem (i.e. for arbitrary T), we have $w(t,s,t-s) \leq w(t,t+T,T) = \mu(t)$ for all $t \geq 0$, for all $s \in [t-T,t]$.*

Again it is instructive to observe what occurs in equation (7) towards the end of a vintage's life. When letting $s \to t+T$, we see that wages in the dying vintage must be weakly decreasing in z. This means that all workers must be taking wage cuts when switching between vintages.

Corollary 4.2 (Wage cuts upon vintage change)**.** *In any solution to the planner's problem (i.e. for arbitrary T), consider the life of an arbitrary worker. For any point t_n, $n > 0$, where the worker switches between vintages, we have $\lim_{\epsilon\downarrow 0} w(t_n-\epsilon, s(t_n-\epsilon), z(t_n-\epsilon)) \geq \lim_{\epsilon\downarrow 0} w(t_n+\epsilon, s(t_n+\epsilon), z(t_n+\epsilon))$.*

Proof. Letting $t \to s+T$ in equation (7), we see that $w(t,t+T,z)$ must be weakly decreasing in z by non-negativity of the Lagrange multipliers. This together with proposition 4.1 implies the result. □

If we take the derivative of (7) with respect to z, we can recover λ (note that differentiability of w is assured since we assumed n to be differentiable)

$$\lambda(t,s,z) = \int_t^{s+T} -\frac{\partial w(\tilde{t},s,z)}{\partial z} d\tilde{t} = -\frac{\partial V(t,s,z)}{\partial z} \tag{9}$$

This shows the relatedness of the planner's problem to the market economy in a clear way: If the discounted value of wages is increasing in experience (or decreasing in the hierarchy z), then the experience constraint holds with equality and no demotions occur. Demotions can only happen when there is no value differential across experience levels inside a vintage.

Taking the derivative of (9) with respect to t, another important insight comes almost for free:

$$\frac{\partial \lambda(t,s,z)}{\partial s} = \frac{\partial w(t,s,z)}{\partial z}. \tag{10}$$

This equation says that if more experience yields higher pay at t (i.e. the left-hand side of (10) is negative), then the differential between the value of careers across experience levels must fade over time and the shadow price on experience in this vintage must decrease – this is to be expected, since ultimately the value differential between the different rungs in the experience ladder must vanish when the vintage dies.

This leads naturally to the question about the wage structure inside a vintage: Is it possible that a worker gets paid less than a less experienced colleague in the same vintage, motivated by high wage prospects later on in his career? It turns out that the answer is no, at least once we invoke log-concavity of the learning curve $f(h)$:

Proposition 4.3 (Wages increasing in experience)**.** *In any solution to the planner's problem (i.e. for any T), wages are increasing in experience in all active vintages. For all points (t,s,z) in the support of n, we have*

$$\frac{\partial w(t,s,z)}{\partial z} \leq 0.$$

A formal proof for this statement is given in appendix A.5. The main argument used in the proof is the following: Suppose a high-experience worker has a lower wage than the workers just below him in the hierarchy. Despite that, he prefers to stay at his experience level. Then future wages must develop even more to his disadvantage, since the gains from the learning curve are decreasing in experience by log-concavity of $f(h)$. But this means that the worker should have chosen demotion in the first place, which is a contradiction.

4.3 The optimal T^* for a stationary economy

So far, all characterizations of the planner's solution were contingent on a fixed T specifying the expiration date of all vintages. However, in the end the planner will have to decide which T to pick to maximize welfare in the economy. In order to do this, define $J(T,n)$ as the value of a (feasible, consistent and continuously differentiable) distribution n for a given T to the planner. Denote by $J^*(T) = \sup_n v(T, n)$ the value of the maximization problem defined in (2). As mentioned before, the global concavity of the problem ensures that the maximizer n^* is unique if the supremum $J^*(T)$ is attained.

However, it turns out that a maximizer cannot exist if we choose T too large. To see this, suppose there exists a solution for the global problem, i.e. there exists a pair (T^*, n^*) such that $J(T^*, n^*) > J(T, n)$ for all pairs (T, n). Then the claim is that we have $J^*(T) = J(T^*, n^*)$ for all $T > T^*$, but that the supremum cannot be reached by any n for $T > T^*$.

To see this, note that the optimal distribution for T^* could also be implemented for all $T > T^*$ by setting $n = n^*$ for all vintages of age $\tau \leq T^*$ and zero else, if we were allowed to use discontinuous functions in the maximization problem. But we are restricted to choose n from the space of continuously differentiable functions, so this is not permissible. However, we can get arbitrarily close to n^* using a sequence of continuously differentiable functions that thins out the density in vintages of age $\tau > T^*$ ever faster and converges to n^* in the end.[3] Hence $J^*(T) = J^*(T^*)$ for all $T > T^*$, but existence of a maximizer breaks down when $T > T^*$.

Because of this non-existence result it is useful to approach T^* from the left side in computations. Here, the following characterization is useful:

Proposition 4.4. (Wages constant in experience in last vintage) *In a global optimum (T^*, n_0^*) to the planner's problem, wages are constant in experience in the last vintage $s = t - T^*$:*

$$w(t, t - T^*, z) = \mu(t) \quad \textit{for all } z \in [0, T^*], \textit{ for all } t.$$

The proof is given in appendix A.6 and uses the following argument: If marginal products are still not aligned in the final vintage T^*, then a skill mix with higher average labor productivity exists that can be obtained by rearranging workers across skills. We can implement this skill mix shortly after T^* by extending the lifetime of a vintage. Notice that such a rearrangement can always be done respecting the experience constraints if enough low-skilled workers leave the vintage. Since the marginal product of low-skilled workers in the final vintage equals the marginal the value of an additional untrained worker to the planner, extending the life of the vintage in this fashion is profitable.

