

Vintage Human Capital and Learning Curves*

Matthias Kredler[†]

2 January 2014

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; this makes the ex-ante homogeneous workers enter different vintages. The continuous-time framework allows me to study the timing decision for the technology phase-out differentially and to derive sharp characterization for wages and the distribution of workers in the dying technology. I show how to posit and solve a planner's problem and construct equilibrium in this way. Consistent with empirical evidence, I show that the experience premium is always positive but diminishes as a technology ages. The connection between workers' learning curves and the technology's progress curve is characterized.

Keywords: vintage human capital, tenure-wage profiles, learning curve, infinite-dimensional state space, Lagrange-multiplier theorem

JEL codes: J01, E24

*This paper is part of my dissertation thesis at New York University. I thank Boyan Jovanovic, Jess Benhabib, Greg Kaplan, Silvana Melitsko and three anonymous referees for discussion and comments. Part of the research was funded by the Spanish Ministerio de Ciencia e Innovación, reference number SEJ2007-62908, and by the Spanish Ministerio de Economía y Competitividad, reference number ECO2012-34581. Accepted for publication at the *Journal of Economic Dynamics and Control*, DOI: 10.1016/j.jedc.2014.01.003.

[†]Departamento de Economía, Universidad Carlos III de Madrid, Calle Madrid 126, 28903 Getafe, Spain. Email: matthias.kredler@uc3m.es.

1 Introduction

1.1 Motivation

There is a huge empirical literature on learning-by-doing that has established that economic agents increase their productivity at a predictable pace as they find out how to best use a new technology (see the surveys by Thompson, 2010, from the micro-economic angle and by Yelle, 1979, from the business-and-management angle). This is true both for individual workers as for firms and entire industries. I will follow Thompson’s (2010) excellent survey and refer to a “learning curve” as increases in productivity that an individual worker exhibits, whereas the “progress curve” refers to the empirical relationship between a firm’s current productivity and cumulative past output of a good.

The first empirical studies on learning curves date back as far as the late 19th century and come from the psychology literature. Progress curves are probably most famous because of Moore’s Law, which describes advances in the production of computer chips remarkably well. However, the first empirical studies of progress curves date back to the time around World War II, when estimating and predicting productivity advances in building new types of aircraft and ships was crucial for a country’s war effort.¹

An aspect that is usually neglected in the above-mentioned empirical studies is the switching decision: when should an economic agent stop learning about an old technology and switch to a newer technology? Workers and firms are routinely faced with the decision if to further exploit their accumulated knowledge in an old technology, or to upgrade to a new, more productive technology that requires learning. Examples in reality are a computer programmer switching to a new programming language, or a car company opening a new plant that implements a revolutionized manufacturing method. The technology-specific skills that agents accumulate over time are called *vintage human capital* in the literature (a term coined by Chari & Hopenhayn, 1991).

This paper aims to study the connection between learning curves and vintage human capital in a theoretical model. I take learning curves for individual workers as given and ask the following questions: when, or if at all, should workers switch to new technologies? How should firms manage the division of tasks between workers of different experience levels? What is the connection between workers’ learning curves and the firm’s/technology’s progress curve? I do this in a setting in which workers of different skill levels are complementary in production. This gives rise to a non-trivial connection between workers’ individual learning curve and the technology’s progress curve, which is a new contribution to the literature (to the best of my knowledge).

I find that workers go through cycles of learning and scrapping of skill. Technologies are shut down once returns to learning are sufficiently low, and all workers from the obsolete vintage re-locate to newer vintages. Unlike most other labor-market the-

¹These observations about the literature are drawn from the surveys of Thompson (2010) and Yelle (1979).

ories, the model predicts that workers experience initial wage cuts upon re-location. Since different skill levels are complementary and existing vintages need a supply of low-skill workers, workers enter technologies of different age cohorts and not only the frontier vintage.

The model has rich predictions on earnings profiles of workers over their careers and on the wage structure in different technologies. I show that within each technology, higher skill is always rewarded by higher wages, but that this skill premium decreases over time and eventually vanishes entirely as the technology becomes obsolete. This is driven by skill becoming less scarce as the technology ages. The model predictions are consistent with evidence provided by Michelacci & Quadrini (2009) and Kredler (2013), who find that entry wages are lower but the premium on experience is higher in fast-growing industries. They are also consistent with the premium on experience being higher in young establishments (see the evidence in Kredler, 2013).

The model also predicts that a technology's progress curve is steeper than individual workers' learning curves. The reason is that a vintage can continuously draw on learning by newly-entering workers who are on the steep part of their learning curves. Thus a firm or industry can avoid decreasing returns in learning for longer than an individual since it continuously changes its skill mix. Finally, I show that a technology's productivity growth must fall below the economy-wide rate of technical change before it becomes obsolete and is phased out.

From the technical point of view, I exploit a continuous-time setting to study the timing decision of a technology phase-out differentially. This leads to sharp characterizations for wages and the distribution of workers in the dying vintage. I show how to state and solve a planner's problem and construct a competitive equilibrium from there. A key problem that has to be overcome is that the distribution of workers over the state space (a choice variable for the planner) has to be linked to a collection of feasible policies by workers. To ensure this, I derive a set of constraints on the distribution that are necessary and sufficient for feasibility. I then use these constraints in a maximization problem and characterize wages and workers' behavior, drawing heavily on the Lagrange-multiplier theorem for infinite-dimensional spaces.

The remainder of the paper is organized as follows: Section 1.2 discusses the relationship of the paper to the literature. Section 2 presents the economic 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.

1.2 Literature review

I first discuss the class of models most closely related to my framework: vintage-human-capital settings with multi-worker firms, which were pioneered by Chari & Hopenhayn (1991). These authors study a discrete-time setting in which workers live for two periods. There are two inputs to production: skilled and unskilled workers.

The key difference to my framework is the short life span of workers, which means that there is no meaningful learning curve and that the predictions on workers' earnings profiles are less rich. This means that Chari & Hopenhayn (1991) cannot address concavity in experience-earnings profiles, which is routinely found in Mincerian earnings regressions in the data. Their model shares some predictions with the one in this paper, such as the declining skill premium in older vintages and the finite life time of technologies. However, there are also key differences: workers never experience wage losses in their setting, whereas they do occur in mine when workers re-locate to a new technology.

Even closer related to the current paper is Kredler (2013), who extends the basic setting of Chari & Hopenhayn (1991) by introducing endogenous human-capital accumulation à la Ben-Porath (1967). Many of the predictions arising from his framework are similar to mine, but they often rely importantly on the human-capital-accumulation channel (see Section 6 for a more detailed comparison between the learning-curve and the endogenous-human-capital-accumulation approach). The current paper is also able to improve upon the theoretical results in Kredler (2013) by deriving sharper characterizations of the wage structure and the link of the progress curve to the learning curve.

Second, there are models where skills are technology-specific, but which differ in other respects from mine. Violante (2002) and Parente (1994) both consider workers who face a trade-off between a learning curve on an old technology and switching to a new, more productive technology. Unlike my framework, both models have single-worker firms. Thus they do not distinguish between a (worker's) learning curve and a (firm's or industry's) progress curve. Also, the predictions on workers' earnings profiles are less rich than those in my framework since interaction between different skill levels is absent.

Third, there is large literature that studies the relationship between *general* human capital (i.e. education) and technology. Bartel & Lichtenberg (1987, 1991) show that high-tech industries in the U.S. employ more high-education workers than other industries. Galor & Tsiddon (1997) and Jovanovic (1998) provide theories for how education is related to workers' sectorial choice and how technological progress can boost economic inequality.² In this literature, there is no re-location decision to new technologies. Also, workers are *ex-ante* heterogeneous and human capital is of the *general-purpose* type. In my setting, however, all heterogeneity is *ex-post* and human capital is *technology-specific*. So these models do not address the substantial heterogeneity in earnings profiles in the data that remains even after controlling for observable characteristics such as education (for evidence, see Abowd et al., 1999, among others).

Finally, there is a large literature on vintage capital of the physical kind. These models differ from mine mainly in the scope of predictions: vintage-capital models make predictions on capital investment and output, but not on worker's career deci-

²Cooley et al. (1997) study a similar type of model, but the allocation of workers with different education levels across technologies is indeterminate in equilibrium. Jovanovic & Nyarko (1996) provide a micro-foundation for the connection of education and learning.

sions and wage patterns. Also, vintage-capital models typically feature only a firm decision, whereas in vintage-human-capital models the non-trivial optimization problem is on the worker side. Once equivalence of equilibrium to a planner's problem is established in a vintage-human-capital model, however, there are many similarities, especially when it comes to technical aspects of the analysis. The Lagrangian approach has been applied to vintage capital by Malcomson (1975). Boucekkine et al. (1997, 1999) use the same approach to study replacement echoes in vintage-capital models, drawing also on techniques pioneered by Solow et al. (1966) and van Hilten (1991).³ Boucekkine et al. (2011) use an alternative variational method borrowed from Hritonenko & Yatsenko (2005), similar to Jovanovic & Yatsenko (2012). All of these papers have in common that they yield optimality conditions for a vintage in terms of ordinary differential equations (ODEs), whereas I am using partial differential equations (PDEs).⁴ This stems from the fact that different vintages of physical capital can be managed separately under the standard assumption that different technologies (say plants or machines) are perfect substitutes in production. In my framework, workers with different skills are complementary to each other, so that workers' problems are inherently interwoven.

Within the physical-vintage-capital literature, Jovanovic & Yatsenko's (2012) model is conceptually closest to mine. As I do for workers, they assume an exogenous schedule for productivity growth for capital. As I do for skills, they specify a standard constant-elasticity-of-substitution (CES) aggregator over capital from all vintages. The difference is that in their model different vintages of capital are complementary, whereas in my model workers with different skill levels *within* a vintage are complementary inputs in production. Furthermore, in my framework all vintages produce a perfectly-substitutable good, so there is no complementarity *between* vintages. This leads to vintages surviving forever in their framework but not in mine.

Finally, there are variations on the (physical) vintage-capital model in which different vintages are linked to each other through non-linear revenue and/or cost functions and thus a PDE approach is needed.⁵ In these models, the firm's total-cost and/or total-revenue functions are assumed to be non-linear, thus linking the optimization problems of the different vintages and necessitating the use of PDEs. Using similar methods to mine, Prskawetz & Veliov (2007) consider a firm's age-specific labor-demand and human-capital investment problem, but they do not consider the labor-supply side as this paper does. Goetz et al. (2008) provide a discussion on the relationship of the vintage-capital literature in economics to the analysis of age-structured systems in mathematical biology, operations research etc., which uses a similar approach (e.g. Brokate, 1985, and Feichtinger et al., 2003).

³In a follow-up paper, Boucekkine et al. (2003) build on this line of research by studying the productivity slowdown under embodied technological change.

⁴These ODEs often involve a delay term stemming from the vintage echoes. Benhabib & Rustichini (1991) and Benhabib & Rustichini (1993) draw on the theory of delayed differential equations (see Bellman & Cooke, 1963) in their analysis.

⁵See Barucci & Gozzi (1998, 2001), Feichtinger et al. (2008, 2006a, 2006b), Goetz et al. (2008) and Saglam & Veliov (2008), among others.

2 Setup

2.1 Technology and firms

Time t is continuous. In every instant s , a new production technology (or *vintage*) arrives that is available to competitive firms for all $t \geq s$. New technologies have higher total factor productivity but are otherwise identical to old technologies. We will refer to the vintages by their date of inception, s .

Workers may enter a vintage s at any point $t \geq s$ in time, thus having different amounts of experience with the technology. I assume that each experience level constitutes a separate input to the production technology, i.e. the vintage of age $\tau \equiv t - s$ uses labor inputs which are differentiated by experience levels $h \in [0, \tau]$.⁶ 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 vintage s – they are “founding members” of their technology. In Section 2.3, I will exactly specify how experience is accumulated.

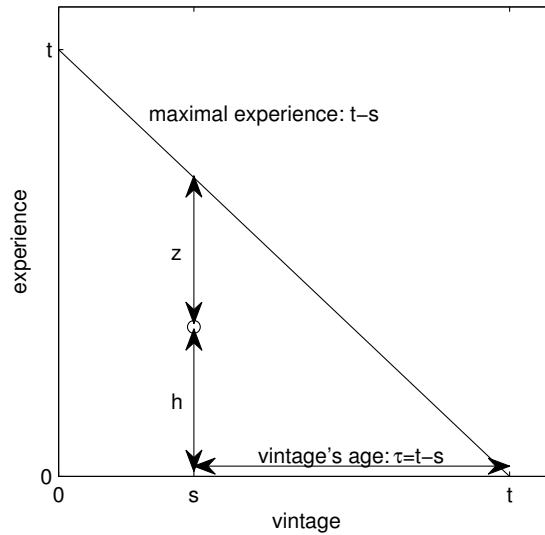


Figure 1: Vintages and experience levels

Figure 1 shows the triangle of possible experience levels for the vintages $s \leq t$ that are active at time t . The upper side of the triangle, which cuts the horizontal axis at t and has a downward slope of 45 degrees, represents the workers who have

⁶By assuming that skills increase smoothly with time, I follow the literature on learning curves. An alternative would be to assume that there is a discrete set of skill levels, and that workers switch between these at certain rates. Continuous skills allow me to use differentiation and thus provide tractability – they may or may not be an accurate description of the world. Note that the continuity assumption on skill is not qualitatively different from the routine continuity assumption we make for many *goods*, although most goods come in discrete quantities in reality (e.g. apples, microprocessors, haircuts).

the maximally-possible experience $h = \tau$ in the vintage of age τ , whereas entering workers with experience $h = 0$ are located on the horizontal axis. The worker in vintage s represented by the circle has experience h (the distance from the horizontal axis).

