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

Kredler, Matthias

New York University

28 July 2008

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

Matthias Kredler[†]

17 October 2009

Abstract

I introduce endogenous human-capital accumulation into an infinite-horizon version of Chari & Hopenhayn's (1991) vintage-human-capital model. Different skill levels inside a vintage are complementary in production. I establish equivalence between competitive equilibrium and a planner's problem, which ensures uniqueness of equilibrium. Returns to skill and tenure premia are highest in young vintages, where skill is scarcest and agents accumulate human capital fastest. As the vintage ages, the skill premium decreases and vanishes entirely upon vintage death. The results are in line with German linked employer-employee data: Young establishments pay higher tenure premia but lower mean wages than old establishments.

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

1   Introduction

This paper introduces endogenous human-capital accumulation into a vintage structure and shows that the resulting framework is still analytically and computationally tractable. As in Chari & Hopenhayn (1991), human capital is tied to a technology and is lost when the technology is phased out.

In each vintage, different levels of human capital are complementary inputs to a constant-returns-to-scale production function used by competitive firms. Unlike in Chari & Hopenhayn's (1991) two-period overlapping-generations model, however, human-capital accumulation is endogenous and the possibly infinite lives of individuals allow for rich patterns in tenure-wage profiles (shown in figure 1).

In equilibrium, the extant skill structure across and inside technologies determines vintage choice of young workers. When imposing an Inada condition on production, all rungs of the skill ladder are filled in all active vintages and there is entry of young workers into all active vintages. Vintages are phased out by an endogenous firm decision. I show equivalence between the market equilibrium and the planner's solution[1], which implies uniqueness of equilibrium.

The premium on technology-specific skills is shown to be highest in the youngest technologies. This result is driven by the relatively more pronounced scarcity of skill in young technologies. As technologies age, the skill premium shrinks and eventually vanishes entirely, a process I refer to as *wage compression*. The high skill premium induces young-vintage entrants to accumulate human capital faster than old-vintage entrants. Fast learning combined with high skill premia leads to faster earnings growth and higher tenure premia in young technologies. This is illustrated in figure 1, which plots earnings profiles along different careers.[2] The longest profiles pertain to workers entering the youngest vintages, the shortest to entrants into almost-dead vintages — note that in equilibrium all workers stay in their vintage until it is phased out.

As for entry wages, these are lowest in new technologies, as may be seen in figure 1 by comparing profiles at tenure $t = 0$. This phenomenon is due to a general-equilibrium effect: Since new technologies offer the most valuable skills, they are more attractive than older technologies ceteris paribus. If there is to be entry into all vintages, however, young vintages' entry wages must fall to the point where workers are indifferent between entering any technology.

---

[1]Due to the continuous-time setting (which is essential to obtain many of the analytic results in the paper), a novel kind of proof is required to do so. The proof is based on the partial differential equations resulting from the planner's problem and agents' optimality conditions from the Hamilton-Jacobi-Bellman equation. The reason conventional proof techniques do not apply is that production of new human capital h is represented by a cost functional on the time-derivative of h, which is not a production function that fits into the classical framework.

[2]Figure 1 is generated for the parameters of the preferred calibration; see section 4 for details.

Also average wages are lowest in young technologies. This occurs because they have the most imbalanced skill mix: Many low-skilled workers are working alongside very few high-skilled ones. Over the life cycle of the vintage, average wages then rise continuously. As more and more workers climb up in the hierarchy, the skill mix becomes more balanced. The vintage is shown to reach maximal productivity upon – and *only* upon – its death.[3]

A positive measure of workers is shown to experience real-wage losses at some point of their vintage career (as is evident from figure 1). This is true despite the fact that skill is never lost. The wage losses are driven by *obsolescence*, the fact that the relative price of skill falls as the vintage ages. Thus, unlike in classical human-capital models à la Ben-Porath (1967), it is not necessary to assume depreciation of human capital in order to obtain downward-bending wage profiles for old workers. When assuming a constant-elasticity-of-substitution (CES) production function, this *obsolescence effect* can be isolated from an *experience effect* and an *organizational-capital* (or vintage-productivity) *effect*.[4]

Once vintages are phased out, workers again experience wage losses as they are re-locating to newer vintages and lose their vintage-specific human capital. In figure 1, this manifests itself in entry wages (the starting points of the profiles) being below exit wages (the final points of the profiles).

Neither type of wage losses occurs in Chari & Hopenhayn (1991), mainly because workers only live for two periods there.[5]

In the special case where different skill levels are perfect substitutes within a vintage, I show that all workers follow the same career pattern: When born, they enter the current frontier vintage and embark upon a unique path of human-capital accumulation. Once the returns from learning about a technology have become small with respect to the gains from switching to a new (more productive) one, workers switch to the newest vintage and repeat the same cycle again. I show that an increase in frontier-productivity growth leads to faster switching between vintages. Through a horizon effect, this leads to lower skill levels for any given given vintage tenure.

When different skills are complementary, however, this horizon effect

---

[3]Since firms are competitive, this is equivalent to the marginal productivity of all skill levels – and thus wags – being equalized upon vintage death.

[4]See equation (19) for the decomposition.

[5]In Chari & Hopenhayn's (1991) setting, figure 1 would show a series of 2-point earnings profiles, each of them increasing. As in my model, profiles in young vintages would have the steepest slope and the lowest entry points. 2nd-period wages would be highest in young vintages and then decrease with vintage age.
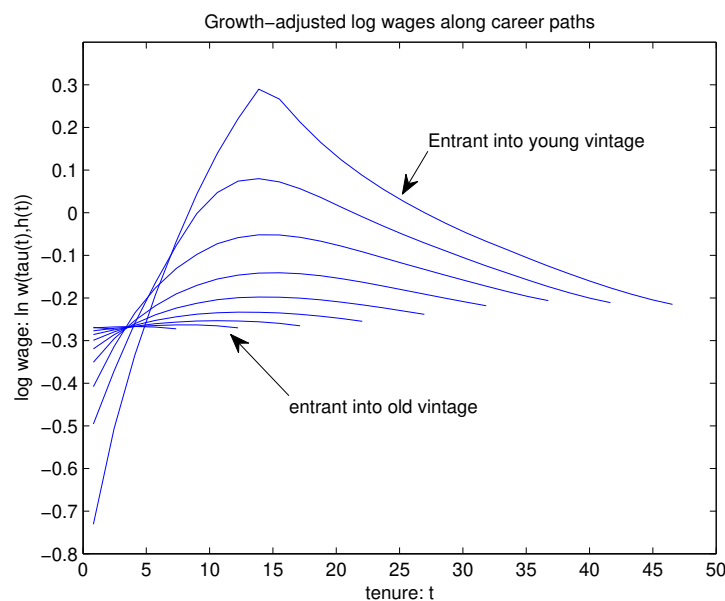
Figure 1: Tenure-earnings profiles over career

may be overcome by a skill-scarcity effect. Numerical exercises show that faster growth still induces shorter vintage lifetimes, but that workers may acquire skills faster when complementarity is sufficiently high. This effect arises because higher growth concentrates workers in young technologies where skill is scarce and skill premia are high. Thus, skill complementarity is necessary for a technological acceleration to lead to more learning-intensive careers.

In German employer-employee-matched data, I find that key predictions of the model are borne out when interpreting vintage age (in the model) as establishment age (in the data): Younger establishments have higher tenure premia, but pay lower wages than older establishments on average. In a calibration exercise, I find that the model yields a reasonable quantitative fit to the earnings structure (by establishment age and tenure) and the worker distribution (by establishment age). The model is also successful in predicting correlations of growth measures and the earnings structure: Fast-growing industries, occupations and establishments display higher tenure premia than slow-growing ones, but pay lower mean wages.[6]

---

[6]Only occupations are an exception to the latter statement: Fast-growing occupations

In relation to the previous literature, the model presented here is closest to Chari & Hopenhayn (1991), but differentiates itself by endogenous human-capital accumulation, workers' infinite life time and the resulting detailed predictions on tenure-wage profiles.