[3]Notice that obeying the experience constraints in the small region after T^* is not an issue here, if the distribution is thinned out at a high rate – there are always many experienced workers around then.

This line of argumentation also provides an easy way to find a lower bound for T^* – if wages are constant in the last technology in use, but one could operate this technology for some more time and afford to pay even higher wages, then definitely this should be done.

Corollary 4.5. (Lower bound for T^*) *It is never optimal for the planner to choose* $T < \bar{T} \equiv \arg\max_\tau Y^*(t, \tau)$.[4]

The proof for this claim, which is given in appendix A.7, uses exactly the same line of argumentation as the proof for proposition 4.4. It is tempting to say that thus the planner should choose $T^* = \bar{T}$, since then the value of the problem $\mu = w(T, T)$ would be maximized. However, it is not clear that there exists a planner's solution where wages are constant in the last vintage for $T = \bar{T}$. One may have to increase vintage age farther to reach this point.

Now, everything is in place to show that the planner's solution is a competitive equilibrium:

Proposition 4.6. (Planner's solution is equilibrium) *A stationary global solution* (T^*, n^*) *to the planner's problem constitutes a stationary competitive equilibrium with the following properties:*

1. *Wages are equal to marginal products for vintages of age* $\tau \leq T^*$*, and they may be chosen to equal* $\mu(t)$ *for* $\tau > T^*$.

2. *The no-crossing measure may be chosen for* μ.

Proof. We have to show that both firms and workers behave optimally given wages. Consistency between the measure on lives and the density n^* is ensured by the construction of the no-crossing measure in proposition 3.1.

First, consider firms. For vintages of age $\tau \leq T^*$, firms behave optimally by setting labor demand to n^* because this sets marginal products equal to wages. I will now show that in vintages $\tau > T^*$ no production occurs since the highest attainable profits are negative. The cost-minimizing skill mix for $\tau > T^*$ is such that marginal products are constant across z since $w(t, \tau, z) = \mu(t)$. Choosing this skill mix, average productivity of a worker is $Y^*(t, \tau)$. However, note that we have $Y^*(t, \tau) < Y^*(t, T^*) = \mu(t)$, since by proposition 4.5 the function $Y^*(\tau)$ is decreasing for $\tau > T^*$. So average productivity under the cost-minimizing skill mix is lower than the average wage. Thus production and labor demand are zero for all $\tau > T^*$.

[4]Note that $\bar{T}$ is independent of t, see the definition (11).

Second, we turn to the worker's problem. Consider the problem of a worker who faces wages as implied by marginal products of the planner's solution. We would like to show that $V(\cdot)$, as defined in (7), is the value function for a worker. For this, we have to establish that for a given starting position (τ_0, z_0), any life l that is feasible from this starting position yields a value $v(l)$ weakly lower than $V(0, \tau_0, z_0)$. So fix (τ_0, z_0) and an arbitrary feasible life l for now. Consider first an alternative life l' that is identical to l but has no demotions: Let z be constant on all careers. Since wages are weakly decreasing in z by proposition 4.3, we have $v(l) \leq v(l')$. So no-demotion lives weakly dominate lives that include demotions.

Since life l was arbitrary, the lives l and thus l' may prescribe that the worker drops out of a vintage before or after its demise at $\tau = T^*$. The goal now is to show that life l' is weakly dominated by a life in which the worker always drops out of a vintage exactly when the vintage's age reaches $\tau = T^*$. To see this, consider the following sequence $\{l'_n\}$ of lives: Life l'_n is identical to life l' until t_n, i.e. the end of the nth career. From t_n on, the worker only pursues "complete careers" in the sense that she always leaves vintages exactly when $\tau = T^*$. For $n = 0$, the worker completes all careers in his life and we have $v(l'_0) = V(\tau_0, z_0)$, as may be seen from combining equations (6) and (8). For $n = 1$, the continuation value of life l'_1 at time t_1 (when leaving the first career) is lower than the continuation value of life l'_0 at t_1 by (7), for which we invoke non-negativity of the Lagrange multipliers. Since l'_1 is identical to l'_0 up to t_1, we have $v(l'_1) \leq v(l'_0)$. By induction on this argument we have $v(l'_{n+1}) \leq v(l'_n)$. Combining the sequence of inequalities, we obtain $v(l'_n) \leq v(l'_0)$ for all n. Finally, note that $v(l'_n)$ converges to $v(l')$, as either $n \to \infty$ or as the agent stays in one vintage forever.[5] This implies that $v(l') \leq v(l'_0) = V(\tau_0, z_0)$.

Since $v(l) \leq v(l')$, we have that $v(l) \leq V(\tau_0, z_0)$ for an arbitrary feasible life l. The value $V(\tau_0, z_0)$ may be attained by life l'_0. Since (τ_0, z_0) was arbitrary, this establishes that $V(\cdot)$ is the value function for the worker.

We still have to establish that the lives prescribed by the no-crossing measure indeed attain the value given by $V(\cdot)$. By equation (9), the level lines of $\bar{N}$ and thus workers' careers can only be decreasing in z in areas where the value function is invariant in z. This means that the demoted workers attain the same value that they would attain if they stayed on the same z-level, and thus are behaving optimally. □

[5]The sequence $\{t_n\}$ goes to infinity since all careers are of the same length for no-crossing lives derived from a stationary density, see the construction of the no-crossing measure in appendix A.1.

As the discussion of the no-crossing measure in section 3.5 showed, there will usually be many measures μ on lives that are consistent with n^* and also support equilibrium. These can be engineered by making workers' paths cross in regions where they are indifferent between promotion and demotion.