It is convenient to introduce another variable which denotes the hierarchy level of skill within a vintage. Complementary to h , define $z \equiv t - s - h$. This variable describes how far a worker is away from the maximum experience level $t - s$ in his technology. In the figure, it is represented by the vertical distance to the maximal experience level. This notation is usually more convenient, since z is constant for a worker who makes full use of his experience. In Figure 1, a worker who accumulates experience in vintage s would increase her distance to the frontier vintage t over time, but her vertical distance z to the maximal-experience worker would stay constant.⁷

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

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

where $\gamma > 0$ and $\rho \in (0, 1)$. $n(t, s, z)$ denotes the density of workers with hierarchy position z in vintage s at time t . Total factor productivity of new vintages increases at rate γ , capturing technological progress. Otherwise new technologies are the same as old ones. The parameter ρ captures the substitutability between different skill levels: skills would be perfect substitutes in the limiting case $\rho = 1$, but they become more complementary as ρ decreases.^{8 9}

f specifies the learning curve, which is a function of experience $h = t - s - z$. $f(h)$ can be interpreted as the number of *efficiency units* of labor that a worker

⁷At least, that is, if she makes full use of her skills – this will become clear in Section 2.3.

⁸Assuming $\rho > 0$ implies that no skill is essential for production and that skills are gross substitutes. This is intuitively reasonable: employing workers of *all* skill levels should not be a necessary condition to produce output, and it should not be too costly to replace one worker by another worker that has slightly different skills. In general, the CES production function also allows for values $\rho < 0$ (gross complements) and the limiting Cobb-Douglas case $\rho = 0$. I exclude these cases not only because they seem less plausible but also for the following technical reason: the above production function implies that productivity grows unbounded as the range of integration goes to zero if $\rho < 0$. To see this, consider the simple case of a production function $X_\tau = [\int_0^\tau m(x)^\rho dx]^{1/\rho}$ with a continuum of skills indexed by x . As an example, consider the input vector $m(x) = 1$. Under this input vector, labor productivity is $X_\tau / \int_0^\tau m(x) dx = [\int_0^\tau 1 dx]^{1/\rho} / \int_0^\tau 1 dx = \tau^{1/\rho} / \tau$, which goes to infinity as $\tau \rightarrow 0$ iff $\rho < 0$. Equilibrium in my model becomes trivial in this case: it is efficient to assign all workers to the newest vintage, production being unbounded. For lack of a reasonable specification of the production function in the case of gross complements I thus restrict the analysis to the case $\rho > 0$. For the same reason I exclude the Cobb-Douglas production function $\exp[\int_0^\tau \ln m(x) dx]$, which obtains as the limit of the CES production function as $\rho \rightarrow 0$ for fixed τ .

⁹Note that as long as there are no capital-skill complementarities, it is inessential that capital is omitted from the production function, at least on a balanced-growth path. To see this, suppose that production at time t in vintage s is given by a CES-aggregator over capital $K(t, s)$ and the labor composite $Y(t, s)$ specified above, i.e. $Z(t, s) = [\alpha K(t, s)^\sigma + (1 - \alpha)Y(t, s)^\sigma]^{1/\sigma}$. If aggregate capital grows at rate γ on the balanced-growth path and its distribution across vintages is stationary, then the ratio $Y(t, s)/K(t, s)$ is constant over time, and marginal products of all skills (i.e. wages) are exactly proportional to those in the specification without capital.

of skill h possesses. f possesses the usual characteristics of learning curves: it is assumed to be continuously differentiable, non-decreasing and weakly log-concave, i.e. $d \ln f(h)/dh$ is non-increasing in h . Log-concavity is a weaker requirement than concavity; it requires the proportional increment $f'(h)/f(h)$ to be decreasing in h , whereas concavity requires the absolute increment $f'(h)$ 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 Workers and preferences

The economy is populated by a measure 1 of workers. Under complete capital markets, i.e. as long as workers can save and borrow against future wages, workers will maximize the present value of wages under any preferences for consumption that satisfy non-satiation, as is well-known. I assume that workers live forever; they thus maximize the criterion

$$\int_0^{\infty} e^{-\beta t} w_t dt,$$

where β is the interest rate or rate of time preference. Since wages will grow at the rate of technical change in equilibrium, we need to impose the familiar $\beta > \gamma$ for the criterion not to diverge.

A worker enters the model in $t = 0$ in position z_0 in some vintage $s_0 \leq 0$, where $z_0 \leq -s_0$. The measure of workers ν over existing vintages at $t = 0$, i.e. the triangle $\{(s, z) : -\infty < s \leq 0, 0 \leq z \leq -s\}$, is exogenously given. It tells us about the initial distribution of experience. The distribution of workers over the (s, z) -space is the economy's state.

While an infinite life is certainly not realistic, this assumption is needed for tractability. If agents had finite lives, then their age would become an additional state variable, making the problem very hard, if not impossible, to analyze. We may identify young workers in the data with those agents in the model who start with no or little experience at $t = 0$, whereas older workers may be seen as those endowed with high levels of experience. Old workers in reality have similar incentives to experienced workers in the model: they possess technology-specific skills that young people do not have, and they will seek to reap the benefits from this skill differential as long as they can. Introducing age explicitly into the model is an interesting and challenging extension that I leave for future research.

2.3 Experience accumulation

Workers can choose both the vintage in which they work and the promotion paths they 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 a worker. A life will consist of a sequence of careers indexed by $n = 0, 1, \dots$. A worker who makes full use of her experience climbs the h -ladder at normalized speed 1, i.e. her z -level stays constant. However, workers are also allowed to remain at tasks they already master and climb the task ladder slower or even drop down in the ladder.¹⁰

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 $z_n(t) : [t_n, t_{n+1}) \rightarrow \mathbb{R}$ 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 worker’s initial endowment comes into play. The starting point for a worker endowed with (z_0, s_0) may be any position $z \in [z_0, t - s_0]$ in vintage s_0 , but has to be $z = -s$ for any vintage $s \neq s_0$.

Figure 2 shows possible careers. The entry points into the careers are fixed to time $t_0 = 0$. A worker chooses a vintage s and enters at position $z = -s$, i.e. with experience $h = 0$, at $t = 0$. The figure only shows careers for workers who do not possess any initial endowment 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).

¹⁰This concept of careers is close to that of industrial and organizational psychology. In this literature, career definitions “revolve around the notion of sequential employment-related experiences through time and across space” according to the survey by Baruch & Bozionelos (2011).

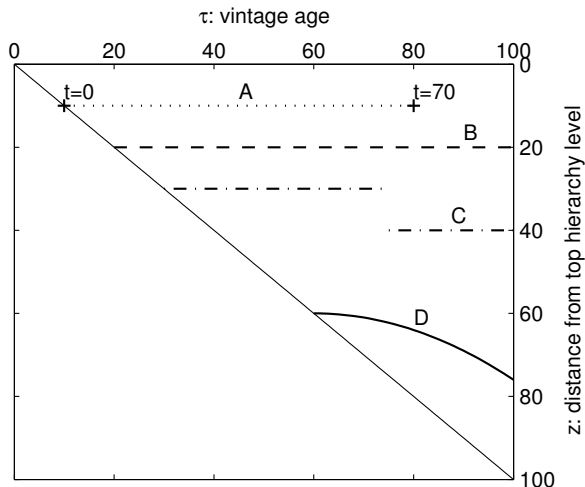


Figure 2: Some possible careers for workers entering at $t = 0$

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 a worker is a sequence of subsequent careers. Precisely, a life is a function $l(t) = [s(t), z(t)]$, $l : [0, \infty) \rightarrow \mathbb{R}^2$, mapping any instant in time t to a vintage $s \leq t$ and a position z that is feasible given the worker's past. s 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 of his life. Since workers do not value leisure, they have no incentive to do so anyway.

3 Competitive equilibrium

In this section, I define a competitive equilibrium for the economy described in the previous section. I first analyze the firm's problem, which is static. I then turn to the worker's problem, which incorporates the interesting dynamic trade-offs in this economy. Finally, I provide the formal definition of a (stationary) competitive equilibrium and characterize some of its properties.

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

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

where $n(t, s, z)$ is experience- z labor demanded by a vintage- s firm at t and $w(t, s, z)$ is the market wage for experience- z labor in vintage s at t . Since the production technology is constant-returns-to-scale, profits $\pi(t, s)$ are zero for all (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 , wages are given by the marginal product of workers in equilibrium:

$$w(t, s, z) = e^{\gamma s} f(t-s-z)^\rho \left(\frac{[\int_0^{t-s} f(t-s-\tilde{z})^\rho n(t, s, \tilde{z})^\rho d\tilde{z}]^{1/\rho}}{n(t, s, z)} \right)^{1-\rho}. \quad (2)$$

3.2 The worker's problem

Given a wage function $w(t, s, z)$ and the initial endowment (s_0, z_0) , the worker has to choose a life l that maximizes the present value of wages. Denote by $l(s_0, z_0)$ a feasible life for a worker with initial endowment (s_0, z_0) and by $L(s_0, z_0)$ the space of all feasible lives for this worker. Define the value of a life l for a worker as

$$v(l(s_0, z_0)) = \int_0^\infty e^{-\beta t} w(t, s(t), z(t)) dt.$$

The least upper bound for this value is the value function, which I denote by

$$V_0(s_0, z_0) = \sup_{l \in L(s_0, z_0)} v(l(s_0, z_0)). \quad (3)$$

3.3 Dynamic competitive equilibrium

Before introducing the equilibrium concept, we need a concept of a measure over workers' lives in order to link it to the density of workers in the firm's problem. Let $L = \bigcup_{(s_0, z_0)} L(s_0, z_0)$ be the space of all workers' lives. Let \mathcal{B} be the σ -algebra generated by sets of the form $\{l : l(t_1) \in B_1, \dots, l(t_n) \in B_n\}$, where $0 \leq t_1 < t_2 < \dots < t_n$ 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 $n(t, s, z)$, and

- a wage function $w(t, s, z)$

which fulfills the following conditions:

- Optimality in production: $n(t, s, \cdot)$ maximizes profits $\pi(t, s)$ given wages $w(t, s, \cdot)$ for all $\{(t, s) : t \geq 0, s \leq t\}$.
- Only optimal careers in μ : any set $A \subset L$ fulfilling $l(s_0, z_0) \in A \Rightarrow v(l(s_0, z_0)) < V_0(s_0, z_0)$ has measure zero.
- Normalization: $\int_L d\mu(l) = 1$.
- μ yields n : for every rectangle $B \subset \mathbb{R}^2$, for all $t \geq 0$, we have

$$\int_{\{l:l(t) \in B\}} d\mu(l) = \int_B n(t, s, z) dz ds.$$

- Consistency with initial conditions: for every rectangle $B \subset \mathbb{R}^2$ of the form $s \in [s_l, s_u]$, $z \in [0, z_u]$, we have

$$\mu(\{l : l(0) \in B\}) \leq \nu(B).$$

The last condition ensures that we never use more experience than prescribed by the initial measure ν at $t = 0$. Recall that workers can drop down from their initial position z_0 in vintage s_0 but never move up.

Note that this equilibrium definition automatically precludes any measure μ giving rise to mass points at any point (t, s, z) since we posit the existence of a density function n . Since I assumed complementarity of inputs in the production technology this is unproblematic. Concentrating too many workers at one skill level is automatically discouraged since it drives marginal returns of this skill 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 on subsets of the (t, s, z) -space where n is zero. On these subsets, the wage schedule must be such that firms optimally set labor demand to zero and workers optimally supply zero labor. This will be relevant for old vintages that have become obsolete.

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 vintages of a fixed age $\tau = t - s$, there is always the same number of workers occupying any fixed rung z in the skill ladder:

Definition 3.2. A *stationary competitive equilibrium* is a dynamic competitive equilibrium that satisfies:

$$n(t, t - \tau, z) = n(0, -\tau, z) \quad \text{for all } t > 0 \text{ given any } \tau \geq 0, z \in [0, \tau].$$

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 that total output in the economy grows at rate γ :

$$\begin{aligned} w(t, t - \tau, z) &= e^{\gamma t} w(0, -\tau, z), \\ Y(t, t - \tau) &= e^{\gamma t} Y(0, -\tau), \\ \bar{Y}(t) &= e^{\gamma t} \bar{Y}(0). \end{aligned}$$

3.5 The relationship between μ and n

It is clear by the definition of competitive equilibrium that any measure $\mu(l)$ over lives induces a unique distribution function $n(t, s, z)$. However, it is not quite clear what the relationship looks like in the opposite direction. Given an arbitrary distribution n , does there always exist a measure μ over feasible lives that yields n ? The answer to this question is no: an arbitrary density might require that some workers accumulate experience faster than is feasible. This leads to the next question: what are the requirements on n that ensure 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 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 two different measures may 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: let both measures spread the unit mass uniformly over the rectangle $\{(s, z) : s \in [-2, -1], z \in [0, 1]\}$ for $0 \leq t < 1$, and uniformly over the rectangle $\{(s, z) : s \in [-2, -1], z \in [1, 2]\}$ for all $t \geq 1$. 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 only takes place in the vintages that are from one to two years 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 $t = 1$, 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 $n(t, \cdot, \cdot)$ for all t , but that the careers constituting the measures are by no means the same.