Somewhat less related is Parente (1994), who studies a ladder model where agents face a trade-off between experience accumulation and obsolescence. However, experience accumulation is of the learning-curve type and the extant skill structure in a technology does not influence agents' decisions since there are no complementarities between workers. The same is true for Violante (2002), where one worker and one entrepreneur are substitutable inputs in a vintage production function and workers accumulate skill according to a learning curve.

In Lucas (1988), the average level of economy-wide (general) human capital has an externality on output and so the extant skill structure *does* influence human-capital-accumulation decisions as in my model. However, human capital is not technology-specific and skill externalities are not priced, whereas in my model the positive effect of skilled workers on the productivity of the unskilled (in the same vintage) is reflected in wages.

Prescott & Boyd (1987) develop an overlapping-generations model of coalitions, where experienced and inexperienced workers face a trade-off between production of output and training of young workers. An important difference between their model and mine is that no reallocation of workers from phased-out technologies to new ones occurs in their model. In a more recent contribution, Garicano & Rossi-Hansberg (2008) model explicitly how tasks are shared within an organization and how organizations grow more complex over time. My framework has no explicit model of task sharing. However, skill complementarity implies that a balanced skill mix is desirable for an organization. Efficient task sharing and efficient skill-accumulation decisions are induced by the competitive wage structure.

An entirely different class of models that is able to generate tenure-related gains in earnings are search models. Burdett & Coles (2003) show in such a setting that firms optimally offer increasing wage schedules in order to prevent costly turnover. The predictions of the model presented here are different, however: In Burdett & Coles (2003) changes in *employer* are vital for the determination of wage profiles, whereas changes in the *technology* a worker uses are crucial in my setting.

The remainder of the paper is organized as follows: Section 2 presents the model and characterizes the competitive equilibrium. Section 3 shows

pay higher mean wages than slow-growing ones.

5

that this equilibrium is equivalent to the solution of a planner's problem. Section 4 presents computational results in a calibrated version of the model. Section 5 concludes and discusses potential further applications of the framework.

2 Model

2.1 Technology

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

The production technology s uses labor inputs that are specific to this technology. The inputs are arranged on a hierarchy and indexed by $0 \leq h \leq 1$. The inputs on this ladder can be thought of as tasks that are increasing in difficulty; tasks with a higher index require more vintage-specific human capital. Section 2.2 will specify exactly how this form of human capital is accumulated by workers.

The production function is supposed to capture the following notions: (i) Newer vintages are more productive holding input ratios equal and (ii) the production function is complementary in its inputs. Specifically, I impose

$$Y(t, s) = e^{\gamma s} \tilde{Y}\big(n(t, s, \cdot)\big)$$

where $n(t, s, h)$ is the density function of workers at time t in vintage s with human capital h and $\tilde{Y}$ is a functional on the space of C^1 functions on $[0, 1]$ with the following properties:

- Constant returns to scale (CRS): $\tilde{Y}(\lambda n) = \lambda \tilde{Y}(n)$.

- The Frechet derivative[7] $\tilde{w}(n)$ exists everywhere, is continuous in n and $\tilde{w}(n) > 0$ for all h, n.

- Weak concavity: $\tilde{Y}\big(\lambda n + (1 - \lambda)n'\big) \geq \lambda \tilde{Y}(n) + (1 - \lambda)\tilde{Y}(n')$ for all $0 \leq \lambda \leq 1$.

---

[7]Recall that the *Frechet derivative* is the generalization of the gradient vector from R^n to infinite-dimensional spaces. In this model, it is a wage function $\tilde{w} : [0, 1] \rightarrow R^+$ which takes h as its argument. In the case of the CES aggregator in (1), it is given by the familiar $f(h)\tilde{Y}^{1-\rho}n(t, s, h)^{\rho-1}$.

The first two properties imply that in a competitive setting, total wage payments exhaust production. An example for such a functional is the constant-elasticity-of-substitution (CES) aggregator

$$\tilde{Y}_{CES}[n(t,s,\cdot)] = \left[\int_0^1 f(h)n(t,s,h)^\rho dh, \right]^{1/\rho}, \tag{1}$$

where $\rho \leq 1$ and $f(h) \geq 0$. Total output in the economy at time t is given by $Y(t) = \int_{-\infty}^t Y(t,s)ds$.

At times, I will additionally invoke the following Inada condition:

Definition 2.1. *(Inada condition)* The production function is said to fulfill an Inada condition if $n(h) \to 0$ implies $\tilde{w}(h) \to \infty$ for all $h \in [0,1]$ and there is a unique element n^* on the interior of the unit simplex $\Delta = \{n : \int_h n = 1\}$ that maximizes output at $\bar{y} \equiv \max_{n \in \Delta} \tilde{Y}(n)$.

Optimality of n^* implies that marginal factor returns must be equalized and we have a constant wage schedule $\tilde{w}(n^*) = \bar{y}$. The CES aggregator in (1) above fulfills the Inada condition if and only if $\rho < 1$.

Competitive firms take wages for all labor inputs as given in each instant. Since the production technology is CRS, profits are zero for any t and s in equilibrium. Workers are paid their marginal product, so $w(t,s,\cdot) = e^{\gamma s}\tilde{w}\big(n(t,s\cdot)\big)$.

2.2 Workers

There is a continuum of agents of mass one. Agents die at a constant rate δ. New agents are born at the same rate δ, keeping total population constant.

Agents have linear utility and discount the future at rate β, where $\beta + \delta > \gamma$. Each agent chooses a work life $\{s(t), h(t)\}_{0 \leq t < \infty}$, which consists of a function $s(t)$ specifying the vintage the agent works for all t and a function $h(t)$ specifying the task he performs at t in vintage $s(t)$. It is required that the vintage already exist at time t, i.e. $s(t) \leq t$, and that $s(t)$ be a measurable function in t.[8]

As for human-capital accumulation $h(t)$, I require that a worker start her work life in position $h = 0$ when she enters the vintage; mathematically I impose that $h(\bar{t}) > 0$ only if there is an interval (a,b) around $\bar{t}$ such that

---

[8]This specification allows for lives with more than countably many vintage changes; a relevant example for such a life is $s(t) = t$.

$s(t) = \bar{t}$ for all $a < \bar{t} < b$.[9] There is no cost of switching between vintages. I will refer to a *career segment* (or short *career*) $l'(t)$ as the maximal open interval $(l'_0(t), l'_1(t))$ around an instant t that is entirely spent in one vintage.[10] If $l'_0(t) = t = l'_1(t)$, the career segment as an open interval is empty and we will not call this degenerate stay in a vintage a career segment. Since segments are open intervals and each of them contains a rational number, there can only be countably many of them in an agent's life.

To capture the notion that human-capital accumulation inside a vintage is costly, I require that the function h be differentiable on all segments and assume that the worker has to pay a flow cost $e^{\gamma s(t)} c(\dot{h}(t))$ on segments, where $\dot{h}$ denotes the time derivative of h and c is a cost functional with the following properties:[11]

- Costless demotion: $c(\dot{h}) = 0$ if $\dot{h} \leq 0$.

- Convexity: $c'(\dot{h})$ is a continuous, strictly increasing function on $(0, \infty)$.

- Inada condition: $\lim_{\dot{h} \to \infty} c'(\dot{h}) = \infty$

An example that satisfies these properties is $c(\dot{h}) = \bar{c} \max\{\dot{h}, 0\}^2 / 2$. No costs accrue for non-segments; observe that for any t that is not on a segment, we must have $h(t) = 0$. This cost may be interpreted as a psychic or monetary cost that the worker incurs when learning about the technology in his spare time or during unpaid overtime at work.