Obviously, it would be of interest to extend proposition 4.6 to a full equivalence result, telling us that that (n^*, T^*) is the *only* distribution that yields a competitive equilibrium. Kredler (2010) shows a partial converse of proposition 4.6 in a closely-related setting with endogenous human-capital accumulation. The proof builds upon the Hamilton-Jacobi-Bellman equations underlying the worker's problem. This avenue of attack is not available here: Since not even continuity is assumed for workers' career paths (the function $z(t)$ can have jumps even while the worker stays in a vintage), standard dynamic-programming techniques do not apply.

5 Numerical example

This section presents results from a numerical approximation to the planner's solution. These results further illustrate the workings of the model. The solution algorithm and parameter values are given in appendix A.8.

Figure 3 shows the equilibrium wage function and the density over the (τ, z)-space in the upper two panels. Since the equilibrium is stationary, the wage function looks the same for all $t \geq 0$ up to the scaling $e^{\gamma t}$. The density stays the same for all t. As predicted by proposition 4.3, wages decrease as we move down the experience ladder in each vintage. However, the wage differentials close as the vintage ages and vanish entirely upon the vintage's death, as proposition 4.4 showed. There is *wage compression*, as in Kredler (2010). The model here shows that wage compression does not depend on late entrants catching up in skills with early entrants by investing in human capital, as one might suspect from Kredler's (2010) setting. Neither is it necessary to assume a bounded task ladder which limits the possibilities of high-experience workers to obtain wage compression.

In the lower left panel we see the log age-earnings profiles when following an agent over her career in a vintage. Future wages are discounted by productivity growth in the economy.[6] As is the case in Kredler (2010), agents who enter young technologies have the lowest entry wages but are

[6]These profiles may be obtained by keeping z constant in the upper-left panel and varying τ. Note that we need not worry about workers being demoted in the end of their careers. Demotions only occur in the region where wages are constant in z, so wage profiles look identical if we take demotions into account.

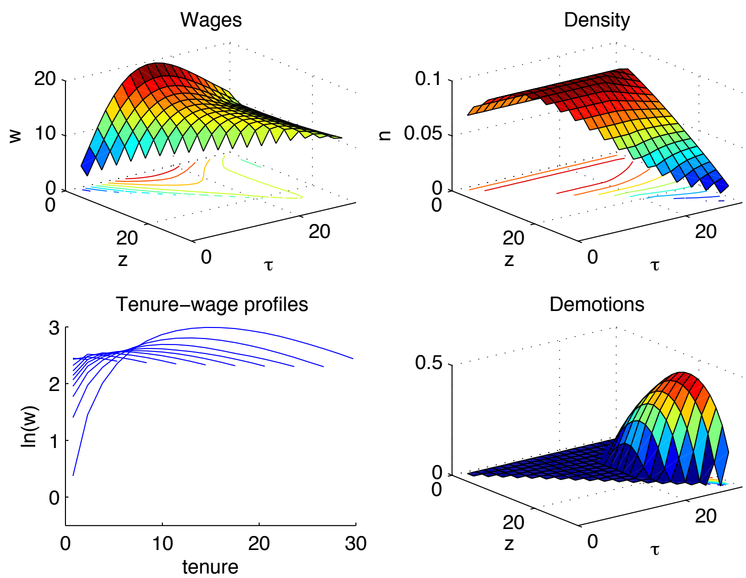

Figure 3: Equilibrium variables

compensated by high wages in the middle of their careers. Late entrants acquire skills that are less useful, since the demise of the vintage makes them obsolete sooner. They have to be compensated by higher entry wages, which means they have flatter wage profiles. This shows that endogenous human-capital accumulation is not necessary to generate overtaking (the crossing of earnings profiles in the lower-left panel). In Kredler (2010), early entrants accumulate human capital at a faster rate than late entrants. This force is not present here since a learning curve is assumed. So the setting here shows that endogenous human-capital accumulation is not necessary to generate overtaking.

The reason for the hump-shaped earnings profiles of early entrants is the following: Young technologies have lower productivity than older ones, which keeps wages low (see the lower-left panel of figure 4 for a plot of vintage productivity). Once the technology matures and becomes more productive, early entrants possess a scarce skill in a productive technology and thus command high wages. In equilibrium, there have to be few enough early entrants to push wages in the middle career stages high enough so that early entry is desirable. The density in the upper-right panel of figure 3 and the entry mass depicted in the lower-left panel of figure 4 indeed show that entry is lower in the youngest vintages than in somewhat more mature ones.

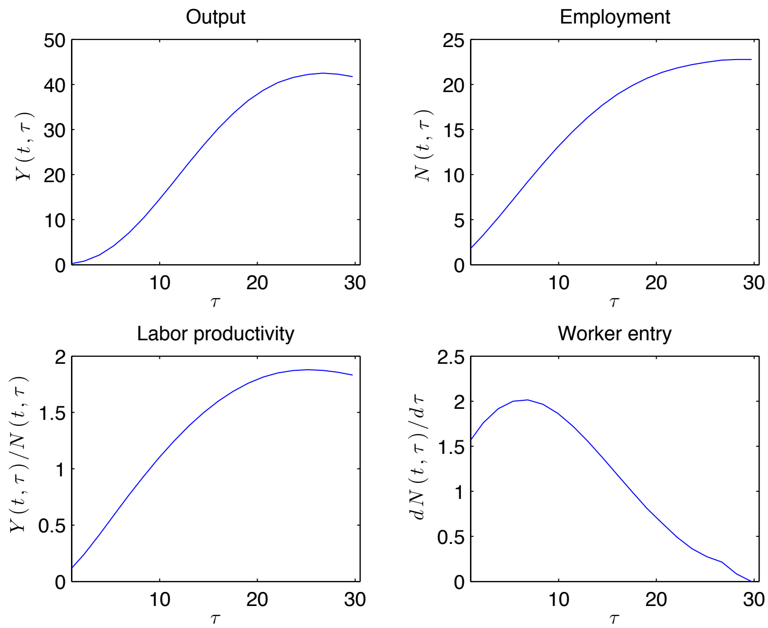

Figure 4: Vintage characteristics

However, as returns to learning slow down, the wage gap between earlier and later entrants closes. This causes workers to move down in the task latter eventually. The size of these downward flows is depicted in the lower-right panel of figure 3. In terms of the planner's problem, positive downward flows mean that the experience constraint is not binding. We see that this is the case for increasingly larger ranges of the task ladder as the vintages become older. In young vintages, skill is scarce and experience constraints bind. Since agents are moving down in the task hierarchy in older vintages, less entry is necessary as the vintage ages, as the lower-right panel in figure 4 shows.