The distribution function n in this example was of an especially simple form. But what happens if we take more elaborate functions n ? 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 *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 for a collection of feasible lives to exist that supports a given density n . The proof shows that μ may be constructed by requiring that workers' career paths never cross. This proof (and all further proofs that are not given in the main text) may be found in the appendix.

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 $n(t, s, z)$ 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 means that workers career paths never cross.*

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 is a degree of freedom in the construction of the measure in subsets of the state space where the level lines of N are strictly decreasing in t . At these points, it would be possible to let the careers of some workers cross in the style of Example 3.1. Crossings cannot be allowed for sets of positive measure, however, whenever N is constant in t .

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 the 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 $V_0(s_0, z_0)$ by the definition of equilibrium. Hence none of the workers would mind to be transferred from the μ' - to the μ_{nc} -universe at $t = 0$. There are no differences between the two universes for firms either since wages are identical. Example 3.1 shows that there will usually exist an entire equivalence class of measures associated with a given n . In the following, I will focus on the no-crossing measure as a representative for this equivalence class.

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 vintage-capital economies considered by Chari & Hopenhayn (1991) and Kredler (2013), it is not obvious that old vintages' productivity must fall behind the frontier in this framework. After all, the gains from factor addition and learning are continuing indefinitely as the vintage ages.

I will now derive a result telling us that under the specified production function, factor addition and learning cannot keep up with total-factor-productivity (TFP) growth of the frontier vintage eventually. The argument will provide the intuition for why learning and factor addition can only postpone but not avoid obsolescence of technologies. Remarkably, this result does not require any assumption on the shape of the learning curve f .

Denote by $L(t, s) \equiv \int_0^{t-s} n(t, s, z) dz$ the labor force in vintage s at t . At t , consider the maximal output in vintage $s \leq t$ that is possible given a unit mass of workers:

$$Y^*(t, s) \equiv \max_{\int_0^{t-s} n(t, s, x) dx \leq 1} Y(t, s). \quad (4)$$

Since there are constant returns to scale in production, this function is an upper bound on vintage productivity, i.e. $Y(t, s)/N(t, s) \leq Y^*(t, s)$ for any n .

In order to achieve the maximum output in (4), marginal productivities across experience levels have to be equalized. This yields the maximizer

$$n^*(t, s, z) = \frac{f(t - s - z)^{\frac{\rho}{1-\rho}}}{\int_0^{t-s} f(t - s - \tilde{z})^{\frac{\rho}{1-\rho}} d\tilde{z}}.$$

Since $\rho \in (0, 1)$, the exponent $\rho/(1 - \rho)$ is always positive. Thus n^* is increasing in skill $h = t - s - z$ (since f is increasing). The more substitutable skills are (the higher ρ), the more the skill distribution n^* is tilted towards the skilled. Indeed, in the limiting case of perfect substitutes ($\rho = 1$) the ideal skill distribution is degenerate and concentrates all workers at the maximally-possible skill $h = t - s$. When making skills more complementary and approaching the Cobb-Douglas case ($\rho \searrow 0$), then the optimal skill mix becomes ever more balanced and skills are used in equal proportions in the limit.

We now use the optimal input schedule n^* from above in the production function to obtain

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

We see that the TFP term $e^{\gamma s}$ 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 we increase vintage age $\tau = t - s$ for fixed t , consider

$$\frac{d \ln Y^*(t, t - \tau)}{d\tau} = -\gamma + \frac{1 - \rho}{\rho} \underbrace{\left[\frac{f(\tau)^{\frac{\rho}{1-\rho}}}{\int_0^\tau f(h)^{\frac{\rho}{1-\rho}} dh} \right]}_{=n^*(t, \tau, \tau)}. \quad (5)$$

We see that the term γ stemming from the exponential decay of TFP in vintage age is constant. Since $\rho \in (0, 1)$, the second term is always positive, and it is proportional to the number of workers used at the highest skill level in the vintage, $h = \tau$. This gives us the intuition for why productivity growth of a technology must vanish eventually. Since skills are complementary, the proportion of workers assigned to the highest skill has to decrease as the vintage ages: low-skilled workers have to complement the top-skill worker at each of the rungs down the skill ladder; the number of these rungs grows larger as the vintage ages, while the optimal input proportions between the ladder rungs stay the same. This means that an ever lower proportion of workers can be assigned to the top skill levels. But it is precisely these top skill levels that yield the learning and factor-addition gains with respect to younger vintages.

To conclude the mathematical argument, note that if f is bounded, then the denominator in the bracketed term will dominate the numerator in (5) eventually and thus $\ln Y^*(t, t - \tau) \rightarrow -\infty$ as $\tau \rightarrow \infty$. In the case that f is unbounded, both the numerator and the denominator go to infinity and we have to invoke L'Hospital's Rule:

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

So also in this case we have $\ln Y^*(t, t - \tau) \rightarrow -\infty$ as $\tau \rightarrow \infty$, and thus productivity approaches zero as $\tau \rightarrow \infty$. Since Y^* is an upper bound on productivity, we have established the following lemma:

Lemma 3.2 (Vanishing productivity of old vintages). *Consider a density n satisfying $L(t, s) = \int_0^{t-s} n(t, s, z) dz > 0$ for all $s \leq t$ given t . Then $\lim_{\tau \rightarrow \infty} Y(t, t - \tau)/L(t, t - \tau) = 0$.*

Remark. It is instructive to compare the case $\rho \in (0, 1)$ to the case of perfect substitutes to see that Lemma 3.2 is a knife-edge result. If $\rho = 1$, then all workers should be assigned to the highest skill level and thus $Y^*(t, t - \tau) = e^{-\gamma\tau} f(\tau)$.¹¹ Then $d \ln Y^*(t, t - \tau)/d\tau = -\gamma + f'(\tau)/f(\tau)$, and thus vintages will not be shut down if the rate of learning $f'(\tau)/f(\tau)$ stays above the rate of technological change γ . This case is different for the following reason: the vintage can assign all workers to the top skill without the marginal product of these workers to decrease. This is not true for $\rho < 1$: there is an Inada condition on each skill level, optimality thus requires to spread workers over all skill levels.

¹¹Formally, we have to consider the production function as a linear functional on the space of measures to do this; the optimal input measure then consists of a mass point of unit mass at the highest skill level.

With this lemma and using an argument from Kredler (2013), it can now be shown that old vintages become obsolete eventually :

Proposition 3.3. (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.: $Y(t, t - \tau) = 0$ for all $\tau > T$.*

In a nutshell, the proof 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 of age $\tau > T$. But Lemma 3.3 shows that very old vintages cannot deliver enough output to provide this value.

The reader may wonder why the complementarities in the production function do not induce firms to maintain vintages alive forever, as is the case in Jovanovic & Yatsenko (2012). It is key here that complementarity in production is between inputs *inside the same vintage*, and not *across different vintages*, as is the case in Jovanovic & Yatsenko (2012). In their setting, capital from different vintages is combined in a CES aggregator, thus there is an Inada condition for each vintage's capital and all vintages are kept alive in equilibrium. In my setting, goods produced by the different vintages are perfect substitutes, which explains the radically different result. The Inada condition in my setting is on skills within a given technology, implying that all rungs of the skill ladder are filled if a vintage is producing positive output. If a vintage is not producing, however, the marginal product of each skill is zero.

4 The planner's problem

4.1 Statement of the problem

Other properties of the equilibrium are better understood from the vantage point of the planner's problem, which I will turn to now.

I will focus on stationary equilibria. Although stationarity implies that the density is time-invariant in the sense that $n(t, t - \tau, z) = n(0, -\tau, z)$, it is important to give the planner the possibility of deviating from the initial density $n(0, \cdot, \cdot)$ when stating the maximization problem. Stationarity is only imposed after deriving the first-order conditions (FOCs) of the planner's problem, meaning that we restrict attention to solutions that satisfy time-invariance. An analogy to the neo-classical growth model may be helpful here to see why we have to give the planner the freedom to deviate from the initial state: when looking for the steady-state level of capital, we do not impose upon the planner to choose a constant level of capital K when stating the problem, but write the capital stock $K(t)$ and its shadow value as functions of time; we then derive the optimality conditions and look for a steady-state level K_{ss} of capital that fulfills these. I thus specify n and the Lagrange multipliers as time-dependent, although they will be constant in a stationary equilibrium (when scaled appropriately in case of the multipliers).

By the construction of the no-crossing measure (Proposition 3.1), we know that any consistent stationary distribution n can be implemented by some measure μ on

(L, \mathcal{B}) , and that any measure yielding the same distribution n gives the same utility to each agent as μ does. I will therefore let the social planner choose the density n instead of a measure μ . 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 for all (s, z) , see again Definition 3.3 and Proposition 3.1. It will prove useful to re-write the density n as the sum of workers who joined the ladder rung z when it was created in the vintage technology plus the additions u to it over time. For vintage s and ladder rung z , we write

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

For ease of exposition, we denote the initial number of workers on rung z in vintage s as

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

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. Using the argument from Lemma 3.3, it is easy to show 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 density n over the set of vintages younger than T for a fixed T . We will later turn to the problem of finding the T^* that is optimal for the planner. This procedure is advantageous because the optimal density n drops down discontinuously to zero at T^* for all z : since there is an Inada condition on each skill within a vintage, n is positive for all z in active vintages and it is zero for all z in inactive vintages. If we had the planner choose n for all $\tau \in [0, \infty)$ instead, we would have to extend the admissible set for the control u to the space of measures in order to engineer the precipitous drop of n at T^* , which unnecessarily complicates the problem. It is thus preferable to split the problem into two steps. This procedure is also clearly valid since $\max_{T, n} J(T, n) = \max_T \{ \max_n J(T, n) \}$ for any criterion $J(\cdot)$.

Given T , define the support of n as $\mathcal{S}_T \equiv \{(t, s, z) : t \geq 0, s \in [t - T, t], z \in [0, t - s]\}$. The planner's problem is then:

$$\max_{n_0 \in \mathcal{N}_0, u \in \mathcal{U}} = \int_0^\infty e^{-\beta t} \int_{t-T}^t e^{\gamma s} \left[\int_0^{t-s} f(t-s-z)^\rho n(t, s, z)^\rho dz \right]^{1/\rho} ds dt \quad (6)$$

$$\text{s.t.} \quad \int_{t-T}^T \int_0^{t-s} n(t, s, z) dz ds \leq 1 \quad \text{for all } t \geq 0, \quad (7)$$

$$\int_0^z u(t, s, \tilde{z}) d\tilde{z} \leq 0 \quad \text{for all } (t, s, z) \in \mathcal{S}_T, \quad (8)$$

$$\text{where } n(t, s, z) = \begin{cases} n(0, s, z) + \int_0^t u(\tilde{t}, s, z) d\tilde{t} & \text{for } s < 0, z \in [0, -s], \\ n_0(s, z) + \int_{s+z}^t u(\tilde{t}, s, z) d\tilde{t} & \text{otherwise,} \end{cases}$$

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

where \mathcal{N}_0 denotes the space of continuous functions on $[-T, \infty) \times [0, T]$ and \mathcal{U} is the space of continuous functions on \mathcal{S}_T . We may call (7) the *total-population constraint* and (8) 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 u .

Note that the problem presented in (6) to (8) is particularly well-behaved: the constraints are linear in n_0 and u , and the objective function is concave in n_0 and u . Thus means that the FOCs are necessary and sufficient.

The Lagrangian for this problem is

$$\begin{aligned} \mathcal{L} = & \int_0^\infty e^{-\beta t} \int_{t-T}^t e^{\gamma s} \left[\int_0^{t-s} f(t-s-z)^\rho n(t,s,z)^\rho dz \right]^{1/\rho} ds dt - \\ & - \int_0^\infty \mu(t) \left[\int_{t-T}^t \int_0^{t-s} n(t,s,z) dz ds \right] dt - \\ & - \int_0^\infty \int_{t-T}^t \int_0^{t-s} \lambda(t,s,z) \left[\int_0^z u(t,s,\tilde{z}) d\tilde{z} \right] dz ds dt. \end{aligned} \quad (9)$$

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 in vintage s on rung 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.3.1 for a formal derivation)

$$\int_{s+z}^{s+T} e^{-\beta t} w(t, s, z) dt = \int_{s+z}^{s+T} \mu(t) dt \quad \text{for all } s \in [-T, \infty], z \in [0, T], \quad (10)$$

where $w(t, s, z)$ is the marginal product of $n(t, s, z)$ in production at time t in vintage s . Equation (10) says that the discounted wage payments over any career without demotions for a worker that stays in the technology until its demise 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 such full-length 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 is equivalent to workers being indifferent between entering different careers in competitive equilibrium at any t . If the value of careers differed, there would be no entry into careers offering inferior values. This also implies that all workers are alike once the initial vintages have disappeared since workers have by then lost any advantage stemming from initial endowments.

When driving $z \rightarrow T$ in (10), i.e. when looking at the workers who enter the vintage shortly before its death, we obtain $\mu(t) = e^{-\beta t} w(t, t-T, T)$.¹² Since workers who

¹²See Lemma A.2 in the appendix for a formal statement.

enter a technology just before it dies do not relax any experience constraints, their marginal productivity must be exactly equal to the shadow value of an additional worker, $\mu(t)$. 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 $\tilde{z} \in [z, t - s]$ at t . Furthermore, it weighs in on the total-population constraints until the demise of vintage s at $s + T$:

$$\int_t^{s+T} e^{-\beta\tilde{t}} w(\tilde{t}, s, z) d\tilde{t} = \int_t^{s+T} \mu(\tilde{t}) d\tilde{t} + \int_z^{t-s} \lambda(t, s, \tilde{z}) d\tilde{z} \quad (11)$$

for all $(t, s, z) \in \mathcal{S}_T$.