Each agent born at $t = 0$ enters the economy with some experience level h_0 for a vintage of age $s_0 \leq 0$, i.e. the first segment may start off with $h_0 \geq h(0) > 0$ if $s(0) = s_0$. There is a density $n_0(\tau, h)$ over these endowments at $t = 0$. New-born workers enter without any endowment, i.e. $h(t) = 0$ for a worker born at $t > 0$.

---

[9]This also means that a worker has to start at zero again even if he had worked in that vintage before but quit it at some point. This assumption is imposed for tractability; in equilibrium, workers would not want to return to vintages they have once left.

[10]Formally, define the end points as $l'_0(t) \equiv \inf\{a \leq t : s(u) = s(t) \; \forall u \in [a, t]\}$ and $l'_1(t) = \sup\{b \geq t : s(u) = s(t) \; \forall u \in [t, b]\}$.

[11]The cost of human-capital accumulation is growing at the pace of total factor productivity (TFP) to ensure stationarity of the economy. This specification entails that the costs of human-capital accumulation relative to productivity in a technology do not change. This is in line with models where workers have to set aside time from productive work in order to accumulate human capital; in such a setting, the opportunity cost of human-capital accumulation is given by the marginal productivity of devoting one's time to productive work instead of learning. The specification here is a modeling shortcut that avoids the explicit modeling of hours.

To summarize, the agent's criterion for a given life l_t starting at t is

$$v(l_t) = \int_t^\infty e^{-(\beta+\delta)(u-t)} \left[w\big(u, s_t(u), h_t(u)\big) - e^{\gamma s_t(u)} c\big(\dot{h}_t(u)\big) \right] du,$$

where it is understood that $\dot{h} = 0$. The value function is defined as $V(t, s, h) = \sup_{l_t(t)=(s,h)} v(l_t)$, where the supremum is taken over all feasible lives starting with endowment (s, h). Since discounting is exponential, optimal policies are time-consistent and $V(t, s, h)$ also gives us the forward-looking value for any agent born before t who finds herself in position (s, h) at t.

2.3 Stationary equilibrium

I will limit the discussion to densities n which have a collection of sets S_n^i in $X \equiv [0, \infty) \times [0, \infty) \times [0, 1]$ as their support $S_n = \cup_i S_n^i$. I require the sets S_n^i to contain an open ball; n is assumed continuous and differentiable on each set S_n^i.[12]

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

$$n(t, s, h) = n(s + \tau, s, h) = \bar{n}(\tau, h).$$

Stationarity immediately implies that wages and production grow at rate γ, i.e. $w(t, s, h) = e^{\gamma t} \bar{w}(\tau, h)$, $Y(t, s) = e^{\gamma t} \bar{Y}(\tau)$ and $Y(t) = e^{\gamma t} \bar{Y}$. By stationarity of the cost functional, also the value function grows at rate γ: $V(t, s, h) = e^{\gamma t} \bar{V}(\tau, h)$. From now on, we will only work with the stationary distribution; I thus drop the bar-notation and write simplify $n(\tau, h)$, $w(\tau, h)$ and so forth.

Definition 2.2. A *stationary competitive equilibrium* is a stationary density $n(\tau, h)$, a measure μ on all possible work lives $l(t) = \{\tau(t), h(t)\}$ and a wage function $w(\tau, h)$ that is continuous on the interior of X such that:

- Compatibility of μ and n: For all Borel sets B in $\mathcal{R}^2$ and for all $u \geq 0$,[13]

$$\int_{t \leq u} e^{-\delta(u-t)} I\big\{ \big(\tau_t(u), h_t(u)\big) \in B \big\} d\mu(l) = \int_B n(\tau, h) d\tau dh.$$

---

[12] The Inada condition 2.1 will naturally lead to such non-degenerate sets S_n^i for the support. Only for the case of a linear production function (i.e. setting $\rho = 1$ in $\tilde{Y}_{CES}$) it will make sense to consider a more general class of sets for S_n, see section 2.5. Note that the specification here allows for densities that drop precipitously down to zero when a vintage dies — which is exactly what occurs in equilibrium. Also, note that feasibility requires that the neighborhoods be connected to points with $h = 0$ or $t = 0$.

[13] $I\{\cdot\}$ denotes the indicator function. The subscript l_t again refers to an agent born at $t \geq .0$. The simple multiplication of the indicator function by the survival function $e^{-\delta(u-t)}$ is valid since death is independent of workers' strategies.

- Optimal labor demand: $n(\tau, \cdot) = \arg\max_{\tilde{n}} \left\{ Y(\tilde{n}) - \int w(\tau, h)\tilde{n}(h)dh \right\} \ \forall \tau$

- Optimal labor supply: Any set A over lives such that $l_t \in A$ implies $v(l_t) < e^{\gamma t}V(\tau_t(t), h_t(t))$ has measure zero under μ.

Note that the definition requires wages to be specified also for regions outside the support S_n of n, i.e. equilibrium must specify a wage schedule in such regions which makes optimal labor supply and optimal labor demand equal zero.

2.4   Properties of equilibrium

We will be looking for a value function $V \in C^1(X)$ that is consistent with a stationary equilibrium. I start to characterize the equilibrium by deriving some properties of the value function. Since workers can always drop down arbitrarily fast in the hierarchy at zero cost and the value function is continuous, we have:

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

Also, workers always have the option to start a new career immediately. So in any position, they must always at least as well off as workers who start an optimal career.

Definition 2.3. Define the maximal value that can be attained by a career starter as $W = \max_\tau V(\tau, 0)$.

Lemma 2.2. (Value equal for all career starters) $V(\tau, 0) = W$ *for all τ and* $V(\tau, h) \geq W$ *for all (τ, h).*

We will now turn to characterizing the support of the equilibrium density S_n. First, observe that the Inada condition ensures that all rungs in the skill hierarchy must be filled if a vintage is in production:

Lemma 2.3. (All jobs filled in producing vintage) *If the Inada condition 2.1 holds, then $Y(\tau) > 0$ implies $(\tau, h) \in \bar{S}_n$.*[14]

This is a consequence of promotion costs being bounded for any position with $\tau > 0$ but wages going to infinity for empty slots in the skill ladder. A formal proof is given in appendix A.1.2.

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

---

[14] $\bar{A}$ denotes the closure of a set A.

Lemma 2.4. (Finite support of technologies) *There exists $T < \infty$ such that $\int_0^1 n(\tau, h)dh = 0$ for all $\tau > T$.*

The proof uses the argument that workers can always secure some positive wage in a frontier vintage without going through training, but that old vintages' productivity goes to zero relative to the frontier. The result is ultimately driven by the the fact that returns to learning are bounded but TFP growth is not.

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

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

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

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

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

$$\int_{\text{all } l'} \tilde{v}(l') \leq \int_0^\infty e^{(\gamma - \beta - \delta)t} \int_{-\infty}^{t-S} \int_0^1 n(t, s, h)w(t, s, h)dhdsdt \leq$$

$$\leq \bar{y}e^{-\gamma S} \int_0^\infty e^{(\gamma - \beta - \delta)t} \left(\int_{s,h} n(t, s, h) \right) dt,$$

where in the last step I used that the upper bound on production for vintages even older than S is at most $e^{-\gamma S}\bar{y}$ for some $\bar{y} < \infty$, see lemma A.1) for a proof.

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

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

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

Definition 2.4. Define the last vintage in production by $T^* \equiv \inf_\tau\{\tau : \int_0^1 n(\tau, h)dh = 0\}$. Note that $T^* < \infty$ is ensured by lemma 2.4.