Finally, the model has interesting predictions on output, employment and productivity of vintages. Figure 4 shows that employment increases as the vintage ages; there is entry into the vintage at all stages. Productivity and thus output also tend to rise with vintage age. The reason for this is three-fold: First, new input factors in the form of the newly-available (high) rungs in the skill ladder are added into production as the vintage ages.[7]

[7]This is evident from the range of integration in the production function becoming

Second, all workers in the vintage are progressing on their learning curves. Finally, marginal productivities are becoming more equalized as the vintage ages, i.e. the skill mix is becoming more favorable.[8]

Also comparative statics with respect to the speed of technological growth γ yield similar results as in Kredler (2010). Faster technological growth lowers T^* – the opportunity cost of sticking with an old technology is larger when the frontier blueprint improves at a faster rate. This makes agents' human capital obsolete more frequently, leading to a higher time-series volatility in wages. On the other hand, the demotion region becomes smaller, so there are fewer vintages with a compressed wage structure. More agents are assigned to new technologies with a high skill premium, leading to an increase in the average tenure premium and a higher cross-sectional dispersion in wages.

6 Conclusions

This paper has presented a model of vintage-specific human capital in which workers accumulate skills according to a learning curve. In equilibrium, there is wage compression, i.e. the premium on skill decreases as the technology ages. Entrants into young technologies start with low entry wages and then experience fast wage growth, whereas entrants into old technologies have high entry wages and low wage growth.

The results are in line with both the theoretical and empirical results in Kredler (2010), who studies an otherwise identical setting with endogenous human-capital accumulation. The similarity of the findings tells us that the most important driving force in this class of models is the scarcity of skill across vintages. As a technology ages, skill becomes more abundant and thus the premium on it decreases. The specific way that workers accumulate skills does not matter much.

One particularly stark result from the model is that experience is not rewarded at all any more in the oldest vintage. At first sight, this result is strikingly at odds with the positive tenure premia that are routinely estimated in the data. However, it has to be kept in mind that the result was derived under the assumption that *all* human capital is vintage-specific. This is most likely not the case in reality. The apparent failure of the model in this respect calls for an extension of the model. The most obvious candidate is the introduction of general human capital alongside vintage human

larger.

[8] In Kredler (2010), only the two latter effects are present.

capital. This is a daunting task, however, since it necessitates the introduction of another state variable into the problem.

References

Ben-Porath, Y. (1967), 'The production of human capital and the life cycle of earnings', *The Journal of Political Economy* **75**(4), 352–365.

Chari, V. V. & Hopenhayn, H. (1991), 'Vintage human capital, growth, and the diffusion of new technology', *Journal of Political Economy* **99**(6), 1142–65.

Hause, J. C. (1981), 'The fine structure of earnings and the on-the-job training hypothesis', *Econometrica* **49**(1), 277.

Kredler, M. (2010), 'Experience vs. obsolescence: A vintage-human-capital model', *Working Paper* .

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

Oksendal, B. (1992), *Stochastic differential equations (3rd ed.): an introduction with applications*, Springer-Verlag New York, Inc., New York, NY, USA.

Parente, S. L. (1994), 'Technology adoption, learning-by-doing, and economic growth', *Journal of Economic Theory* **63**(2), 346–369.

Violante, G. L. (2002), 'Technological acceleration, skill transferability, and the rise in residual inequality', *Quarterly Journal of Economics* **117**(1), 297–338.

A Additional proofs

A.1 Proof for proposition 3.1: Construction of the no-crossing measure

Proof. The no-crossing measure is based on the idea that we can engineer a distribution over L where agents' positions in the z-hierarchy of a vintage s with respect to the other workers in s do not change over the life cycle of the vintage. When demotions occur, a worker drops somewhat in the hierarchy, but never crosses the career paths of other agents – these are pushed down in the hierarchy as well.

A *no-crossing career* is constructed as follows. Consider a worker entering vintage s at time t_n in position $t_n - s$. Define implicitly the hierarchy position $z_{nc}(t)$ of this worker's no-crossing career by $N(t, s, z_{nc}(t)) = const = N(t_n, s, t_n - s)$ for $t \geq t_n$. This specifies that the mass of workers that are above the worker in vintage s stays constant. We let the career z_{nc} end once there is less than a mass c of workers left in the vintage. Figure 5 shows the no-crossing careers for a stationary density. For a stationary density, the density over the (τ, z)-triangle is constant and we can thus see all no-crossing careers in a two-dimensional graph. The figure uses the notation $\bar{N}(\tau, z) \equiv N(0, -s, z) = N(t, t - s, z)$.