By non-negativity of the Lagrange multipliers λ , Equation (11) implies that the value of discounted wages at any point of a career is decreasing in z ; so workers with higher skills are better off. It also means that “promotion weakly dominates demotion”. Workers will never strictly prefer to move down in the hierarchy (they may be indifferent, however).

When adding wages of the worker after the demise of the vintage to both sides of Equation (11), 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 (10) that any concatenation of full careers must yield discounted wages equal to the sum over the multipliers $\mu(t)$ ’s until infinity:

$$\underbrace{\int_t^{s+T} e^{-\beta\tilde{t}} w(\tilde{t}, s, z) d\tilde{t} + \int_{s+T}^{\infty} \mu(\tilde{t}) d\tilde{t}}_{\equiv V(t, s, z)} = \int_t^{\infty} \mu(\tilde{t}) d\tilde{t} + \int_z^{t-s} \lambda(t, s, \tilde{z}) d\tilde{z}. \quad (12)$$

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

Combining equations (11) and (12) will now tell us something about the wages of career starters. Consider two workers: 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 worker, as we 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 worker with the long career must have been lower than $\mu(t)$. The following proposition, which is formally proven in the appendix, summarizes:

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

We can also make statements about wages towards the end of the vintage's life. Letting $t \rightarrow s + T$ in Equation (11) we see that wages must be non-increasing in z , meaning that there is a non-negative premium on experience at T .¹³ This then implies that no worker can experience an increase in earnings when switching to a newer technology in view of Proposition 4.1. The numerical exercise will show that workers usually experience wage *losses* when re-locating. This is remarkable in the sense that in most other models of job choice (such as on-the-job search or occupational choice in span-of-control models), job transitions are associated with earnings increases.

Corollary 4.2 (No wage gains upon vintage change). *In any solution to the planner's problem (i.e. for any T), consider the life of an arbitrary worker. For any point t_n , $n > 0$, where the worker switches between vintages, we have $\lim_{\epsilon \searrow 0} w(t_n - \epsilon, s(t_n - \epsilon), z(t_n - \epsilon)) \geq w(t_n, s(t_n), z(t_n))$.*

Proof. Proposition 4.1 and Lemma A.3 together immediately imply the result. \square

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 for 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 and consistent) 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 (6)-(8). As mentioned before, the global concavity of the problem ensures that the maximizer n^* is unique if the supremum $J^*(T)$ is attained. The following characterization will prove useful to find T^* :

Proposition 4.3. (Wages constant in experience in last vintage) *In a global optimum (T^*, n^*) to the planner's problem, wages are constant in experience in the last vintage:*

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

The proof uses the following argument. If marginal products are not equalized in the final vintage of age T^* , then a firm may be created just behind this final vintage, say of age $T^* + \epsilon$. This firm should re-arrange the skill mix such that marginal products are equalized, which will lead to productivity strictly larger than in the age- T^* vintage. The suitable rearrangement of skills is always feasible when choosing the scale of the firm small enough. Since the marginal product of low-skilled workers in the final vintage equals the marginal value of an additional untrained worker to the planner, extending the life of the vintage in this fashion must be profitable for the planner.

¹³For a formal statement, see Lemma A.3 in the appendix.

This line of argumentation also provides a way to establish 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 4.4. (Lower bound for T^*) *It is never optimal for the planner to choose $T < \bar{T} \equiv \arg \max_{\tau \geq 0} Y^*(0, -\tau)$.*

It is tempting to say that thus the planner should choose $T^* = \bar{T}$, since this would maximize wages $w(t, t - T, \cdot)$ in the final vintage and thus the shadow value of an untrained worker, $\mu(t)$. However, it is not clear that there exists a planner’s solution where wages are constant in the last vintage for $T = \bar{T}$. Numerical exercises show that usually T has to be increased beyond \bar{T} in the optimum.

Building on the previous results, one can now show that the planner’s solution is a competitive equilibrium. The key here is to establish that V as defined by Equation (12) is workers’ equilibrium value function, which in turn requires to show that no-demotion careers and staying in a vintage until its demise is always optimal. See Appendix A.3.8 for the proof and some auxiliary results.

Proposition 4.5. (Planner’s solution is equilibrium) *A stationary global solution (T^*, n^*) to the planner’s problem is a stationary competitive equilibrium with the following properties:*

1. *Wages are equal to marginal products for all $(t, s, z) \in \mathcal{S}_{T^*}$, and they may be chosen to be $w(t, s, z) = e^{\beta t} \mu(t)$ for all $(t, s, z) \notin \mathcal{S}_{T^*}$.*
2. *The no-crossing measure may be chosen for μ .*

Setting wages in inactive vintages equal to those in the dying vintage, $e^{\beta t} \mu(t)$, makes it impossible for firms to break even in technologies older than T^* in view of Proposition 4.4, thus optimal labor demand is zero. Workers are indifferent between staying in a vintage after it has reached age T^* and entering new careers by Equation (12), so labor supply of zero is also consistent with optimality. There is a wide range of other possibilities to choose wages in inactive vintages to achieve market clearing. Furthermore, note that there are many other measures μ on lives that support an equilibrium associated with (T^*, n^*) if the experience constraint does not bind everywhere, as the discussion of the no-crossing measure in Section 3.5 showed.

Obviously, it would be of interest to extend Proposition 4.5 to a full equivalence result, telling us that that (n^*, T^*) is the *only* distribution consistent with competitive equilibrium. Kredler (2013) established a partial converse of Proposition 4.5 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, standard dynamic-programming techniques do not apply.

4.4 Careers and wages

This section derives some further results on careers and the equilibrium wage structure. First, one can show that the premium on skill within a vintage decreases over the lifetime of a vintage. This is mainly due to the fact that the learning curve flattens out as workers accumulate experience. This prediction of the model is consistent with the evidence in Kredler (2013) and Michelacci & Quadrini (2009), who find that the premium on experience is higher in fast-growing industries (new vintages in the model) than in slow-growing ones (old vintages), and that newer establishments exhibit higher experience premia.

Proposition 4.6 (Wage compression). *Consider any solution to the planner's problem (for any T). Define the promotion set $\mathcal{P} = \{(t, s, z) : \lambda(t, s, z) > 0\}$ and the demotion set $\mathcal{D} = \{(t, s, z) : \lambda(t, s, z) = 0\}$. Then wages become more compressed over the lifetime of a vintage in the sense that:*

1. *Inside any open ball $B_\epsilon(t, s, z) \subset \mathcal{P}$, the experience-wage differential $\hat{w}(t, s, z; \Delta z) \equiv \ln w(t, s, z) - \ln w(t, s, z + \Delta z)$ is decreasing in t for $\Delta z \in (0, \epsilon)$. Furthermore, $\hat{w}(t, s, z; \Delta z) \rightarrow 0$ as $t \rightarrow (s + T)$.*
2. *Inside any open ball $B_\epsilon(t, s, z) \subset \mathcal{D}$, w is invariant in z . Thus w is differentiable in z and $\frac{\partial w}{\partial z} = 0$ on $B_\epsilon(t, s, z)$.*

A related question is the following: is it possible that a worker gets paid a lower wage 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. The intuition is as follows: the premium on skill is zero when the vintage dies (see Proposition 4.3), so it cannot be negative at any point before since wages become more compressed over the lifetime of the vintage (see Proposition 4.6).¹⁴

Proposition 4.7 (Wages non-decreasing in experience, form of demotion set). *Suppose that w is continuously differentiable on \mathcal{S}_T . Then in any solution to the planner's problem (for any T):*

$$\lambda(t, s, z) = \int_t^{s+T} -\frac{\partial w(\tilde{t}, s, z)}{\partial z} d\tilde{t} = -\frac{\partial V(t, s, z)}{\partial z}, \quad (13)$$

$$\frac{\partial \lambda(t, s, z)}{\partial t} = \frac{\partial w(t, s, z)}{\partial z} \leq 0. \quad (14)$$

The demotion set $\mathcal{D} = \{(t, s, z) : \lambda(t, s, z) = 0\}$ is of the form $\mathcal{D} = \{(t, s, z) : t \geq d(s, z)\}$ for some function $d(s, z)$.

¹⁴I was not able not prove this stronger result without resorting to assuming differentiability of w . The technical problem is that continuity of the controls only implies continuity, but not differentiability of n and w . This is not sufficiently strong to infer more about the shape of the sets \mathcal{P} and \mathcal{N} . The numerical solutions indicate that these sets are well-behaved (they are connected sets and are separated by a continuous frontier) and that this should not be an issue in practice. For the sake of the formal proof, however, the example of the Cantor Function (or Devil's Staircase), shows that caution is warranted when inferring from a local property that holds on subsets of a domain to global properties of a function, even when the local property holds almost everywhere.

Proof. Equation (13) follows from taking the derivative of Equation (11) in z . Taking the derivative of (13) in t then shows that λ is differentiable in z (invoking continuous differentiability of w) and yields Equation (14).

I now show that the inequality in Equation (14) must hold. Since w is differentiable, we have

$$\frac{\partial w(t, s, z)/\partial z}{w(t, s, z)} = - \lim_{\Delta z \rightarrow 0} \frac{\hat{w}(t, s, z; \Delta z)}{\Delta z}.$$

It then follows from Point 1 in Proposition 4.6 that $(\partial w/\partial z)/w$ is a non-decreasing function in t . This must now be true for *all* (t, s, z) , since we assumed w to be differentiable and either Point 1 or Point 2 of Proposition 4.6 must hold. Finally, since at vintage death we have $\partial w/\partial z|_{t=s+T} = 0$ for all z by Proposition 4.3, we can go backward in time and conclude from the above that $\partial w/\partial z \leq 0$ for all $t < s + T$, all z (since $w > 0$ is positive).

Finally, by Equation (14), this also implies that λ is non-increasing in t for fixed (s, z) . This implies the statement on the shape of the set \mathcal{D} . \square

It is worth to have a closer look to Equations (13) and (14) in Proposition 4.7. Equation (13) 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 (decreasing in z), then the experience constraint holds with equality and no demotions occur. Demotions are only possible when there is no value differential across experience levels inside a vintage.

Equation (14) says that if the skill premium is positive, 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. Thus, for each ladder rung z in vintage s , the experience constraint will be non-binding from some $t \leq s + T$ on, so demotions will only occur towards the end of the vintage's life.

Finally, one can prove that the wage profiles of any two workers who enter different careers at the same point in time must cross. This property has been called “overtaking” in the literature on labor markets (see Hause, 1981). In the numerical solutions we will see that it is always entrants into younger vintages who overtake workers in older vintages: they start off their careers with lower wages, but they are compensated by higher wage later on.

Proposition 4.8 (Overtaking). *Consider two workers a and b who earn wages $w^a(t)$ and $w^b(t)$ in equilibrium. If both start a career at \bar{t} and $w^a(\bar{t}) > w^b(\bar{t})$, then there exists $t' > \bar{t}$ such that $w^a(t') < w^b(t')$.*

Proof. Since both workers could have entered the other worker's career (and lead the subsequent life) at \bar{t} , they must be indifferent between the two lives since both behave optimally in equilibrium. Thus $\int_t^\infty e^{-\beta \tilde{t}} w^a(\tilde{t}) d\tilde{t} = \int_t^\infty e^{-\beta \tilde{t}} w^b(\tilde{t}) d\tilde{t}$. This immediately implies the claim. \square

4.5 The progress curve

Finally, I analyze the connection between the worker's individual learning curve f and his vintage's progress curve. Let $L(t, s) \equiv \int_0^{t-s} n(t, s, z) dz$ denote the total number of workers in vintage s at t , and let $F_s(t) \equiv Y(t, s)/L(t, s)$ be productivity in vintage s at time t .

Proposition 4.9 (Progress curve). *In any global stationary solution to the planner's problem:*

1. *If all experience constraints bind at t in vintage $s < t$, i.e. $N_t(t, s, z) = 0$ for all z , then*

$$\frac{F'_s(t)}{F_s(t)} = \int_0^{t-s} W(t, s, z) \frac{f'(z)}{f(z)} dz + \frac{1}{\rho} W(t, s, t-s) - \frac{n(t, s, t-s)}{N(t, s)},$$

where $W(t, s, z) = \frac{w(t, s, z)n(t, s, z)}{Y(t, s)}$ is skill z 's share of the wage bill.

2. *Vintage progress eventually drops to or below the rate of technical change:*

$$\frac{F'_s(s+T)}{F_s(s+T)} \leq \gamma.$$

Point 1 of the proposition decomposes progress in a vintage into contributions from its different workers. The first term involving the integral is an average over the learning gains of all workers in the vintage, weighted by the contribution of each skill to the technology's total wage bill. This means that the vintage's productivity growth tends to exceed the learning gains of its most experienced workers. This is because the vintage's progress is an average of all individual learning gains f'/f , which are decreasing in the worker's skill. A vintage can thus maintain productivity growth longer than an individual worker because it continues to draw on the fast learning of newly-entering workers even as it ages; individual workers run into decreasing returns of learning more quickly.

The other two terms are correction terms that account for the addition of new workers to the vintage. The first term is positive and captures the benefits of adding new skill inputs to the vintage. The last term adjusts for the increase in the number of workers and is negative. The decomposition is valid while all experience constraints bind; we will see in the numerical example that this is the case for a large range of young technologies.