In order to further characterize S_n, it will be useful to know more about the wage structure in the oldest technology. Consider the problem of a worker who optimizes his career with respect to the switching point $\bar{t}$ when he quits a vintage:

$$\max_{\bar{t}} \int_0^{\bar{t}} e^{-(\beta+\delta-\gamma)t} w[\tau(t), h(t)]dt + e^{-(\beta+\delta-\gamma)\bar{t}}W$$

Since w is continuous, differentiating with respect to τ yields that $\bar{t}$ can only be optimal if $w[\tau(t), h(t)] = (\beta + \delta - \gamma)W$, where the right-hand side is the flow value of starting a new career. If the wage was still higher than that, the worker should stay in the vintage at least a bit more; if it was lower, quitting a bit earlier would make him better off. We summarize:

Lemma 2.5. (Final career wage) *At the end of any career segment l_0 wages tend to the flow value of starting a new career, i.e. $\lim_{t\to l_1} w[\tau(t), h(t)] = (\beta + \delta - \gamma)W$.*

Corollary 2.6. (Flat wage structure in oldest technology) *For all $(T^*, h) \in S_n$, we have $w(T^*, h) = (\beta + \delta - \gamma)W$. If the Inada condition 2.1 holds, this implies that vintages attain maximal productivity upon their death.*

For vintages $\tau > T^*$ that are out of production, the equilibrium definition 2.3 requires us to specify a wage structure that makes it undesirable for both workers and firms to use those vintages. There are many possible choices for w in this region; one of them is $w(\tau, h) = e^{-\gamma(t-T^*)/2}w(T^*, 0)$. Workers will strictly prefer W to any career behind T^*, and firms would not break even for $\tau > T^*$ — even at optimal factor-input ratios, TFP decays faster with τ than the wage bill does. Also, continuity of w is ensured as required.

Reasoning along these lines shows that there cannot be any holes in the support of n along the τ-direction:

Lemma 2.7. (No holes in vintage space) *Suppose the Inada condition 2.1 holds. Then, if both $Y(\tau_0) > 0$ and $Y(\tau_1) > 0$, also $Y(\tau) > 0$ for all $\tau_0 < \tau < \tau_1$.*

This of course implies that all these in-between vintages τ have points (τ, h) in the support S_n.

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

Lemma 2.7 together with lemma 2.3 implies that the closure of S_n must be a rectangle $[T_0, T^*] \times [0, 1]$ if the Inada condition 2.1 holds. Section 2.5 establishes that there cannot be holes in the τ-direction either when labor inputs are perfect substitutes. Arguments in sections 2.5 and 2.12 will finally show that we must of course have $T_0 = 0$.

We will now seek a further characterization of the equilibrium studying the worker's behavior on career segments. The Hamilton-Jacobi-Bellman equation (HJB) for an interior point of a career segment is the following first-order partial differential equation (PDE):

$$-V_\tau(\tau, h) = w(\tau, h) - (\beta + \delta - \gamma)V(\tau, h) + \max_{\dot{h}} \left\{ -c(\dot{h}) + \dot{h} V_h(\tau, h) \right\} \quad (2)$$

where partial derivatives are denoted by subscripts. The equation says the following: If we know the value function for a given τ on the entire h-ladder, we can get the value a bit left of this τ by letting the agent choose the optimal career slope $\dot{h}$. The optimal slope depends on the marginal value of skill V_h and the cost of learning. The change in the value function a small step to the left (keeping h fixed) is the gain the agent obtains from moving up in the hierarchy (the term inside the max-operator) and another term, which is the difference between the current wages and the flow value of $V(\tau, h)$ under the discount factor $\beta + \delta - \gamma$ (which is modified for economic growth).

The first-order condition (FOC) for $\dot{h}$ corresponding to the HJB (2) is

$$c'\big(\dot{h}(\tau, h)\big) = V_h(\tau, h), \quad (3)$$

where a unique solution for $\dot{h}$ is assured whenever $V_h > 0$ by the assumptions on c. Since c is convex, the FOC implies that greater value differentials in the hierarchy induce faster human-capital accumulation. Given the boundary condition $V(T^*, h) = W$ for all h, equations (2) and (3) together determine the optimal policies of an agent who takes the wage function w as given.

It is worthwhile noting that the term in V_h (i.e. the max-operator) in the HJB (2), is the Legendre transform of $c(\cdot)$ and hence a convex function in V_h. This implies that the HJB will be non-linear in V_h.[15]

Sometimes, it will be convenient to work with the Euler equation, which tells us how the marginal value of human capital V_h changes along an optimal career path. Differentiate (2) with respect to h and use the envelope condition (3) to obtain

$$\frac{dV_h}{dt} = \dot{h}(V_h)_h + (V_h)_\tau = (\beta + \delta - \gamma)V_h - w_h, \qquad (4)$$

where an agent's career is parameterized by time by t: $d\tau = dt$ and $dh = \dot{h}dt$. The dependence of the various functions on (τ, h) is suppressed for the sake of clarity. We can solve (4) as an ordinary differential equation in t along an agent's optimal career path and see that the marginal value of human capital equals the discounted integral of marginal wage gains over a career:

$$V_h(t) = \int_t^T e^{-(\beta + \delta - \gamma)(u-t)} w_h\big(\tau(u), h(u)\big) du, \qquad (5)$$

where T is the end of the career segment and $V_h(T, h) = 0$ since $V(T, h) = W$ for all h, i.e. the marginal value of skill is zero at the end of a career. This suggests that the incentives for human-capital accumulation are strongest in the beginning of a career, making human-capital accumulation front-loaded and decreasing over segments.

I now proceed to characterize how the density $n(\tau, h)$ evolves given the optimal local behavior of agents characterized by (3). Inside S_n, n must obey the following PDE:

$$n_\tau(\tau, h) + \dot{h}(\tau, h)n_h(\tau, h) = -\left[\delta + \dot{h}_h(\tau, h)\right] n(\tau, h), \qquad (6)$$

where the notation $\dot{h}_h = \frac{\partial \dot{h}}{\partial h}$ is used. This PDE says the following: When following an agent's optimal career path, the density thins out at the death rate plus the divergence of the promotion policies $\dot{h}_h$ inside the hierarchy. Section 3.1 will deliver a derivation of this equation.[16] For a given boundary condition $n(\tau, 0) = n_0(\tau)$ on $\tau \in [T_0, T^*]$, we may solve this PDE throughout S_n to obtain n given $\dot{h}$.

---

[15]For example, the Hamiltonian equals $V_h^2/2\bar{c}$ in the quadratic case.

[16]The equation is the usual mass-transport equation for densities in a deterministic context; it may be seen as a special (non-stochastic) case of the Kolmogorov forward equation.

To summarize, the HJB (2) with its boundary conditions characterizes workers' optimal strategies given wages. Equation (6) tells us how the resulting decisions by workers in (3) translate into a density n (once we know how new-borns enter vintages).[17] Optimality of firms' decisions implies that wages w on the support of n are given by the Frechet derivative of $Y(\tau)$ with respect to $n(\tau, \cdot)$. Wages w then feed back again into workers' HJB.

An equilibrium is a worker distribution n satisfying (6), where policies $\dot{h}$ are determined through workers optimal behavior characterized by (2). Equation (2), in turn, uses the marginal products of workers induced by n as the wage function w.

2.5 Sub-case: Perfect substitutes

The next two subsections will be concerned with the characterization of the wage structure and human-capital accumulation. Specifically, the following objects are of interest: the *skill premium* inside a given vintage (i.e. $w(\tau, \bar{h}) - w(\tau, \underline{h})$ for $\bar{h} > \underline{h}$), the *(vintage-)tenure premium* (defined exactly like the skill premium, but conditioning on vintage tenure instead of h as the independent variable) and the intensity of *skill accumulation* $\dot{h}(\tau, h)$. I will also study how these objects depend on technological growth by comparing steady states for different values of γ, ceteris paribus.