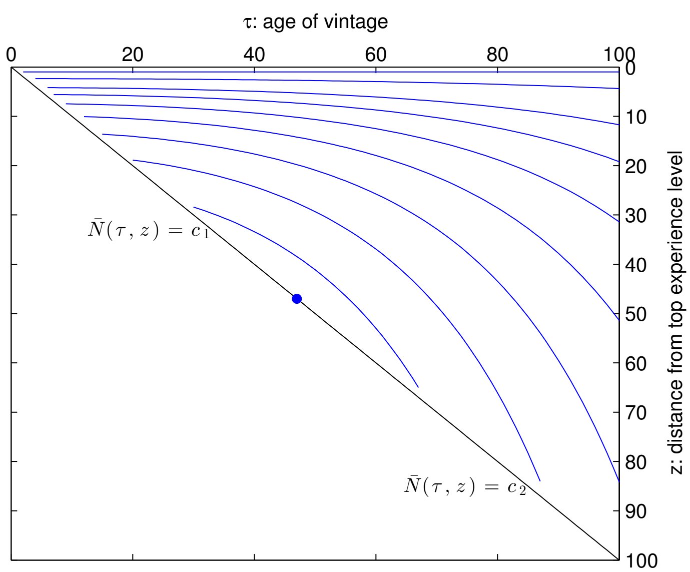

Figure 5: Construction of no-crossing measure

The *no-crossing lives* l_{nc} are an extension of the no-crossing careers. Once a worker leaves a vintage, we have to insert him into a new career again. We insert workers by maintaining their ordering in the following sense: Workers who exit an older vintage are inserted into younger vintages. If there is large number of workers leaving one vintage when it shuts down, then the workers with lower z-levels are inserted into younger vintages. This means that workers who have just had a long career once will tend to have a long careers again. In a stationary setting as in figure 5, this means that a worker will always be assigned to enter the vintage of the same age τ throughout his whole life. The worker will always follow the no-crossing career associated with the same constant c. [9]

[9]Notice that even when $n(\cdot)$ spreads the mass of workers over the infinite triangle, almost every career – i.e. all except for the one with $N = 0$ for all τ – will ultimately

We can now decide for any given number of rectangles $\{A_n\}_{n=1}^N \in \mathbb{R}^2$ if a given no-crossing life l_{nc} crosses these rectangles at a given sequence of dates $t = t_1, \ldots, t_N$, which is necessary for the construction of the measure μ. However, we still have to be able to tell which measure of lives crosses the sets to have a complete description of the stochastic process. In other words, we have to define a measure on the space of lives $(L, \mathcal{B})$. We will do so in the following manner: For each life l in a given set of lives B, check if it is a no-crossing life (for a given density n). Then, take the measure over all no-crossing lives in B by tracking them back to their starting points $\{l_{nc}(0) : l_{nc} \in B\}$ and checking how much mass the initial density ν assigns to this set of starting points. Formally, the measure of any set $B \subset L$ is set to

$$\mu(B) = \nu\big(\{(\tau_0, z_0) : \text{there is } l_{nc} \in B \text{ s.t. } l_{nc}(0) = (\tau_0, z_0)\}\big).$$

This measure obviously yields $n(\cdot)$ in the sense of definition 3.1; furthermore, the careers defined by the no-crossing measure all conform to the definition of a feasible career in 2.1 if $N(t, s, z)$ is a weakly decreasing function in t at all (t, s, z). This proves that for any consistent n, we can construct a permissible measure μ as stated in the proposition.

To see that it is *not* possible to construct a permissible measure μ if n is inconsistent (i.e. $N(t, s, z)$ is increasing in t for some (t, s, z), simply observe that this would mean that experience has been created out of nothing at or before this point: There are more agents working at ladder rung z or higher at time t then before, so some workers must have violated the non-increasingness constraint on z. This concludes the proof. □

Remark: Another strategy for a proof is to apply Kolmogorov's extension theorem for stochastic processes, which states that given a family of consistent finite-dimensional distributions there exists a probability space and a stochastic process defined on it having these distributions. See Oksendal (1992) for a statement of this theorem. However, to make use of the theorem one would have to take a stand on the joint distributions of l for multiple points in time $t_1, \ldots, t_n$, which would require a similar construction as in the proof given here.

A.2 Proof for lemma 3.2: Finite support of technologies

Proof. I will first establish that that the maximum attainable output in a vintage goes to zero in comparison to the output level attainable with new vintages when $\tau \to \infty$. To do this, calculate the maximal output attainable by the vintage of age τ when using a unit mass of workers and ignoring experience constraints:

$$Y^*(t, \tau) \equiv \max_{\int_0^\tau n(t,\tau,x)dx \leq 1} Y(t, \tau) \tag{11}$$

have to end, and the worker has to enter a new vintage again. If that was not the case, the integral of the mass of workers above this specific career would grow without bound: $\int_0^T \int_0^{z:N(t,s,z)=c} n(t, s, \tilde{z}) d\tilde{z} dt = Tc > 1$ for $T > 1/c$.

To achieve the maximum output in (11), marginal productivities across experience levels have to be equalized; this yields the following maximizer:

$$n^*(t,\tau,h) = \frac{f(h)^{\frac{\rho}{1-\rho}}}{\int_0^\tau f(h)^{\frac{\rho}{1-\rho}}dh}$$

When using this to calculate Y^*, we obtain

$$Y^*(t,\tau) = e^{\gamma(t-\tau)}\left[\int_0^\tau f(x)^{\frac{\rho}{1-\rho}}dx\right]^{\frac{1-\rho}{\rho}}.$$

So the TFP term $e^{\gamma(t-\tau)}$ makes old vintages less productive compared to new ones, but gains from factor addition (the range of the integral) and learning (increasingness of f) are working in favor of old vintages' productivity. To see which force prevails as τ grows large, consider the following derivative:

$$\frac{\partial Y^*(t,\tau)/\partial\tau}{Y^*(t,\tau)} = \frac{\partial \ln Y^*(t,\tau)}{\partial\tau} = -\gamma + \frac{1-\rho}{\rho}\frac{f(\tau)^{\frac{\rho}{1-\rho}}}{\int_0^\tau f(x)^{\frac{\rho}{1-\rho}}dx}.$$