Point 2 of the proposition tells us that progress in a vintage must eventually fall below the rate of economy-wide technical change. The intuition is simple: if a technology could still outpace aggregate growth by training existing workers and adding new ones, then it would be preferable to reap these gains and to refrain from shutting down the technology. We will see in the numerical example that technologies are actually kept in use for quite some time after their productivity growth has dropped below γ . It turns out that it is worthwhile to reap the gains from the past investments in learning for some time before switching to a new technology.

5 Numerical example

This section presents results from a numerical solution to the planner’s solution, which further illustrate the workings of the model. The solution algorithm is described in Appendix B.

5.1 Baseline example

A unit of time in model is taken to equal one year. I choose standard values for the discount rate ($\beta = 0.04$) and the growth rate of the economy ($\gamma = 0.03$). As for the learning curve, I normalize efficiency units of inexperienced workers to 1 and choose the following parsimonious specification: $f(h) = 1 + A\sqrt{h}$. The slope parameter is set to $A = 0.5$ so that yearly learning gains are approximately 10% for entering workers (see Figure 5). This is on the order of what Mincer regressions find for labor-market entrants’ earnings growth.¹⁵ Finally, I choose $\rho = 0.7$ so that technologies survive for a realistically long time ($T^* = 38.9$), yet don’t peak too early in their development. The most productive vintage is 24.2 years old in the calibration, which is roughly in line with the time it took high-tech companies like Microsoft, Apple and Google to join the ranks of the highest-valued companies. Table 1 shows more statistics of interest for the baseline calibration.

Figure 3 displays the equilibrium wage function and the density as a function of vintage age $\tau = t - s$ and z 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 factor $e^{\gamma t}$. The density stays the same for all t . As Proposition 4.7 tells us, wages are decreasing in experience in (almost) all vintages. However, the premium on experience becomes compressed and vanishes entirely upon vintage’s death, as Proposition 4.6 claimed.

For young vintages wages are strictly increasing in experience (up to vintage age 20 in the baseline example). Workers with higher experience are strictly better off than lower-experience workers, which corresponds to the experience constraint being binding for all z in the planner’s problem in these technologies. Workers are using all experience available to them; as we see in the lower-right panel of Figure 3, there are no demotions in these vintages.

However, at some point wage compression has advanced to the point where the wage schedule would become decreasing in skill if no demotions occurred. This is possible since late entrants are scarcer than early entrants – entry tends to decrease as the vintage ages, as the lower-right panel of Figure 4 shows. This effect tilts the wage structure in favor of lower skills, and thus counteracts the learning curve.

In the numerical examples this flattening in the wage hierarchy first occurs for intermediate points in the skill hierarchy, in the baseline calibration for $z \simeq 15$. We can clearly see this from the plot of demotions in the lower-right panel of Figure 3. Interestingly, in all numerical examples, the first demotions occur in the heyday of the technology, i.e. around the point in time when the vintage dominates all other

¹⁵See Heckman et al. (2003), for example.

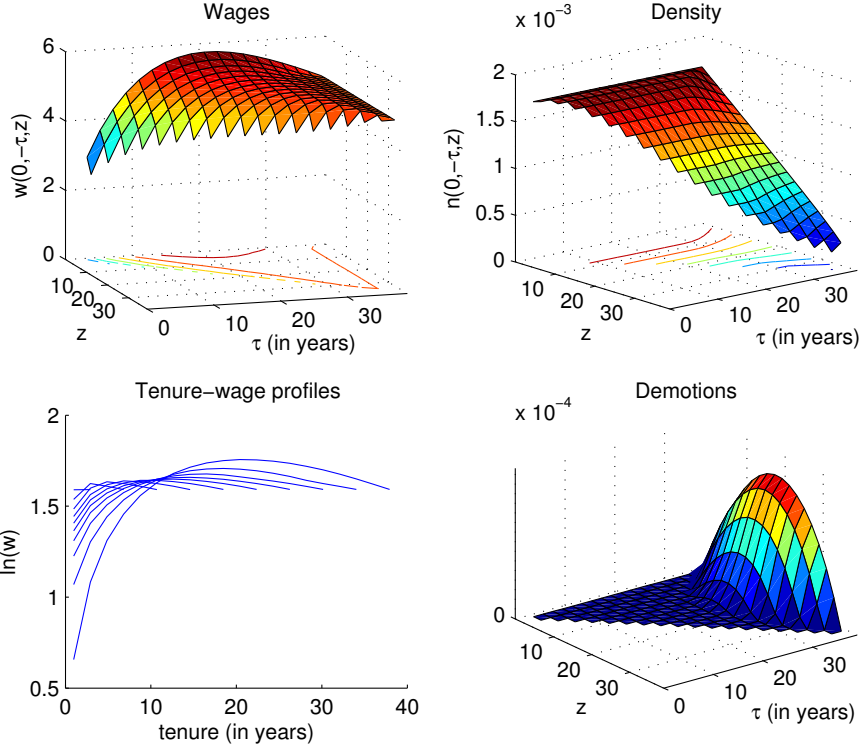


Figure 3: Equilibrium variables

Vintages are labeled by their age $\tau = t - s$ in years. All variables are evaluated at $t = 0$ in the stationary equilibrium. Wages grow at rate γ over time, density and demotions are time-invariant. The tenure-wage profiles depict $\ln[e^{-\gamma t} w(t, s_0, z_0(t))]$ for different careers (s_0, z_0) starting at $t_0 = 0$ as a function of tenure $t \in [0, t_1)$.

vintages in terms of productivity (see the lower-left panel of Figure 4 for a plot of vintage productivity by age). In the baseline calibration, this is for vintage age $\tau \simeq 25$.

Since high-skill workers always have the option to move down in the hierarchy, however, the wage structure never tilts in favor of the lower skills. Instead, there are just enough demotions to maintain the wage hierarchy flat in z . Correspondingly, the experience constraint is slack in this subset of the state space in the planner's problem, as Proposition 4.6 tells us. As the vintage ages, the constraint binds for fewer and fewer skill levels, and eventually demotions are ubiquitous as $\tau \rightarrow T^*$. Corresponding to this, we see that the wage hierarchy is completely flattened out in these moribund vintages.

The lower left panel of Figure 3 shows the log age-earnings profiles when following an agent over her career in a vintage, discounted by economic growth. Workers who enter young technologies have the lowest entry wages but are compensated by high wages in the middle of their careers. Indeed, the baseline model predicts that workers who entered the frontier vintage are the top earners in the economy 20 years later.

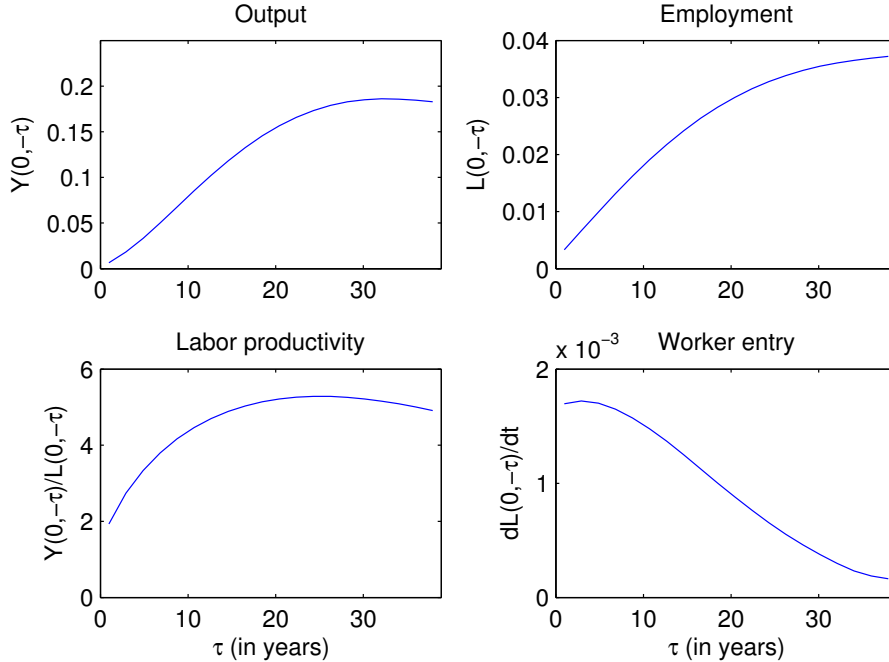


Figure 4: Vintage characteristics

Vintage output $Y(0, -\tau)$ and total vintage employment $L(0, -\tau) = \int_0^\tau n(0, -\tau, z) dz$ are evaluated at time $t = 0$. In the stationary equilibrium, output and productivity grow at rate γ ; employment and entry are time-invariant.

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 are induced by the low level of entry in such vintages (see the lower-right panel of Figure 4). This also means that late entrants have flatter wage profiles.

The earnings profiles exhibit a pronounced hump shape, especially for early entrants. Why is this? First, young technologies have lower productivity than older ones, which keeps wages low (see the lower-left panel of Figure 4). Once the technology matures and becomes more productive, early entrants possess a scarce skill in a productive technology and thus command high wages, in addition to having progressed on their learning curves.

However, as returns to learning slow down for a large number of workers, skill becomes abundant and the premium on it decreases, making the earnings profile bend down.¹⁶ Another force behind this flattening is that vintage productivity growth slows down as the vintage ages.

Finally, the model has interesting predictions on output, employment and productivity of vintages. Figure 4 shows that employment increases as the vintage ages;

¹⁶If wages are not discounted by economic growth, then their slope is always positive in the numerical examples, but they are still concave. So the model does not generate wage losses during workers' careers, as Kredler's (2013) does.

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 become available to the technology as it ages (the highest skill levels). 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.

Figure 5 compares a vintage's growth rate of productivity, $F'_0(\tau)/F_0(\tau)$, with the learning curve (the growth rate of efficiency units) of the first entrants into this vintage, $f'(\tau)/f(\tau)$. As Point 1 in Proposition 4.9 suggested, the vintage can maintain higher growth rates than individual workers since it can continuously draw on new workers who are not subject to decreasing returns to learning and because it increases its skill range. The baseline model predicts that vintage progress is almost twice as large as the individual learning rate throughout the technology's life. Finally, we see that productivity growth falls below the rate of technical change in the economy, as predicted by Point 2 of Proposition 4.9, and that it stays below this benchmark for almost 15 years. Despite these low growth rates, however, old vintages are among the most productive in the economy, as Figure 4 showed.

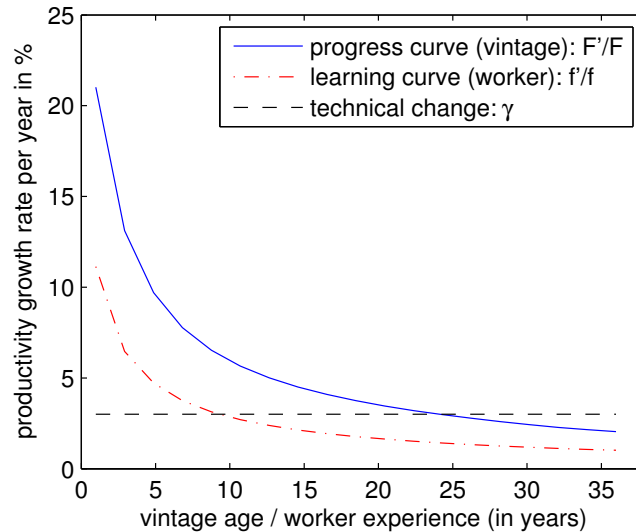


Figure 5: Progress curve vs. learning curve

5.2 Comparative statics

How do the predictions of the model change if there are changes to the production function, e.g. if different skills are more substitutable? What happens if there is faster technological growth or faster learning? This section answers these questions by carrying out comparative-statics exercises with respect to the model parameters. Table 1 shows how the model's (steady-state) predictions change when we increase

the parameters ρ , γ , β or A , one at a time.

	T^* (oldest vintage)	age of most productive vintage	entry into vintages age $< T^*/2$	average career length	st.dev. of $\ln w$
baseline	38.9	24.2	76%	14.0	0.61
$\rho = 0.8$	30.1	18.2	87%	8.4	0.41
$\gamma = 0.035$	32.5	20.4	75%	11.8	0.57
$\beta = 0.06$	40.6	24.2	76%	14.6	0.60
$A = 1$	43.5	26.7	82%	14.0	0.76

Table 1: Comparative statics

Learning curve: $f(h) = 1 + A\sqrt{h}$. The baseline model's parameters are $\gamma = 0.03$, $\beta = 0.04$, $\rho = 0.7$ and $A = 0.5$. The four rows below the baseline results report results when increasing one parameter at a time with respect to the baseline model. The statistics are the age of the oldest vintage (T^*) in years, the age of the most productive vintage ($\arg \max_{\tau} Y(0, -\tau)$) in years, the proportion of worker entry going into vintages of age $\tau < T^*/2$, the average career length $[t_{n+1} - t_n]$ in years, and the standard deviation of log-wages over all workers. All statistics are time-invariant in the stationary equilibrium.