We first turn our attention to the special case where different skill levels are perfect substitutes: Take the CES-aggregator (1) with $\rho = 1$ as the production function. Then, wages are independent of the distribution of workers across the skill hierarchy: $w(\tau, h) = e^{-\gamma\tau} f(h)$[18]. Throughout this section, we will assume a standard learning curve and require that $f'(h) > 0$ and $f''(h) < 0$.

With a linear production function, the problem essentially reduces to one of partial equilibrium: It is sufficient to solve a worker's problem, let every worker follow her optimal policy and collect the results in an equilibrium density n.[19]

A first result is that a worker will always choose to switch to the newest vintage when she relocates:

---

[17]The entry density $n_0(\tau)$ may be freely chosen if $V_\tau(\tau, 0) = 0$ for all $\tau \in [T_0, T^*]$, i.e. if workers are indifferent between all careers.

[18]Recall that wages are adjusted by frontier-productivity growth.

[19]Note that typically all agents will follow the same path and thus n would be a *measure* that cannot be represented by a density function. However, this is unproblematic since the production function (1) and the PDE (6) still make sense, the former as a linear functional on measures and the latter in the weak sense.

Lemma 2.8. (Always enter newest vintage under substitutability) *If $\tilde{Y} = \tilde{Y}_{CES}$ with $\rho = 1$, then for any optimal life $h(t) = 0 \Rightarrow \tau(t) = 0$ at the beginning of segments and $\tau(t) = 0$ almost everywhere on non-segments.*

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

We will now be concerned with the question how technological progress (in the form of a change in γ) affects agents' decisions. Using the solution to the Euler equation (5), we obtain

$$V_k(t) = \int_t^{T^*} e^{-(\beta+\delta)(u-t)} f'\big(h(u)\big) du, \tag{7}$$

where T^* is the optimal switching point to a new career. Since the wage gains from human-capital accumulation are decreasing in h by the concavity assumption on f, this entails that workers in a lower hierarchy position have stronger incentives to learn ceteris paribus. Another consequence is that the h-profile will always steeper for a worker with a longer horizon. The following lemma (proven in appendix A.1.3) will be sufficient to prove the formalization of these insights:

Lemma 2.9. (Paths cross at most once) *If $\tilde{Y} = \tilde{Y}_{CES}$ with $\rho = 1$, then for two optimal careers $h(t)$ and $g(t)$ we have:*

$$h(t) \geq g(t) \text{ and } \dot{h}(t) < \dot{g}(t) \Rightarrow h(s) > g(s) \text{ for all } s < t.$$

Since equation (7) shows that technological growth is inessential for the intertemporal incentives of human-capital accumulation while *inside* a vintage – it only affects the optimal switching point T^* –, we can use the above lemma to compare optimal careers for different values of γ. As the *(vintage-)tenure premium*, we define the ratio of the wage of an tenure-t worker in a vintage to the wage of a career starter. In a stationary context, this equals $p_\gamma(t) = e^{-\gamma t} w(h(t))/w(h(0))$, where the optimal path h of course depends on γ through $T^*(\gamma)$.

Proposition 2.10. (Shorter horizon lowers tenure premia) *Assume $\tilde{Y} = \tilde{Y}_{CES}$ with $\rho = 1$. Suppose that $\gamma' \neq \gamma$, but fix all other parameters. Then $T^{*'} > T^*$ implies $p_{\gamma'}(t) > p_\gamma(t)$ for all $0 < t \leq T^*$.*

16

Proof. Without loss of generality, take two career segments h and h' in vintage $s = 0$ starting with $h'(0) = h(0) = 0$ and $T^{*'} > T^*$. Now suppose that $h'(T^*) \leq h(T^*)$. First, note that $\dot{h}(T^*) = 0$ but $\dot{h}'(T^*) > 0$ by equation (7) and the fact that $c'(\dot{h}) = V_h$. By lemma 2.9, the two paths cannot cross again for any $0 \leq t > T^*$. But this is a contradiction to $h(0) = h'(0) = 0$. By the same argument, the two paths cannot intersect at any other point $0 < t < T^*(\gamma)$. So we must have $h'(t) \geq h(t)$ and so $w(h'(t)) \geq w(h(t))$, which implies the desired result. $\quad\square$

This result shows that it is impossible that the life span of technologies shortens and simultaneously we see an increase in the experience premium. Figure 2 illustrates the intuition for the result: Workers with a shorter planning horizon in their technology have fewer incentives to invest in technology-specific knowledge and thus increase their productivity at a slower pace, leading to a lower experience premium.
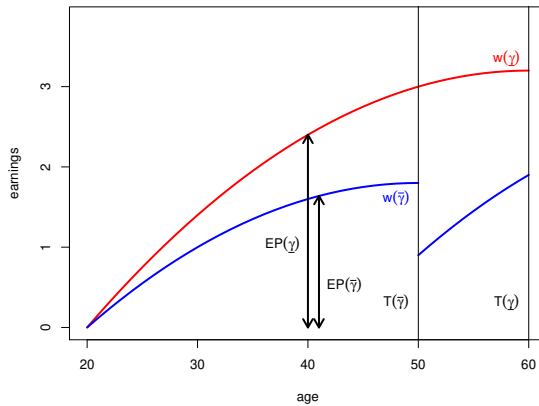

Figure 2: Horizon effect under substitutability

To show that faster technological growth leads to faster scrapping of technologies, define the discounted value of a career segment of length T by

$$K(T) = \int_0^T e^{-\rho t} \big[w\big(h_T(t)\big) - c\big(\dot{h}_T(t)\big) \big] dt, \tag{8}$$

where $h_T(\cdot)$ is the optimal skill-accumulation path for a career of length T.

Proposition 2.11. (Faster growth shortens careers and lowers tenure premia) *Assume* $\tilde{Y} = \tilde{Y}_{CES}$ *with* $\rho = 1$. *Then, if the function* $K(\cdot)$ *in (8)*

is twice differentiable, the following hold: $T^*(\gamma)$ is non-increasing in γ and strictly decreasing whenever $T^*(\gamma) > 0$; the tenure premium $p_\gamma(t)$ is decreasing in γ for any $0 < t \leq T^*$.

One may find this result somewhat counter-intuitive: In a world where different skill levels are perfect substitutes, it tells us we should expect that faster technical growth to lead to a less learning-intensive careers; this is in the sense that for each given tenure, agents would accumulated less knowledge in the high-growth than in the low-growth world.[20] The following sections will show that this result need not hold when different human-capital levels are complementary and give some intuition on how it can indeed be overturned.

2.6 Wage structure under complementarity

This section will further characterize the wage structure for the case where labor inputs of different skill levels are complementary. I start with the following observation:

Lemma 2.12. (Vintage T^* has highest entry wage) *If the Inada condition 2.1 holds, then $w(T^*, 0) \geq w(\tau, 0)$ for all $0 \leq \tau < T^*$.*

Proof. By lemma 2.6, $w(T^*)/(\beta + \delta - \gamma) = W$, i.e. always working in the oldest vintage as an unskilled worker is an optimal strategy. Suppose $w(\tau, 0) > w(T^*, 0)$ for some τ. Then always working in position $(\tau, 0)$ would give value $(\beta + \delta - \gamma)w(\tau, 0) > W$, which contradicts W being the maximal attainable value for a career starter. $\qquad\square$

Intuitively, entry wages have to be lower in young technologies for the following reason: Entering a new technology provides experience that will be valuable in the future. So, barring any offsetting wage differential, all workers would choose to enter new technologies. However, under the Inada condition some workers are also needed in low-skill tasks in the oldest technologies. In order for both entry options to be equally attractive, entry wages in young technologies have to be lower than in old technologies.