We see that the term γ stemming from the exponential decay of TFP in vintage age is constant. If $\rho \leq 0$, then the second term is also negative (or zero) and thus $\ln Y^*(t,\tau) \to -\infty$ as $\tau \to \infty$. If $\rho > 0$, then the second term is always positive. If $f(\tau)$ is bounded, then the denominator will dominate the numerator eventually and thus $\ln Y^*(t,\tau) \to -\infty$ as $\tau \to \infty$. In the case that $f(x)$ is unbounded, both the numerator and the denominator go to infinity and we have to invoke L'Hospital's Rule:

$$\lim_{\tau\to\infty}\frac{1-\rho}{\rho}\frac{f(\tau)^{\frac{\rho}{1-\rho}}}{\int_0^\tau f(\tau)^{\frac{\rho}{1-\rho}}dx} = \lim_{\tau\to\infty}\frac{f(\tau)^{\frac{\rho+(\rho-1)}{1-\rho}}}{f(\tau)^{\frac{\rho}{1-\rho}}} = \lim_{\tau\to\infty} f(\tau)^{-1} = 0.$$

So also in this case TFP growth eventually dominates the returns from learning and factor addition, independently of the form of f. So we have established that in all cases $\ln Y^*(t,\tau) \to -\infty$ as $\tau \to \infty$, and thus $Y^*(t,\tau) \to 0$ as $\tau \to \infty$.

Now, observe that in a stationary equilibrium it is always possible for a worker to obtain some wage $e^{\gamma t}\delta$, for some $\delta > 0$, at time t. This wage may be obtained by always working as a zero-experience worker in some vintage $\tau' > 0$ close to the frontier. If we see workers staying in an old vintage, then they must be better off doing this then obtaining $e^{\gamma t}\delta$. In other words, for any career $z_n(t)$ on $[t_n, t_{n+1}]$ in the vintage τ (note that we may have $t_{n+1} = \infty$) chosen in equilibrium, we have

$$\int_{t_n}^{t_{n+1}} e^{-\beta t} w(t,\tau,z_n(t))dt \geq \int_{t_n}^{t_{n+1}} \delta e^{(\gamma-\beta)t}dt. \tag{12}$$

If this was not the case, the worker would not be behaving optimally and should choose the alternative plan of working in the newer vintage. We will now integrate up the inequality (12) over all careers in vintages of age T and older. To do this, we truncate careers with $t_n < T$ at T and observe that the above inequality must still hold (since the worker should leave the vintage otherwise). We obtain

$$\frac{1}{\gamma-\beta}\int_T^\infty Y(0,\tau)d\tau \geq \frac{\delta}{\gamma-\beta}\int_T^\infty \int_0^\tau n(\tau,z)dzd\tau.$$

On the left-hand side, we have the discounted value of all future wage payments in vintages of age T and older, which must equal discounted total output in these vintages since firms make zero profits. Since the economy is stationary, output in vintages of age T and older must grow at rate γ. On the right-hand side, we see the value of the outside option that workers have when working in a new vintage. Now observe that for any fixed $\delta > 0$, per-worker output in vintages $\tau \geq T$ can be pressed below δ when choosing T large enough since $Y^*(0,\tau) \to 0$ as $\tau \to \infty$ (by the argument in the beginning of the proof). But this means that the above inequality will be violated for T large enough. This completes the proof that the density $n(\cdot)$ must have finite support in any stationary equilibrium. □

A.3 Deriving the FOCs in the planner's problem

To obtain first-order conditions for n_0, we have to disturb n_0 by some $h_0 \in \mathcal{N}_0$, take the Frechet derivative and set to zero. See Luenberger (1973), for example, for a statement of the Lagrange multiplier theorem in infinite-dimensional spaces. Perturbing the optimal $n_0(s,z)$ by some $h_0(s,z)$, the Frechet derivative of the Lagrangian (5) is

$$\int_{s,z}\left[\int_{s+z}^{s+T} e^{-\beta t} w(t,s,z)dt\right] h_0(z,t) - \int_{z,t}\left[\int_{t+z}^{t+T} \mu(s)ds\right] h_0(z,t) = 0. \tag{13}$$

Similarly, perturbing the optimal $u(t,s,z)$ by some $h(t,s,z)$, we obtain

$$\int_{t,s,z}\Bigg[\underbrace{\int_{s}^{t+T} w(\tilde{s},t,z)d\tilde{s}}_{=V(t,s,z)}\Bigg] h(t,s,z) = \int_{t,s,z}\left[\int_{s}^{t+T} \mu(\tilde{s})d\tilde{s}\right] h(t,s,z) \tag{14}$$

$$+ \int_{t,s,z}\left[\int_{z}^{s-T} \lambda(t,s,\tilde{z})d\tilde{z}\right] h(t,s,z).$$

Since these two equations must hold for any perturbations h_0 and h, we obtain the first-order conditions (6) and (7) as claimed in the text. If any of the two was not true at any point, the equalities could be violated could be violated, respectively, by choosing h_0 or h zero except for a small area around this point.

A.4 Proof for proposition 4.1: Entry wages highest in oldest vintage

Proof. Consider an entrant into vintage $s \in [t-T,t)$ at a fixed time t. The entrant's value function over the course of the career is given by $V(t+\epsilon, s, t-s)$, where we vary $\epsilon \in [0, T-t+s]$. Equation (8) tells us that the infinitesimal change in value at the beginning of the career is given by $dV/d\epsilon|_{\epsilon=0} = -w(t,s,t-s)$.