First, consider the effects of changes to the production function. Making workers more substitutable by increasing ρ from 0.7 to 0.8 has dramatic effects. Vintages are phased out almost 9 years earlier in the stationary equilibrium, and the most productive vintage is now 18.2 instead of 24.2 years old. Also, more workers (87%) enter vintages that are less than half as old as the last vintage than in the baseline calibration (76%). Why are workers now more reluctant to enter old vintages? The reason is that fewer low-skill workers are needed in old vintages to complement high-skill workers since complementarities are weaker. In the extreme case of a linear production function ($\rho = 1$), indeed, entry would be fully concentrated in the newest vintage.¹⁷ Weaker complementarities also imply that the addition of new inputs increases the vintage's productivity by less, as Point 1 in Proposition 4.9 shows. This reduces the gains from prolonging a technology's life and thus leads to quicker shut-down. Indeed, the numerical experiments show that the progress curve for $\rho = 0.8$ starts at about the same level as in Figure 5, but then drops faster, the learning curve remaining unchanged. Since there is less productivity growth inside a vintage, also workers' returns to experience in a technology decrease. This brings down overall wage dispersion considerably, as we see in the last column of Table 1.

The second experiment shows the consequences of an increase of the pace of technical change from 3.0% to 3.5% yearly. Since the frontier technology is now improving faster, the opportunity cost of staying with old technologies rises, and vintages become obsolete 6.4 years earlier than in the baseline scenario. Consistent with this, workers' careers shorten by more than 2 years on average. In Figure 5, the progress curve is essentially unchanged, but the dashed reference line (γ) shifts up. The intersection between the two shifts left, meaning that the vintage with the

¹⁷See Kredler (2013) for a proof in a setting with endogenous human-capital accumulation.

highest labor productivity is around 4 years younger. Wage dispersion decreases somewhat, but the quantitative effect is not strong. This is because there are two competing effects: on the one hand, the career-earnings profiles are flattened because older vintages are penalized more by their lower TFP – this dampens wage dispersion. On the other hand, there is a larger mass of workers employed in young technologies where wages are low, which increases dispersion.

The third experiment considers an increase in β from 4% to 6%. This can be interpreted as an increase in the interest rate, which makes workers discount future earnings more. It turns out that the quantitative effects of this change are modest. Since there is no change to production technology, the shape of the progress curve and the wage structure for a given vintage remain almost unchanged with respect to the baseline. This is evident in the most productive vintage’s age being unaffected. However, since workers are essentially more impatient, they now shy away from entry into young vintages with low entry wages and back-loaded returns. When faced with the decision of switching technologies, they prefer staying in old vintages with decent wages in the present to entering young technologies, which offer high wages only in the far future after an initial spell of low wages. As a consequence, vintages are phased out 1.7 years later and careers are 0.6 years longer on average. Wage dispersion goes slightly down since there is now a larger range of old vintages with a flat wage structure.

Finally, I consider an increase in the slope of the learning curve from $A = 0.5$ to $A = 1.0$. The learning curve shifts up in Figure 5, and in consequence the progress curve does so to. The most productive vintage is now 2.5 years older than in the baseline economy. In the left two panels of Figure 3, output and productivity of a vintage keep growing for a longer time. Vintages shut down 4.6 years later, and tenure-wage profiles peak later and become steeper. Entry into younger vintages becomes more attractive. All this manifests itself in a sharp increase in wage dispersion.

6 Conclusions

I conclude with a brief comparison of the learning-curve approach to the human-capital approach. There is a large literature that studies the implications of endogenous human-capital accumulation, which stands in contrast to the exogenous learning curve used in this paper. Particularly related is Kredler (2013), who studies a vintage setting in which skill accumulation is endogenous. His models has two important predictions that are in line with the current paper, but which rely importantly on the human-capital-accumulation mechanism. First, *overtaking* occurs as entrants into new technologies accumulate human capital faster than entrants into older technologies . Second, *wage compression* takes place inside vintages as entrants invest more in human capital than incumbents, the high-skilled being held back by an upper bound on human capital. The current paper shows that neither overtaking nor wage compression depend on the particular assumptions made on skill accumulation, but that another force is paramount in world where skill is technology-specific: the *relative*

scarcity of skill. The key mechanism is that as a technology ages, skill becomes more abundant and the premium on it decreases.

References

- Abowd, J. M., Kramarz, F. & Margolis, D. N. (1999), ‘High wage workers and high wage firms’, *Econometrica* **67**(2), 251–333.
- Bartel, A. & Lichtenberg, F. (1987), ‘The skill distribution and competitive trade advantage of high-technology industries’, *Advances in Industrial and Labor Relations* **4**, 161–76.
- Bartel, A. & Lichtenberg, F. (1991), ‘The age of technology and its impact on employee wages’, *Economics of Innovation and New Technology* **1**(3), 215–231.
- Barucci, E. & Gozzi, F. (1998), ‘Investment in a vintage capital model’, *Research in Economics* **52**(2), 159–188.
- Barucci, E. & Gozzi, F. (2001), ‘Technology adoption and accumulation in a vintage-capital model’, *Journal of Economics* **74**(1), 1–38.
- Baruch, Y. & Bozionelos, N. (2011), Career issues, in S. Zedeck, ed., ‘APA handbook of industrial and organizational psychology, Vol 2: Selecting and developing members for the organization.’, American Psychological Association, pp. 67–113.
- Bellman, R. E. & Cooke, K. L. (1963), ‘Differential-difference equations’, *Academic, New York* .
- Ben-Porath, Y. (1967), ‘The production of human capital and the life cycle of earnings’, *The Journal of Political Economy* **75**(4), 352–365.
- Benhabib, J. & Rustichini, A. (1991), ‘Vintage capital, investment, and growth’, *Journal of Economic Theory* **55**(2), 323–339.
- Benhabib, J. & Rustichini, A. (1993), A vintage capital model of investment and growth: Theory and evidence, in R. A. Becker & L. W. McKenzie, eds, ‘General Equilibrium, Growth, and Trade: The Legacy of Lionel McKenzie’, London: Academic.
- Boucekkine, R., Del Rio, F. & Licandro, O. (2003), ‘Embodied technological change, learning-by-doing and the productivity slowdown’, *The Scandinavian Journal of Economics* **105**(1), 87–98.
- Boucekkine, R., del Rio, F. & Licandro, O. (1999), ‘Endogenous vs exogenously driven fluctuations in vintage capital models’, *Journal of Economic Theory* **88**(1), 161–187.

- Boucekkine, R., Germain, M. & Licandro, O. (1997), ‘Replacement echoes in the vintage capital growth model’, *Journal of Economic Theory* **74**(2), 333–348.
- Boucekkine, R., Hritonenko, N. & Yatsenko, Y. (2011), ‘Scarcity, regulation and endogenous technical progress’, *Journal of Mathematical Economics* **47**(2), 186–199.
- Brokate, M. (1985), ‘Pontryagin’s principle for control problems in age-dependent population dynamics’, *Journal of Mathematical Biology* **23**(1), 75–101.
- Chari, V. V. & Hopenhayn, H. (1991), ‘Vintage human capital, growth, and the diffusion of new technology’, *Journal of Political Economy* **99**(6), 1142–65.
- Cooley, T. F., Greenwood, J. & Yorukoglu, M. (1997), ‘The replacement problem’, *Journal of Monetary Economics* **40**(3), 457–499.
- Feichtinger, G., Hartl, R. F., Kort, P. M. & Veliov, V. M. (2006*a*), ‘Anticipation effects of technological progress on capital accumulation: a vintage capital approach’, *Journal of Economic Theory* **126**(1), 143–164.
- Feichtinger, G., Hartl, R. F., Kort, P. M. & Veliov, V. M. (2006*b*), ‘Capital accumulation under technological progress and learning: a vintage capital approach’, *European Journal of Operational Research* **172**(1), 293–310.
- Feichtinger, G., Hartl, R. F., Kort, P. M. & Veliov, V. M. (2008), ‘Financially constrained capital investments: the effects of disembodied and embodied technological progress’, *Journal of Mathematical Economics* **44**(5), 459–483.
- Feichtinger, G., Tragler, G. & Veliov, V. M. (2003), ‘Optimality conditions for age-structured control systems’, *Journal of Mathematical Analysis and Applications* **288**(1), 47–68.
- Galor, O. & Tsiddon, D. (1997), ‘Technological progress, mobility, and economic growth’, *American Economic Review* **87**(3), 363–82.
- Goetz, R.-U., Hritonenko, N. & Yatsenko, Y. (2008), ‘The optimal economic lifetime of vintage capital in the presence of operating costs, technological progress, and learning’, *Journal of Economic Dynamics and Control* **32**(9), 3032–3053.
- Hause, J. C. (1981), ‘The fine structure of earnings and the on-the-job training hypothesis’, *Econometrica* **49**(1), 277.
- Heckman, J. J., Lochner, L. & Todd, P. E. (2003), Fifty years of Mincer earnings regressions, IZA Discussion Papers 775, Institute for the Study of Labor (IZA).
- Hritonenko, N. & Yatsenko, Y. (2005), ‘Turnpike and optimal trajectories in integral dynamic models with endogenous delay’, *Journal of Optimization Theory and Applications* **127**(1), 109–127.

- Jovanovic, B. (1998), ‘Vintage capital and inequality’, *Review of Economic Dynamics* **1**(2), 497–530.
- Jovanovic, B. & Nyarko, Y. (1996), ‘Learning by doing and the choice of technology’, *Econometrica* **64**(6), 1299–1310.
- Jovanovic, B. & Yatsenko, Y. (2012), ‘Investment in vintage capital’, *Journal of Economic Theory* **147**(2), 551–569.
- Kredler, M. (2013), ‘Experience vs. obsolescence: A vintage-human-capital model’, *Journal of Economic Theory* **in print**.
- Luenberger, D. G. (1973), *Optimization Using Vector Space Methods*, AddisonWesley, New York, NY, USA.
- Malcomson, J. M. (1975), ‘Replacement and the rental value of capital equipment subject to obsolescence’, *Journal of Economic Theory* **10**(1), 24–41.
- Michelacci, C. & Quadrini, V. (2009), ‘Financial markets and wages’, *The Review of Economic Studies* **76**(2), 795–827.
- Øksendal, B. (2003), *Stochastic differential equations*, Springer.
- Parente, S. L. (1994), ‘Technology adoption, learning-by-doing, and economic growth’, *Journal of Economic Theory* **63**(2), 346–369.
- Prskawetz, A. & Veliov, V. M. (2007), ‘Age-specific dynamic labor demand and human capital investment’, *Journal of Economic Dynamics and Control* **31**(12), 3741–3777.
- Saglam, C. & Veliov, V. M. (2008), ‘Role of endogenous vintage specific depreciation in the optimal behavior of firms’, *International Journal of Economic Theory* **4**(3), 381–410.
- Solow, R. M., Tobin, J., von Weizsäcker, C. C. & Yaari, M. (1966), ‘Neoclassical growth with fixed factor proportions’, *The Review of Economic Studies* **33**(2), 79–115.
- Thompson, P. (2010), ‘Learning by doing’, *Handbook of the Economics of Innovation* **1**, 429–476.
- van Hilten, O. (1991), ‘The optimal lifetime of capital equipment’, *Journal of Economic Theory* **55**(2), 449–454.
- Violante, G. L. (2002), ‘Technological acceleration, skill transferability, and the rise in residual inequality’, *Quarterly Journal of Economics* **117**(1), 297–338.
- Yelle, L. E. (1979), ‘The learning curve: Historical review and comprehensive survey’, *Decision Sciences* **10**(2), 302–328.

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 $N(t, s, z_{nc}(t)) = c = N(t_n, s, t_n - s)$ 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 $N(t_n, s, t_n - s)$ of workers left in the vintage. Figure A.1 shows the no-crossing careers for a stationary density where vintages are identified by age $\tau = t - s$. For a stationary density, the density over the (τ, z) -triangle is constant and we can thus depict all no-crossing careers in a two-dimensional graph. The figure uses the notation $\bar{N}(\tau, z) \equiv N(t, t - \tau, z)$.

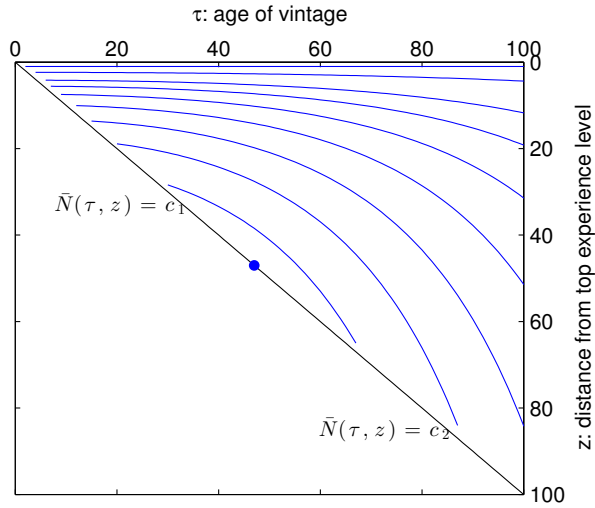


Figure A.1: Construction of no-crossing measure

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. 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 A.1, 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 .¹⁸

¹⁸Notice that even when n 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 have to end, and the worker

We are now able to decide for any given number of rectangles $\{A_n\}_{n=1}^N \in \mathbb{R}^2$ if a given no-crossing life l_{nc} passes these rectangles at a given sequence of dates $t = t_1, \dots, t_N$, which is necessary for the construction of the measure μ . However, we also need to be able to specify which mass of lives passes through these rectangles 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

$$\mu(B) = \nu(\{(s_0, z_0) : \text{there exists } l_{nc} \in B \text{ s.t. } l_{nc}(0) = (s_0, z_0)\}).$$

This measure obviously yields n 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 Definition 2.1 if N is a non-increasing function in t at all points (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 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 workers located at ladder rung z or higher at time t than before, so some workers must have violated the non-increasingness constraint on z . This concludes the proof. \square

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 Øksendal (2003) 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, \dots, t_n , which would require a similar construction as in the proof given here.