It turns out that on the top of the skill hierarchy, the converse is true:

Lemma 2.13. (Wage explosion for skilled in young technologies) *If the Inada condition 2.1 holds, then $\lim_{\tau \to 0} w(\tau, 1) = \infty$ and $Y(\tau) > 0$ for all $\tau \in (0, T^*)$.*

---

[20]Of course, agents in a high-growth world might still learn more in total since they switch to new vintages more often.

A proof is given in the appendix A.1. The intuition behind the result is that people with high skills in very young technologies must have worked very hard to acquire these skills. Thus, those workers have to be compensated by very high wages. Note that this is ensured if only very few people take such steep paths. When scarce enough a factor, the skilled in young technologies can earn unbounded returns under the Inada condition.

Collecting the previous results yields:

Corollary 2.14. (Wage Compression) *Suppose the Inada condition 2.1 holds. Then the wage difference between high-human-capital and low-human-capital workers is highest in the youngest vintages and lowest in the oldest vintages, i.e. $w(\tau, 1) - w(\tau, 0) \to \infty$ as $\tau \to 0$ and $w(\tau, 1) - w(\tau, 0) \to 0$ as $\tau \to T^*$.*

Proof. The first statement follows from $w(\tau, 0) \leq w(T^*, 0)$ for all $\tau < T^*$ (see lemma 2.12) and $w(\epsilon, 1) \to \infty$ (see lemma 2.13). The second statement follows from lemma 2.14. $\square$

Intuitively, the wage structure is compressed because experience becomes less scarce over the life cycle of a vintage. It is easier to acquire skills over a long time than to master a technology that was barely invented. With a view to tenure premia, note that wage compression opens the possibility that a technological acceleration can occur alongside an increase in tenure premia. Since the wage structure is steeper in young technologies, a shortening of the vintage horizon T^* can send more workers into steep earnings paths, increasing the average tenure premium.

Another consequence of the discussion above is the following:

Corollary 2.15. (Obsolescence/wage losses) *There is a positive measure of careers with $dw(\tau(t), h(t))/dt < 0$ for some t. Furthermore, agents who quit their vintage start their new career with a wage weakly lower than their last wage in the old career.*

Proof. The first statement follows from the reasoning laid out in lemma 2.13: There is a positive measure of agents with high human capital $h \in [1 - \epsilon, 1]$ in young vintages $\tau \in (0, \epsilon]$ with a high wage $w(\tau, h) > M$, M large, which must experience wage losses once they leave the high-wage region. The second statement is an obvious consequence of lemma 2.12 and 2.6. $\square$

Note that the first type of wage losses (those occurring during a career) cannot occur when skills are perfect substitutes. These wage losses during a

vintage career are remarkable since they occur without human-capital depreciation — an assumption often invoked in Ben-Porath-type models in order to obtain downward-bending wage profiles for old workers. Here, agents do not lose any of their skill over their vintage career; the reason for the wage losses is that the relative price for skill falls over time, a phenomenon referred to as *obsolescence*.

Finally, note that the second type of wage loss, which stems from the loss of vintage human capital due to a vintage change, is not due to an exogenous shock (an assumption sometimes made in human-capital models), but stems from an endogenous decision. The worker accepts a temporary wage loss in order to obtain skills in a new technology which pay off later in his work life.

3 The planner's problem

This section shows that the competitive equilibrium characterized in the previous section is equivalent to the solution of the following planner's problem: Let the planner weigh the utility of an agent born at t with $e^{-\beta t}$. Since it costs the planner $e^{-\delta(u-t)}$ units of time-u output to supply one unit to each surviving member of a cohort born at t and since utility is linear for all agents, it is easy to see that the planner's criterion is then to choose a function $n(t, s, h)$ (which we require again to be C^1 on a given support S_n) to maximize

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

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

3.1 Optimal promotion strategy

It turns out that the optimal promotion strategy is such that agents' career paths inside a vintage never cross. A formal proof for this statement, which builds on a discrete approximation technique, is given in appendix A.2. Intuitively, if a positive measure of agents crossed each other's way, then one could improve upon the strategy by maintaining the ordering inside the vintage, making agents go shorter paths and hence lowering total cost for the planner.

In the following, it will prove useful to work with the anti-cdf $N(t, s, h) \equiv \int_h^1 n(t, s, \tilde{h}) d\tilde{h}$. In a scheme where agents' paths do not cross, this function

must decrease at the death rate δ when we evaluate it along an agent's path staying in a fixed vintage s. A first-order approximation following a career line $\{h(t), \tau(t)\}$ yields:

$$N_t(t, s, h) + \dot{h}(t, s, h)N_h(t, s, h) = -\delta N(t, s, h), \tag{9}$$

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

Re-arranging equation (9) gives us an expression for the career slope $\dot{h}$ that the planner should choose given that she wants to implement a given n:

$$\dot{h}(t, s, h) = \frac{N_t(t, s, h) + \delta N(t, s, h)}{n(t, s, h)}. \tag{10}$$

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

3.2 The planner's first-order conditions

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

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

Lagrangian is then[21]

$$\mathcal{L} = \int_0^\infty e^{-\beta t} \left[\int_{t-T}^t Y(t,s) - e^{\gamma s} \left(\int_0^1 c\big[\dot{h}(t,s,h)\big] n(t,s,h) dh \right) ds \right] dt +$$

$$+ \int_{t,s,h} e^{-(\beta-\gamma)t} \left[\nu(t,s,h)\left(\dot{h} - \frac{\dot{N} + \delta N}{n} \right) + \right.$$

$$+ \lambda(t,s,h)\left(n_0(s,h) + \int_0^t u(\tilde{t},s,h)d\tilde{t} - n(t,s,h) \right) +$$

$$+ \eta(t,s,h)\left(\dot{N}(t,s,h) - \int_h^1 u(t,s,\tilde{h})d\tilde{h} \right) +$$

$$+ \xi(t,s,h)\left(N(t,s,h) - \int_h^1 n(t,s,\tilde{h})d\tilde{h} \right) +$$

$$\left. + \mu(t)\left(1 - \int_{t-T}^t \int_0^1 n(t,s,h)dhds \right) dt \right],$$

where the Lagrange multipliers are scaled by $e^{-(\beta-\gamma)t}$ to render them stationary. The set of constraints linked to the multipliers ν is taken from equation (10). The constraints connected to μ enforce that total population not exceed the bound 1. The rest of the constraints link the various variables related to the density n.