We will now compare this to how the outside option, i.e. the value of a career starter, changes over time. Define $W(\epsilon) = \int_\epsilon^\infty \mu(\tilde{t})d\tilde{t}$, which by equation (6) is the value for an agent without experience. We have $dW/d\epsilon|_{\epsilon=t} = -\mu(t)$. Since $V(t,s,z) \geq W(t)$ for all t, $s \in [t-T, t]$, $z \in [0, t-s]$ by non-negativity of the Lagrange multipliers in equation (8), it must be that $dV/d\epsilon|_{\epsilon=t} \geq dW/d\epsilon|_{\epsilon=t}$. Using the derived expressions for the derivatives, this implies that $w(t,s,t-s) \leq \mu(t)$. Finally, letting $s \to t+T$ in equation (6) implies that $w(t,T,T) = \mu(t)$. □

A.5 Proof for proposition 4.3: Wages increasing in experience

Proof. By way of contradiction, suppose there were some point (t^*, s^*, z^*) at which $\frac{\partial w(s,t,z)}{\partial z} > 0$. By the continuous differentiability of $n(t,s,z)$, also $w(t,s,z)$ is continuously differentiable. Continuity of $\frac{\partial w}{\partial z}$ implies that there is an interval $[s^*, \bar{s})$ on which $\frac{\partial w(t,s,z)}{\partial z} > 0$, where $\bar{s}$ is defined as $\bar{s} = \min\{s > s^* : \frac{\partial w(t,s,z)}{\partial z} \leq 0\}$. By (10), $\frac{\partial \lambda}{\partial s} > 0$ for all $s \in [s^*, \bar{s})$ in this interval. This, in turn, directly implies that $\lambda > 0$ everywhere on the interval – if λ was zero at some point, it would have to become negative in the immediate neighborhood, which is impossible since Lagrange multipliers must be positive. So the experience constraints hold with equality throughout the interval.

Now, observe that by the continuous differentiability of $n(\cdot)$, (9) implies that $\lambda(\cdot)$ is continuous in its arguments. Hence, for any fixed $s' \in [s^*, \bar{s})$ there must be an open ball around (s', t, z) where $\lambda > 0$, and hence $N(t,s,z) = \int_0^z n(t,s,\tilde{z})d\tilde{z}$ does not change with s. But then neither its derivatives $n(t,s,z)$ can change within this ball, which is equivalent to writing that $\partial n(s',t,z)/\partial s = 0$ for all s'.

Since the density is invariant on the interval $[s^*, \bar{s})$, any changes in the wage differential across experience levels over time must come from experience alone. But if wages are already decreasing in experience at s^*, then matters will only become worse for the experienced as agents climb up in the skill ladder: If returns to experience $f(h)$ are log-concave, then the the marginal return to experience decreases as experience h increases. To see this, consider the elasticity of wages with respect to z:

$$\begin{aligned} \hat{w}_z(t,s,z) \equiv \frac{\partial \ln w(t,s,z)}{\partial z} &= \frac{\partial w(t,s,z)/\partial z}{w(t,s,z)} = \\ &= -\rho \frac{f'(s-t-z)}{f(s-t-z)} - (1-\rho)\frac{\partial n(t,s,z)/\partial z}{n(t,s,z)} \end{aligned} \tag{15}$$

We now want to establish the claim that $\partial \hat{w}_z/\partial s < 0$ on the interval $s \in [s^*, \bar{s})$. Since there are no changes in $n(\cdot)$, the second term on the right-hand side of (15) is invariant in s. The first term on the right-hand side must increase in s if f is log-concave: If $\ln f(h)$ is concave, then $\frac{\partial \ln f}{\partial h} = f'(h)/f(h)$ is a decreasing function in h. Then $-\rho f'(s-t-z)/f(s-t-z)$ is increasing in s, so $\frac{\partial \hat{w}_z}{\partial s} > 0$. Economically speaking, higher experience gives a worker always smaller advantages over the ranks just below as he climbs the career ladder, since returns to experience are decreasing.

To complete the proof, observe that $\hat{w}_z(\bar{s}, t, z) = \hat{w}_z(s^*, t, z) + \int_{s^*}^{\bar{s}} \partial\hat{w}_z/\partial s > 0$ (since $\hat{w}_z(s^*, t, z)$ by the assumption that w is decreasing in z at (s^*, t, z)). But this contradicts the definition of $\bar{s}$. □

A.6 Proof for proposition 4.4: Wages constant in experience in last vintage

Proof. In this proof, we will denote densities using the age of a vintage: $n(t, \tau, z)$ since it is more convenient for the sake of this argument.

Suppose, by way of contradiction, that marginal products were not constant across z for $\tau = T^*$ at t. This implies that per-worker output in vintage $\tau = T^*$ stays below its unconstrained optimum $Y^*(t, T^*)$ – recall that marginal productivities have to be equalized to obtain this maximum. Since $Y^*(t, \cdot)$ is a continuous function, there is an interval $I_\delta = (T^*, T^* + \delta)$ such that $Y^*(t, \tau) > Y(t, T^*)/\int_0^\tau n(t, \tau, z)dz$ for all $\tau \in I_\delta$.

Consider now the strategy of relocating a small amount of workers to an interval $[T^*, T^* + \delta]$ to obtain optimal (or at least very-close-to-optimal) output $Y^*(\tau)$ on this interval. This relocation will be feasible by extending the life-time of vintages beyond T^*.[10] I will now argue that it is feasible to obtain output per worker arbitrarily close to $Y^*(t, \tau)$ on the interval $\tau \in [T^*, T^* + \delta]$.

First, we will see that it is always possible to implement a maximizing experience mix n^* and produce $Y^*(t, \tau)$ for $\tau = T^* + \epsilon$ for any given $\epsilon > 0$. Note first that $n(t, T^*, \cdot)$ must be lower-bounded, i.e. $n(t, T^*, z) \geq c$ for all z for some $c > 0$. If this was not the case, marginal productivity would grow infinite for some z otherwise, which is clearly not optimal. It must then be possible to choose some low level $n(t, T^* + \epsilon, z) < c$ and equalize factor returns, so that output per worker is maximal; this is clear from the experience constraint (4).