A.2 Proof for Proposition 3.3 (finite support of technologies)

Proof. 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 since production is positive. If we see workers staying in an old vintage, then they must be better off doing this than obtaining $e^{\gamma t} \delta$. In other words, for any career $z_n(\cdot)$ in vintage s_n over the interval $[t_n, t_{n+1})$ chosen in equilibrium, for any $t \in [t_n, t_{n+1})$ we must have

$$\int_t^{t_{n+1}} e^{-\beta \tilde{t}} w(\tilde{t}, s_n, z_n(\tilde{t})) d\tilde{t} \geq \int_t^{t_{n+1}} \delta e^{(\gamma - \beta)\tilde{t}} d\tilde{t}, \quad (\text{A.1})$$

otherwise the worker should leave vintage s_n at t and work for wage $e^{\gamma \tilde{t}} \delta$. Consider now all career segments that occur in vintages of age T and older: i.e. for any career satisfying $t_n - s_n > T$ take the segment $[\max\{t_n, s_n + T\}, t_{n+1})$. When integrating up the inequality (A.1) over all such career segments, we find

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

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 $T > 1/c$.

where we invoke stationarity of the equilibrium. 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 for wage δ .

Now suppose that the support of technologies was not finite. Then for any fixed $\delta > 0$, per-worker output in vintages $\tau \geq T$ can be pressed below δ when choosing T large enough by Lemma 3.2. But this means that the above inequality will be violated for T large enough, so workers do not behave optimally and thus the density n must have finite support. \square

A.3 The planner's problem

A.3.1 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 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 (9) is

$$\int_{-T}^{\infty} \int_0^T \left[\int_{s+z}^{s+T} e^{-\beta t} w(t, s, z) - \mu(t) dt \right] h_0(s, z) dz ds = 0, \quad (\text{A.2})$$

and it must equal zero in the optimum. Similarly, perturbing the optimal $u(t, s, z)$ by some $h(t, s, z) \in \mathcal{U}$, we obtain

$$\begin{aligned} \int_{S_T} \left[\int_t^{s+T} e^{-\beta \tilde{t}} w(\tilde{t}, s, z) d\tilde{t} \right] h(t, s, z) &= \int_{S_T} \left[\int_t^{s+T} \mu(\tilde{t}) d\tilde{t} \right] h(t, s, z) \\ &+ \int_{S_T} \left[\int_z^{t-s} \lambda(t, s, \tilde{z}) d\tilde{z} \right] h(t, s, z). \end{aligned} \quad (\text{A.3})$$

Since Equations (A.2) and (A.3) must hold for any perturbations h_0 and h , we obtain the first-order conditions (10) and (11) as claimed in the text. If any of the two was not true at any point, the equalities could be violated, respectively, by choosing h_0 or h zero except for a small neighborhood around this point.

A.3.2 Continuity of n and w

Lemma A.1 (n and w continuous). *The functions n and w are continuous for any controls $n_0 \in \mathcal{N}_0$, $u \in \mathcal{U}$.*

Proof. n is continuous since it is an integral over the continuous functions u and n_0 , see the statement of the planner's problem in (6)-(8). w then inherits continuity from n and f , see the wage equation (2). \square

A.3.3 Shadow value $\mu(t)$

Lemma A.2 (Shadow value of inexperienced worker). *In any solution to the planner's problem (i.e. for any T), for all t we have*

$$\mu(t) = e^{-\beta t} w(t, t - T, T).$$

Proof. Taking the limit of (10) as $z \rightarrow T$ implies the result since w is a continuous function by Lemma A.1. \square

A.3.4 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 (12) tells us that the infinitesimal change in value at the beginning of the career is given by $dV/d\epsilon|_{\epsilon=0} = -e^{-\beta t} w(t, s, t - s)$, where we invoke continuity of w (see Lemma A.1).

We will now compare this to how the outside option, i.e. the value of a career starter, changes over time. Define $W(\epsilon) = \int_{t+\epsilon}^{\infty} \mu(\tilde{t}) d\tilde{t}$, which by Equation (10) is the value for an agent without experience. We have $dW/d\epsilon|_{\epsilon=0} = -\mu(t)$. Since $V(t, s, z) \geq W(t)$ for all $(t, s, z) \in \mathcal{S}_T$ by non-negativity of the Lagrange multipliers in Equation (12), it must be that $dV/d\epsilon|_{\epsilon=0} \geq dW/d\epsilon|_{\epsilon=0}$. Using the derived expressions for the derivatives, this implies the claim $w(t, s, t - s) \leq e^{\beta t} \mu(t)$. To conclude the proof, observe that $w(t, t - T, T) = e^{\beta t} \mu(t)$ by Lemma A.2. \square

A.3.5 Wages at T

Lemma A.3 (Wages at T). *In any solution to the planner's problem (for any T), $w(t, t - T, z)$ is non-increasing in z for all t .*

Proof. Since w is continuous by Lemma A.1, taking the limit of Equation (11) as $t \rightarrow s + T$ implies the claim by non-negativity of the Lagrange multipliers λ (note that the integral over μ vanishes). \square

A.3.6 Proof for Proposition 4.3 (wages constant in experience in last vintage)

Proof. Fix some t . Suppose, by way of contradiction, that marginal products were not constant across z in vintage $t - T^*$ at t . This implies that per-worker output in vintage $t - T^*$ stays below its unconstrained optimum $Y^*(t, t - T^*)$ as defined in (4) – recall that marginal productivities have to be equalized to attain this maximum. Since $Y^*(t, s)$ is a continuous function, there exists $\delta > 0$ such that $Y^*(\tilde{t}, t - T^*) > Y(t, t - T^*)/L(t, t - T^*)$ for all $\tilde{t} \in (t, t + \delta]$, i.e. it is feasible to attain higher productivity with vintage $t - T^*$ slightly after t .¹⁹

Consider now the strategy of retaining a small mass of workers in vintage $t - T^*$ in the time interval $\tilde{t} \in (t, t + \delta]$ to obtain optimal productivity $Y^*(\tilde{t}, t - T^*)$ on this interval. I will now argue that it is feasible to obtain output per worker arbitrarily close to $Y^*(\tilde{t}, t - T^*)$ for $\tilde{t} \in (t, t + \delta]$.

First, we will see that it is always possible to implement a maximizing experience mix in an arbitrary short amount of time, i.e. one can achieve productivity $Y^*(t + \epsilon, t - T^*)$ for any $\epsilon > 0$ given any experience distribution $n(t, t - T^*, \cdot)$ at shut-down. Note first that $n(t, t - T^*, \cdot)$ must be lower-bounded, i.e. there exists $c > 0$ such that $n(t, t - T^*, z) \geq c$ for all z . If this was not the case, marginal productivity would grow infinite for some z , which is clearly not optimal. It must then be possible to choose some small $b > 0$ and set $n(t + \epsilon, t - T^*, z) = bn^*(t + \epsilon, t - T^*, z) < c$ for all z so that output per worker is maximal; this is clear from the experience constraint (8).

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 \tilde{n} for vintage $t - T^*$ that leaves $\tilde{n}(\tilde{t}, t - T^*, z) = n(\tilde{t}, t - T^*, z)$ unchanged for $\tilde{t} \leq t$ (for all z), but then jumps to the optimal skill mix, i.e. $\tilde{n}(\tilde{t}, t - T^*, z) = bn^*(\tilde{t}, t - T^*, z)$ for some constant $b > 0$ for all $\tilde{t} > t$ (all z). But \tilde{n} is not a feasible policy since the planner is restricted to use continuous

¹⁹We have used the notation $L(t, s) = \int_0^{t-s} n(t, s, z) dz$ for the labor force in vintage s at t here.

density functions. However, \tilde{n} may be arbitrarily well approximated using a continuous policy u , which means that a smooth transition can 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. So we may choose ϵ arbitrarily small.

Since ϵ may be chosen arbitrary small compared to δ , we can construct a deviation such that workers in vintage $(t - T^*)$ increase the planner's criterion for all $\tilde{t} \in (t, t + \delta]$. By continuity, they are more productive than the frontier vintage at \tilde{t} , i.e. $w(\tilde{t}, t - T^*, z) > Y(\tilde{t}, \tilde{t} - T^*)/N(\tilde{t}, \tilde{t} - T^*)$ for all z since wages are aligned. Since wages are non-increasing in z in the final vintage by Lemma A.3, $w(\tilde{t}, \tilde{t} - T^*, T^*) \leq Y(\tilde{t}, \tilde{t} - T^*)/N(\tilde{t}, \tilde{t} - T^*) < w(\tilde{t}, t - T^*, z)$ for the lowest skill level in the final vintage. By Lemma A.2, $w(\tilde{t}, \tilde{t} - T^*, T^*) = e^{\beta \tilde{t}} \mu(\tilde{t})$, thus $e^{-\beta \tilde{t}} w(\tilde{t}, t - T^*, z) > \mu(\tilde{t})$ for all z . This means that the marginal marginal contribution $e^{-\beta \tilde{t}} w(\tilde{t}, t - T^*, z)$ by workers in vintage $t - T^*$ at \tilde{t} to the planner's criterion is strictly larger than $\mu(\tilde{t})$, which is the marginal value of an additional untrained worker. This shows that we have found a profitable deviation from the proposed maximizer (T^*, n^*) and it cannot be a global optimum, which concludes the proof. \square

A.3.7 Proof for Corollary 4.4 (lower bound for T^*)

Proof. If $T < \bar{T}$, then there exists $\delta > 0$ such that for all $\tilde{t} \in (t, t + \delta]$ we have $Y^*(\tilde{t}, t - T) > Y(t, t - T)$, for fixed t . Then by the same argument as in the proof for Proposition 4.3, T cannot be globally optimal. \square

A.3.8 Proof for Proposition 4.5 (planner's solution is equilibrium)

Definition A.1 (No-demotion career). We call $\{s_n, t_n, t_{n+1}, z_n(\cdot)\}$ a *no-demotion* career if $z_n(t) = z_n(t_n)$ for all $t \in (t_n, t_{n+1})$.

Definition A.2 (Full-length career). Consider the planner's problem for a given T . We call $\{s_n, t_n, t_{n+1}, z_n(\cdot)\}$ a *full-length* career if $t_{n+1} = s_n + T$.

Lemma A.4 (Value for inexperienced worker). *Consider a solution to the planner's problem (for any T) and the wage function implied by this solution on \mathcal{S}_T . For an inexperienced worker with initial conditions (s_0, z_0) such that $z_0 = -s_0$, any life consisting of full-length no-demotion careers yields value $\int_0^\infty e^{-\beta t} \mu(t) dt$.*

Proof. For a full-length no-demotion career, Equation (12) implies $\int_{t_n}^{t_{n+1}} e^{-\beta t} w(t, s_n, t_n - s_n) dt + \int_{t_n}^{t_{n+1}} e^{-\beta t} \mu(t) dt$ since the range of integration in the integral over λ vanishes. Concatenating careers and taking the limit $t \rightarrow \infty$ yields the desired result. \square

Lemma A.5 (Promotion always optimal). *Consider a stationary solution to the planner's problem (for any T). Given the wage function w implied by this solution on \mathcal{S}_T and wages $w(t, s, z) = \mu(t)$ for all $(t, s, z) \notin \mathcal{S}_T$, any life l is weakly dominated by a corresponding no-demotion life l' , where we set $s'(t) = s(t)$ for all t and $z'_n(t) = z_n(t_n)$ for $t \in [t_n, t_{n+1})$, for all n .*

Proof. Consider an arbitrary career $z : [0, d) \rightarrow \mathbb{R}$ in a vintage $s_0 \leq 0$ – by stationarity of the solution having the career start at $t = 0$ is without loss of generality. Since the function z is decreasing and measurable, we can approximate it arbitrarily well by a sequence of step careers $\{z^{(k)}\}_{k=1}^\infty$ that are constant on disjoint intervals, see the blue dashed line in Figure A.2 for an example. Construct career $z^{(k)}$ as follows. Let the height of the steps be equally spaced between $z(0)$ and $z(d)$, i.e. define a decreasing sequence of numbers $z_i^{(k)} =$

$z(0) - i[z(0) - z(d)]/k$ for $i = 0, 1, \dots, k$. Then choose the t -coordinates of the steps as $t_i^{(k)} = \sup\{t : z(t) \geq z_{i+1}^{(k)}\}$. The step career is then defined by setting $z^{(k)}(t) = z_i^{(k)}$ if $t \in [t_i^{(k)}, t_{i+1}^{(k)})$. In Figure A.2, the blue dashed line shows an example with $k = 4$ steps.