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

$$\eta(t,s,h) = e^{-\gamma\tau}c'\big(\dot{h}(t,s,h)\big) \tag{11}$$
$$\nu(t,s,h) = e^{-\gamma\tau}c'\big(\dot{h}(t,s,h)\big)n(t,s,h)$$
$$\xi(t,s,h) = \delta e^{-\gamma\tau}c'\big(\dot{h}(t,s,h)\big).$$

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

$$\lambda(t,s,h) = w(t,s,h) - e^{-\gamma\tau}c\big(\dot{h}(t,s,h)\big) + e^{-\gamma\tau}\dot{h}(t,s,h)c'\big(\dot{h}(t,s,h)\big) -$$
$$- \mu(t) - \delta \int_0^h \eta(t,s,\tilde{h})d\tilde{h}. \tag{12}$$

where we recognize in the terms involving $c(\cdot)$ the Hamiltonian from the value function (2) in the worker's problem. The last remaining derivative is

---

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

the one with respect to to $u(t, s, x)$, which will prove crucial to obtain the PDE that is equivalent to the HJB (2):

$$\int_\tau^T e^{-(\beta-\gamma)(\tilde{\tau}-\tau)} \lambda(\tilde{\tau}, h) d\tilde{\tau} = \int_0^h \eta(\tau, \tilde{h}) d\tilde{h}. \tag{13}$$

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

When substituting the expressions for the Lagrange multipliers (11) and (12) into (13) and imposing stationarity, one obtains

$$\int_\tau^T e^{-(\beta-\gamma)(\tilde{\tau}-\tau)} \Big[w(\tilde{\tau}, h) - e^{-\gamma\tilde{\tau}} c\big(\dot{h}(\tilde{\tau}, h)\big) + \dot{h}(\tilde{\tau}, h) c'\big(\dot{h}(\tilde{\tau}, h)\big) - \mu - $$
$$- \delta \int_0^h \eta(\tilde{\tau}, \tilde{h}) d\tilde{h} \Big] d\tilde{\tau} = \int_0^h e^{-\gamma\tau} c'\big(\dot{h}(\tau, \tilde{h})\big) d\tilde{h} \equiv \Lambda(\tau, h). \tag{14}$$

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

Directly from (14), we can get the following insights: First, when $\tau \to T$, the left-hand side and with it the marginal cost of human-capital accumulation $c'(\dot{h})$, and hence $\dot{h}$ itself, go to zero. This says that one should not accumulate human capital anymore just before the vintage shuts down, which also implies that $w(T, h)$ must be weakly increasing in h by non-negativity of the multipliers η. Second, when we let $h \to 0$, the right-hand side of (14) goes to zero and we see that $\lambda(\tau, 0) = 0$ for all τ. This says that for all entry jobs the value function must be equalized. Third, when we let both $\tau \to T$ and $h \to 0$ and use the insights from above, we obtain $w(T, 0) = \mu$. This says that $w(T, 0)$ is the reference wage of the economy: It does not provide any valuable experience, so it has to be just as attractive per se as any other career (in flow terms).

Now, take the derivatives of Λ in (14) in both dimensions to see how this excess-value function behaves on the interior:

$$\Lambda_h(\tau, h) = e^{-\gamma\tau} c'\big(\dot{h}(\tau, h)\big) \tag{15}$$

$$-\Lambda_\tau(\tau, h) = w(\tau, h) - e^{-\gamma\tau} c(\dot{h}) + e^{-\gamma\tau} \dot{h} c'(\dot{h}) - \mu - (\beta + \delta - \gamma) \Lambda(\tau, h). \tag{16}$$

23

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

$$-V_\tau(\tau, h) = w(\tau, h) - e^{-\gamma\tau}c(\dot{h}) + \dot{h}V_h - (\beta + \delta - \gamma)V(\tau, h), \qquad (17)$$

where we use $V_h = \Lambda_h = c'(\dot{h})$. When imposing the boundary conditions $V(\tau, 0) = V(T, h) = 0$ for all τ and for all h, this system is the same as the agent's HJB (2) and its boundary conditions in the decentralized problem. Subsection 3.5 will discuss equivalence of the planner's problem to the competitive equilibrium more carefully; before, it is useful to analyze the effects of variations in T.

3.3   Uniqueness

A fundamental concavity argument allows us to establish uniqueness of the planner's solution:

Proposition 3.1. (Solution to planner's problem is unique) *If $\tilde{Y}\big(n_{t,s}(\cdot)\big)$ is strictly concave in $n_{t,s}(h)$[22], then $J(n)$ is strictly concave in n and there is a unique density $n(t, s, h)$ that maximizes $J(n)$.*

Proof. Suppose there were two maximizers n_1 and n_2. Clearly, a convex combination $n_\lambda = \lambda n_1 + (1 - \lambda)n_2$ would also be feasible. Implementing n_λ in terms of promotion costs would be at least as cheap as implementing λn_1 and $(1 - \lambda)n_2$ separately and adding up the costs. Output, however, will be strictly larger for each fixed pair (t, s) by the concavity assumption on $\tilde{Y}(\cdot)$, which implies the desired result. $\square$

It is worthwhile to note that this argument does not hinge on the assumption of n being continuous or differentiable, nor on any restriction on S_n.

If Y is not strictly concave, matters are slightly more complicated. Take the example from subsection 2.5 with a linear production function: Uniqueness of the planner's problem depends on uniqueness of the partial-equilibrium solution for the agent. If the agent's problem has a unique solution for any starting value of h, then the solution to the planner's problem is unique.

Existence of equilibrium is not a problem computationally, but could not be established formally without making an equicontinuity assumption on the function space for n; see section A.3 for a discussion.

---

[22]For the CES case, this is equivalent to assuming $\rho < 1$.

3.4 Varying T

So far, we had fixed the maximal vintage age T and imposed it on the planner; we will now be concerned with varying T and finding the optimal T^* under the assumption that $\tilde{Y}$ is strictly concave. By the concavity argument in lemma 3.1, there is at most one T^* for which the planner's criterion is maximized. An argument analogous to the proof for 2.4 shows that $T^* < \infty$. However, it is very hard to further characterize T^*. Computationally, it may be found by finding the optimal n for each fixed T and then pick the value T^* that yields the highest value to the planner. The following discussion describes regularities and problems that arose in this process.

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

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

3.5 Equivalence to competitive equilibrium

The following proposition establishes that the global solution to the planner's problem is a competitive equilibrium and a partial converse of this statement:

Proposition 3.2. (Equivalence of planner's solution and competitive equilibrium) *The stationary (global) solution to the planner's problem with T^* is a competitive equilibrium (CE). Any stationary CE is also a solution to a planner's problem for some $T \leq T^*$. There is no CE with $T > T^*$.*

Proof. I will first show that the global solution to the planner's problem constitutes a CE. Set wages $w(\tau, h) = \partial Y(\tau)/\partial n(\tau, h)$ for $\tau \leq T^*$ and

$w(\tau, h) = w(T^*, 0) = \mu$ for all $\tau > T^*$, all h. This implies that firms optimally choose not to produce for $\tau > T^*$ since even the cost-minimizing input combination leads to losses. For $\tau \leq T^*$, $n(\tau, h)$ is an optimal input choice and profits are zero. For agents, the HJB (17) and its boundary conditions imply that any career segment which fulfills $\dot{h} = V_h$ everywhere is an optimal strategy with starting value μ. This weakly dominates any career segment in vintages (T^*, ∞). One may then insert agents into careers to engineer the entry density $n(\tau, 0)$ since agents are indifferent between all careers. Equation (6) ensures that the density n reproduces itself given the optimal decisions of agents.

Second, I prove that any CE is a solution to the planner's problem for some $T \leq T^*$. To start, note that the worker's HJB (2) and the corresponding optimal policy (3) in competitive equilibrium have their exact counterparts in equations (17) and (16) for the planner's problem. Equation (14) follows by integrating from the boundary over τ and h, which in turn is equivalent to (13). Since the first-order conditions (11) and (12) can be used to define the Lagrange multipliers, equation (13) already ensures that all first-order conditions for the Lagrangian hold for any competitive equilibrium.