Second, note that the transition to optimal per-worker output can be made in arbitrarily short time, i.e. any $\epsilon > 0$ may be chosen. To see this, consider a density function that leaves the density $n(t, T^*, z)$ unchanged for $\tau \neq T^*$, but then jumps to the optimal skill mix, i.e. we have $Cn^*(t, T^*, z)$ for some constant $C > 0$ for all $\tau > T^*$. This is not a feasible policy since the planner is restricted to use continuously-differentiable density functions. However, this discontinuous policy may be arbitrarily well approximated by continuously-differentiable functions, which means that a smooth transition may be engineered for arbitrary $\epsilon > 0$. Also, the transition can clearly be chosen such that experience constraints are respected by thinning out the density fast enough, see the argument above.

Third, consider the following policy for vintages $\tau \in [T^* + \epsilon, T^* + \delta]$: Maintain the density on the rungs $z \in [0, T^* + \epsilon]$ as it is at $T^* + \epsilon$ and add arbitrary worker

[10] This is possible for most vintages; only for very young vintages it might not be feasible to extend the support of the density up to $\tau = T^* + \delta$ since initial conditions do not allow it.

flows on the newly-created rungs $z > T^* + \epsilon$. By continuity of the production function, output in the vintages $[T^* + \epsilon, T^* + \delta]$ may be kept arbitrarily close to optimal production $Y^*(t, \tau)$ in these vintages when choosing δ small enough.

The profitable deviation is now constructed as follows: We can choose δ small enough so that per-worker output in vintages $[T^* + \epsilon, T^* + \delta]$ is as close to $Y^*(t, T)$ as we wish. We then choose ϵ small enough so that joint output per worker in the vintages $[T^*, T^* + \delta]$ is very close to $Y^*(t, T)$. Finally, we have to thin out the density in vintages $\tau \leq T^*$ in order to free workers to work in vintages $\tau > T^*$.

The workers in vintages $\tau > T^*$ increase the planner's criterion since their marginal productivity is larger than $\mu(t)$, which gives us the marginal value of an additional untrained worker. Average productivity is larger in $\tau > T^*$ than in T^*, which in turn is larger than $\mu(t) = w(t, T^*, T^*)$ – recall that marginal products are lowest for low-experience workers in the last vintage by the planner's optimality condition (7).

Hence, (T^*, n_0^*) cannot be a global optimum, which concludes the proof. □

A.7 Proof for corollary 4.5: Lower bound for T^*

Proof. By way of contradiction, suppose that a policy with $T < \bar{T}$ was optimal. By proposition 4.4, marginal products of workers must be equalized for $\tau = T$. Now consider the possibility of extending the life of vintages to $\tau = T + \delta$ for $\delta > 0$. This time, maintain the density on the ladder rungs $z \in [0, T]$ constant and add arbitrary positive worker flows for $z \in [T, T + \delta]$. For this policy, denote by $Y_e(t + T + \tau, T + \tau)$ the per-worker output of the fixed vintage over time. The function Y_e is continuous and differentiable in τ. Consider now the maximal per-worker output for this vintage, which is $Y^*(t + T + \tau, T + \tau)$. When regarding Y_e and Y^* as functions of τ, we see that Y^* is an upper envelope for Y_e and $Y_e(t + T, T) = Y_e^*(t + T, t)$, so the derivatives of Y_e and Y_e^* must coincide for $\tau = 0$. Since Y^* has positive slope in τ (we assumed $T < \bar{T}$), output per worker can be made higher for vintages $T + \tau$ than for T by the deviation we considered. By the same argument as in the proof for proposition 4.4, this implies that T cannot be optimal. □

A.8 Numerical solution algorithm

The parameters chosen for the numerical example are $\gamma = 0.1$, $\beta = 0.2$ and $\rho = 0.5$. I choose a concave form for the learning curve, specifying that also inexperienced workers are productive: $f(h) = 0.2 + h^{0.7}$. The idea behind the algorithm to solve for the planner's solution is similar to the one in Kredler (2010).

For fixed T, the state space is discretized on a grid for τ, z. Workers' strategies are represented by a choice of the vintage they enter and demotion decisions: Workers can either stay at the current z-level or drop one level down. For a given guess for workers decisions, the resulting stationary density can be computed. This density gives us a wage function. Following proposition 4.3, the optimal policy response of workers to wages is now obtained as follows: If wages are increasing in human capital at a given tuple (τ, z), it is assumed that the planner's constraint binds at this point and all workers are kept at the current z-level. If wages are decreasing, a negative slope of workers' z-function is obtained which is proportional to the derivative of the wage function at this point. Just as in chapter 1, entry densities into vintages are adjusted along the way such that vintages with high entry values receive more workers in the next iteration. When varying workers' policies across iterations by small-enough amounts, the algorithm converges: Entry values into all vintages are equalized, and demotions occur if and only if the wage function is invariant in z.

T^* is then obtained by increasing T until wages in the oldest vintage are invariant in z. It turns out that the algorithm usually does not make wages exactly equal: It is hard to get wages at the lowest z-point as low as the other wages. When choosing T large enough such that this occurs, then often the wage function becomes increasing in z for some z.

The algorithm converges for most parameter values, but it is less stable than the one considered in Kredler (2010). This is probably due to the non-smoothness of the problem: The decision if to stay at the same z-level is discrete, which one can interpret as an infinite cost of raising one's z coupled with a zero cost of lowering it. In Kredler (2010), workers face a smooth cost functional that penalizes the slope of their careers and decisions at all grid points are interior decisions, which seems to promote stability in the algorithm.