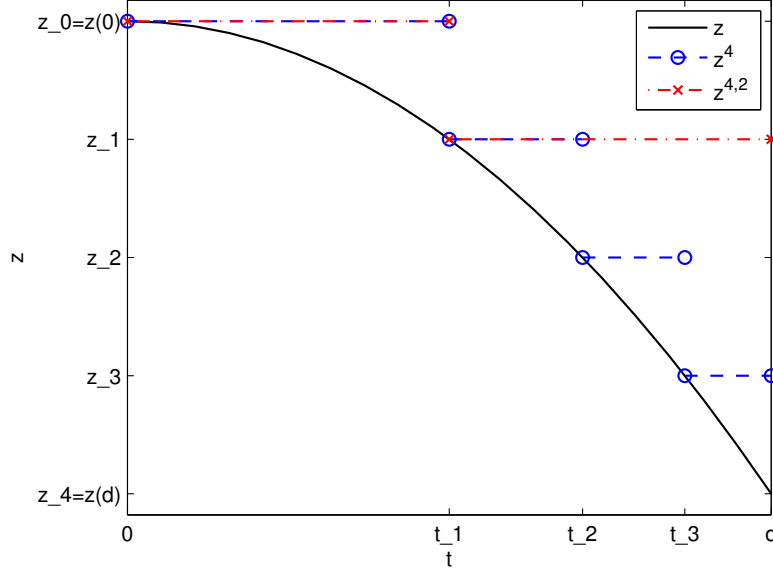


Figure A.2: Step careers

We will now show that the no-demotion career $z^{(1)}$ which sets $z^{(1)}(t) = z(0) \forall t \in [0, d]$ is at least as good as any of the step careers $z^{(k)}$, $k \geq 2$, in a planner's solution. To do this, define a sub-sequence of careers $z^{(k,i)}$ which is identical to $z^{(k)}$ on the first i steps but then stays constant. See Figure A.2 for the example $z^{(4,2)}$ (the dash-dotted red line).

For fixed k , we now note that staying on a higher ladder rung must always be weakly preferred to being on a lower rung. In active vintages, i.e. for $d \leq s_0 + T$, we have $\int_{t_i^{(k)}}^d e^{-\beta t} w(t, s_0, z_i^{(k)}) dt \geq \int_{t_i^{(k)}}^d e^{-\beta t} w(t, s_0, z_{i+1}^{(k)}) dt$ by Equation (11) and non-negativity of λ in active vintages. Since wages in inactive vintages are invariant in z , the inequality also holds if $d > s_0 + T$. Since the careers $z^{(k,i)}$ and $z^{(k,i+1)}$ are identical up to $t_i^{(k)}$, this implies that career $z^{(k,i)}$ weakly dominates $z^{(k,i+1)}$. By induction, it follows that $z^{(k,1)} = z^{(1)}$ weakly dominates $z^{(k,k)} = z^{(k)}$. Now, since $z^{(k)}$ converges to z and w is a continuous function, we have

$$\begin{aligned} \int_0^d e^{-\beta t} w(t, s_0, z(0)) dt &\geq \lim_{k \rightarrow \infty} \int_0^d e^{-\beta t} w(t, s_0, z^{(k)}(t)) dt \\ &= \int_0^d e^{-\beta t} w(t, s_0, z(t)) dt. \end{aligned}$$

We have shown that a demotion career is always weakly dominated by a full-promotion career. Thus when replacing all demotion careers in a life l by no-demotion careers, this alternative life weakly dominates l . \square

Lemma A.6 (Full-length no-demotion careers optimal). *Consider a stationary solution to the planner's problem (for any T). Suppose a worker with initial conditions (s_0, z_0) faces wages as implied by the planner's solution on \mathcal{S}_T and wages $w(t, s, z) = \mu(t)$ for all $(t, s, z) \notin \mathcal{S}_T$. Then any life l consisting of full-length no-demotion careers is optimal, and $V(s_0, z_0)$ as defined by Equation (12) is the worker's value function.*

Proof. Consider an arbitrary life l'' that is feasible given (s_0, z_0) . The goal is to show that any life l fulfilling full-length and no-demotion for all careers is weakly preferred to l'' . First, note that a life l' that replaces all demotion careers in l'' by no-demotion careers is weakly preferred to l'' by Lemma A.5.

It remains to show that l' is in turn weakly dominated by l . Construct the following sequence of lives $\{l'_n\}_{n=0}^\infty$. l'_n is identical to l' up to career $n - 1$, i.e. $l'_n(t) = l'(t)$ for all $t \in [0, t_n)$. However, the worker stays in vintage s_n until its demise at $t = s_n + T$ in life l'_n (whereas she may leave the vintage before or after in life l'). For the careers in l'_n after n , pick any concatenation of full-length and no-demotion careers.

I will now show that l'_n is weakly preferred to l'_{n+1} . First, if career n in life l' is full-length, then l'_n and l'_{n+1} yield the same value: they are identical up to t_{n+1} , and from t_n on both must yield value $\int_{t_n}^\infty e^{-\beta t} \mu(t) dt$ by Lemma A.4. Second, if career n is longer than full length, then l'_n and l'_{n+1} yield the same value since $w(t, s, z) = \mu(t)$ for inactive vintages. Third, if career n is shorter than full-length, then l'_n is weakly preferred to l'_{n+1} since Equation (12) implies $v(l'_n) - v(l'_{n+1}) = \int_{t_n}^{s_n+T} \lambda(t, s_n, z_n) dt \geq 0$ by non-negativity of the Lagrange multipliers λ , and invoking again Lemma A.4 for continuation values from career $(n + 1)$ onward.

Since l'_n is weakly preferred to l'_{n+1} for all n , by induction l'_0 is weakly preferred to l'_N for any N , and $v(l'_N) \rightarrow v(l')$ as $N \rightarrow \infty$. Note that l'_0 contains only full-length no-demotion careers, and it is weakly better than any other feasible life (since l' weakly dominates l'' , and l'' was arbitrary). Furthermore, note that we could have chosen any l'_0 for which the worker stays in her initial z -position until her initial vintage dies (i.e. $z_0(t) = z_0$ for all $t \in [0, s_0 + T]$) and consists of full-length no-demotion careers thereafter. Thus this proves the claim, invoking Equation (12) for statement about the value function. \square

Proof for Proposition 4.5 (planner's solution is equilibrium):

Proof. We have to show that both firms and workers behave optimally given the wage function w specified by the proposition. 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 active vintages $s \in [t - T^*, t]$ at time t firms behave optimally by setting labor demand to n^* because this sets marginal products equal to wages. In inactive vintages $s < t - T^*$ no production occurs since the highest attainable profits are negative. The cost-minimizing skill mix for these vintages is such that marginal products are constant across z since $w(t, s, z) = \mu(t)$ for all z . Choosing this skill mix, average productivity of a worker is $Y^*(t, s)$. However, note that we have $Y^*(t, t - \tau) < Y^*(t, t - T^*) = \mu(t)$ for vintages of age $\tau > T^*$, since by Proposition 4.4 we have $T^* \geq \bar{T} = \arg \max_{\tau \geq 0} Y^*(t, t - \tau)$ and by the fact that $Y^*(t, t - \tau)$ is decreasing in τ for $\tau > \bar{T}$. So average productivity under the cost-minimizing skill mix is lower than the average wage, thus production and optimal labor demand are zero for all vintages $s < t - T^*$.

Second, we study the worker's problem. Consider the problem of a worker who faces wages as given in Point 1 of the proposition. By Lemma A.6, V as defined in (11) is the value function for the worker, and it can be attained by a life that consists of full-length no-demotion careers. We still have to establish that the lives prescribed by the no-crossing measure indeed attain the value given by V – note that they may contain demotion and

less-than-full-length careers. By Equation (12), we have $\partial V/\partial z = -\lambda$. Thus V is invariant in z if and only if the experience constraint is slack. This means that the level lines of N and thus workers' careers can only be decreasing in t in neighborhoods where the value function is invariant in z . This implies that the demoted workers attain the same value that they would attain if they stayed on the same z -level, and thus they are behaving optimally.

It remains to show that the careers prescribed by the no-crossing measure are optimal even if they are not full-length. Suppose the no-crossing measure requires a worker to drop out of vintage s_n at $t_n < s_n + T$. Then by Lemma A.6, entry into any full-length no-demotion career is optimal, thus also staying in s_n at level $z = t_n - s$ is optimal, since this is an entry point. So the worker is not worse off by quitting the vintage than by staying in his vintage until its death, which concludes the proof. \square

A.3.9 Proof for Proposition 4.6 (wage compression)

Proof. To prove Point 1 of the proposition, first note that N and thus n are invariant in t inside any ball $B_\epsilon(t, s, z) \subset \mathcal{P}$ since the experience constraint binds. Hence, from the wage equation (2) we can calculate $d\hat{w}/dt = \rho[f'(t - s - z)/f(t, s, z) - f'(t - s - z - \Delta z)/f'(t - s - z - \Delta z)]$. We see that, as claimed, $d\hat{w}/dt < 0$ on B_ϵ since f'/f is a decreasing function by log-concavity of f . The second statement in Point 1 is an immediate consequence of Proposition 4.3.

In order to prove Point 2, take an open ball $B_\epsilon(t, s, z) \subset \mathcal{D}$. Note that inside $B_\epsilon(t, s, z)$, the worker's value function V is invariant in z for fixed (t, s) by (12): when changing z , the right-hand side does not change since $\lambda(t, s, z) = 0$ by definition of \mathcal{D} . Thus the partial derivative $\partial V/\partial z$ exists, and furthermore $\partial V/\partial z = 0$ throughout B_ϵ . Now varying t in (12) we see that also wages w must be invariant in z inside B_ϵ – otherwise (12) would be violated when varying t on B_ϵ since the right-hand side does not change in t . This implies that $\partial w/\partial z$ exists and that $\partial w/\partial z = 0$ throughout B_ϵ , which concludes the proof. \square

A.3.10 Proof for Proposition 4.9 (progress curve)

Proof. I first prove Point 1 in the proposition. Given (t, s) , $N_t(t, s, z) = 0$ for all z implies $n_t(t, s, z) = 0$ for all z . Now, taking the derivative of $Y(t, s)$ in t and invoking the wage equation (2) imply the claim.

To prove Point 2, consider without loss of generality vintage $s = 0$. Any worker who spends a career segment I contained in a time interval $[t, T]$, $t \in [0, T)$, in vintage 0 must obtain wages $\int_I e^{-\beta t} w(t, s(t), z(t)) dt \geq \int_I e^{-\beta t} \mu(t) dt$ as Equation (12) shows. Integrating over career segments of all workers in vintage 0 in a time interval $[t, T]$, we obtain $M(t) \equiv \int_t^T e^{-\beta \tilde{t}} [Y(\tilde{t}, 0) - \mu(\tilde{t})L(\tilde{t}, 0)] d\tilde{t} \geq 0$ for all $t \in [0, T]$. The derivatives of the function $M(\cdot)$ defined here are

$$\begin{aligned} M'(t) &= -e^{-\beta t} [Y(t, 0) - L(t, 0)\mu(t)], \\ M''(t) &= \beta M'(t) - e^{-\beta t} [Y_t(t, 0) - L_t(t, 0)\mu(t) - L(t, 0)\mu'(t)]. \end{aligned}$$

The second derivative exists and is continuous in a stationary equilibrium since Y and N inherit continuous differentiability in t from n (which itself is continuously differentiable in t since u is continuous), and since $\mu(t) = e^{\gamma t}\mu(0)$ in a stationary equilibrium. By Proposition 4.3 and the fact that wage payment exhausts the firm's revenue, we have $\mu(T) = Y(T, 0)/L(T, 0)$. Thus $M(T) = M'(T) = 0$. Now, applying the Fundamental Theorem of Calculus twice we find that $M(t) = \int_t^T \int_u^T M''(x) dx du$. Since $M(t) \geq 0$ for all $t \in [0, T]$, this implies $M''(T) \geq 0$.

Straightforward manipulations then yield

$$\frac{F'_0(T)}{F_0(T)} = \frac{Y_t(T, 0)}{Y(T, 0)} - \frac{N_t(T, 0)}{N(T, 0)} \leq \frac{\mu'(T)}{\mu(T)},$$

where we again invoke the facts that $\mu(T) = Y(T, 0)/L(T, 0)$ and that $M'(T) = 0$. Since in a stationary equilibrium $\mu(t) = e^{\gamma t}\mu(0)$, the claim in Point 2 follows. \square

B Numerical solution algorithm

B.1 Description of the algorithm

The idea behind the algorithm to solve for the planner's solution is similar to the one in Kredler (2013). 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 decision: 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 then be computed. This density gives us a wage function from Equation (2).

Workers strategies are then updated in the next step. Following Proposition 4.7, the optimal policy response of workers to the wage function 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 -(career-)function is calculated, being proportional to the derivative of the wage function at this point. As in Kredler (2013), also entry densities into vintages are adjusted such that vintages with high entry values receive more workers in the next iteration. The sensitivities of workers' promotion and entry decisions as a response to the wage structure are tuning parameters of the algorithm. When setting the tuning parameters low enough, the algorithm converges (setting them too low, however, increases computation time). Convergence is defined as entry values into all vintages being ϵ -close, and that demotions occur if and only if the wage function's slope in z is ϵ -close to zero.

T^* is then obtained by increasing T until average productivity in the dying vintage is equal to the flow value of entry into new careers. This criterion turns out to be a smooth function of T in practice, so that T^* can easily be found by a Newton-type algorithm. However, care has to be taken not to choose T too large, since then the algorithm may fail to converge; see the following subsection for a discussion for why this is the case.

B.2 Discussion: no continuous maximizer for $T > T^*$

The computational difficulties that were encountered when increasing T too much are most likely linked to the fact that the supremum in the maximization problem $J^*(T) = \sup_n v(T, n)$ in (6) cannot be reached by any continuous control n whenever $T > T^*$. To see this, suppose there exists a solution to 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 admissible 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 continuous, 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.²⁰ Hence $J^*(T) = J^*(T^*)$ for all $T > T^*$, but existence of a maximizer breaks down for $T > T^*$.

Because of this non-existence result it is useful to approach T^* from below in computations.

²⁰Notice that obeying the experience constraints in the neighborhood behind T^* is not an issue: if the distribution is thinned out at a high rate, then there are always enough experienced workers available.