This means that any competitive equilibrium is a stationary point of the Lagrangian.[23] However, there can be at most one stationary point for a given T since J is a concave function and the set of permissible n is convex. Hence this stationary point must be the global maximum of the planner's problem corresponding to the T induced by the respective CE. As the discussion in 3.4 showed, no such maximizer exists for $T > T^*$, which means that there cannot be any CE with $T > T^*$. □

It is hard to formally rule out competitive equilibria with $T < T^*$. If there is such a CE, then it must be that $\mu_T > \mu_{T^*}$ since these multipliers equal wages in the last vintage. This seems to suggest that $J_T > J_{T^*}$, which would be a contradiction to T^* being associated with a global maximizer. However, as the discussion in 3.6 will show, J also includes the excess value for agents already born at $t = 0$ starting with $h(0) > 0$, which is not comprised in the multiplier μ.[24]

---

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

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

3.6 Vintage productivities

Thinking along the lines of the planner's problem also proves useful in assessing vintage productivities. First, note that we can decompose the planner's criterion by integrating over the single agents' values:

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

where the first equality decomposes the value for the measure one of agents alive at $t = 0$ according to $V = W + \Lambda$ and uses the fact that $\Lambda = 0$ for all agents born later. We can juxtapose this decomposition and the decomposition of J into production and promotion costs:

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

where we write $Y = Y(0)$ and $C = C(0)$. Since $\bar{\Lambda} \geq 0$ and $C \geq 0$, this equation says that labor productivity in the last vintage $w(T,0)$ is lower than average labor productivity Y in the overall economy. Furthermore, it gives us upper bounds for both C and Λ that can be empirically assessed by observing productivity in dying vintages and average productivity in the economy.

4 Calibration and computational results

For the calibration, I choose the CES aggregator defined in (1) for the production function and a quadratic specification for the cost of human-capital accumulation: $c(\dot{h}) = \bar{c} \max\{\dot{h}, 0\}^2/2$. Finding a competitive equilibrium amounts to solving the system of PDEs and integral equations consisting of the agent's HJB (2) and FOC (3) with boundary conditions $V(T^*, h) = V(\tau, 0) = W$ for all h and all τ, the PDE (6) (which describes the evolution of n) and the following wage equation:[25]

$$w(\tau,h) = e^{-\gamma\tau} f(h) \left(\frac{\tilde{Y}(\tau)}{n(\tau,h)} \right)^{1-\rho} \tag{18}$$

---

[25]Note that this system is non-standard in the following respects: First, wages are determined non-locally; they depend not only on the density in the immediate (τ, h)-neighborhood of the agent but also on h-levels in the same vintage that are far away from the agent. Second, we are dealing with a system where the boundary condition $\bar{W}$ is unknown.

I propose a solution algorithm which uses a discretization scheme as in standard lattice methods and attacks the problem of endogeneity of the boundary values with an algorithm inspired by the way a real economy might oscillate around a steady state when some inertia is present. Section B provides the complete documentation of the algorithm.

The model is calibrated to yearly data. I identify vintage age in the model with the age of an establishment in the data; the underlying assumption here is that new establishments incorporate the newest available technology. The death rate δ is set to 0.092; this number is chosen to match the rate at which workers are displaced from establishments in the data (6.7% yearly) plus the rate at which they leave the labor force (2.5%, obtained from an expected labor-market participation of 40 years for a 20-year old[26]).[27] β is set to 0.015 to obtain a standard yearly discount rate of $\beta + 0.025 = 0.04$, where 0.025 is the exit rate from the labor force.

The most parsimonious functional form possible is assumed for returns to experience: $f(h) = A$, where $A = 1$ is chosen as a normalization. Experiments with the specification $F(h) = 1 + a$ showed that identification inside the parameter tuple $(\bar{c}, a)$ was weak and that the specification did not improve upon the fit of the more parsimonious model.

The remaining parameters $\bar{c} = 0.77$, $\gamma = 0.049$ and $\rho = 0.71$ were chosen to minimize the mean squared deviation of the six model moments given in table 1 to their counterpart in the data. A grid search was used to minimize the penalty function. The data used here are the German IAB employment sample, a large employer-employee-matched panel, see appendix C for details.

Table 1 shows that that the model is qualitatively in line with the main features of the earnings distribution in the data: Wages are higher in older vintages on average, but the experience premium is lower in old vintages (implying that entry wages are highest in old vintages). Vintages survive for $T^* = 49$ years. The model over-predicts both the increase of earnings with vintage/establishment age and the decrease of the tenure premium with vintage/establishment age, but is able to generate a reasonable quantitative fit to the orders of magnitude observed in the data.

---

[26]I take the expected retirement age of 60 years for Germans from the German association of retirement insurers, see www.deutsche-rentenversicherung.de.

[27]It is easy to see that a model where agents are displaced from a vintage with probability δ_p and die at rate δ_d yields the same allocations as a model with death rate $\delta = \delta_p + \delta_d$ — human-capital-accumulation decisions do not have any effect on a worker's life after a vintage-displacement shock.

Table 1: Calibration targets

Moment	Establishments	Data	Model
Difference of median log earnings	< 10 years old	−0.13	−0.16
to establishments > 20 years old	10-20 years old	−0.13	−0.05
	< 10 years old	0.41	0.41
5-year tenure premium	10-20 years old	0.34	0.21
	> 20 years old	0.28	0.08
Fraction of workers	> 20 years old	0.56	0.58

(The first five statistics are in log points. The 5-year tenure premium is calculated from the predicted values of a regression of log earnings on a quartic in establishment tenure within the respective group of establishments.)

4.1 Equilibrium properties

Figure 3 is a summary of the results. The career lines in the upper-left panel show agents' equilibrium trajectories through the (τ, h)-space. Agents in young vintages make the hardest efforts to climb the skill hierarchy. This is in line with value differentials in the hierarchy being highest in these vintages, as the value function in the lower-left panel shows. The lower-right panel illustrates that wage compression is a process that happens all the way from the newest to the oldest vintages. The skill premium is highest in the youngest vintages and continuously shrinks as the vintage ages. The density function in the upper-right panel is in line with wages: Workers with high experience in young technologies are the scarcest factor in the economy, whereas skill scarcity vanishes in old vintages as workers press up in the skill hierarchy from below.

Figure 4 shows more variables of interest. In the upper-right panel, we see that entry into vintages is hump-shaped and strictly greater than zero even for the oldest vintages.[28] As apparent in figure 3, late entrants are compensated for learning the least useful skills by the highest entry wages in the economy. The upper-left panel illustrates that despite positive entry, total employment is decreasing in vintage age in the end because incumbent workers are displaced at a faster rate than entrants replace them.

The lower-right panel in figure 4 illustrates labor productivity by vintage age. The pattern is reminiscent of the hump-shaped, back-loaded return profiles that are typical for organization-capital models (see Atkeson & Kehoe, 2005, for example). Young vintages are unproductive because they have a

---

[28]A straightforward calculation shows that the entry density is $m(\tau) = n(\tau, 0)/\dot{h}(\tau, 0)$.
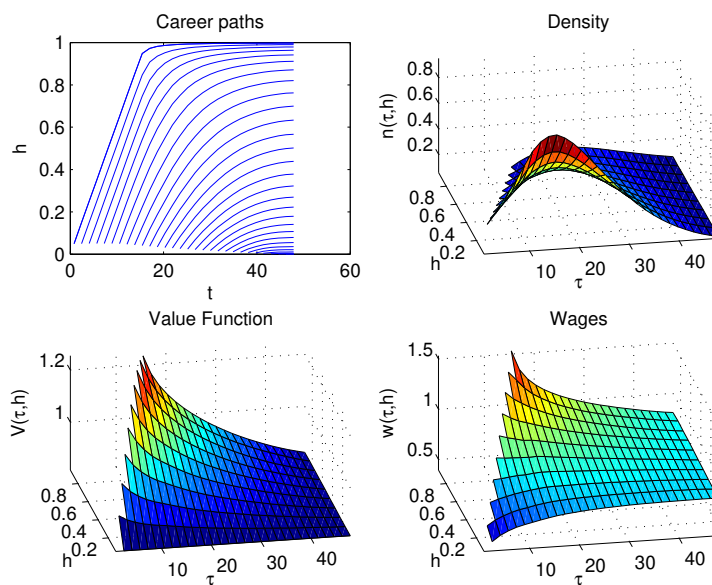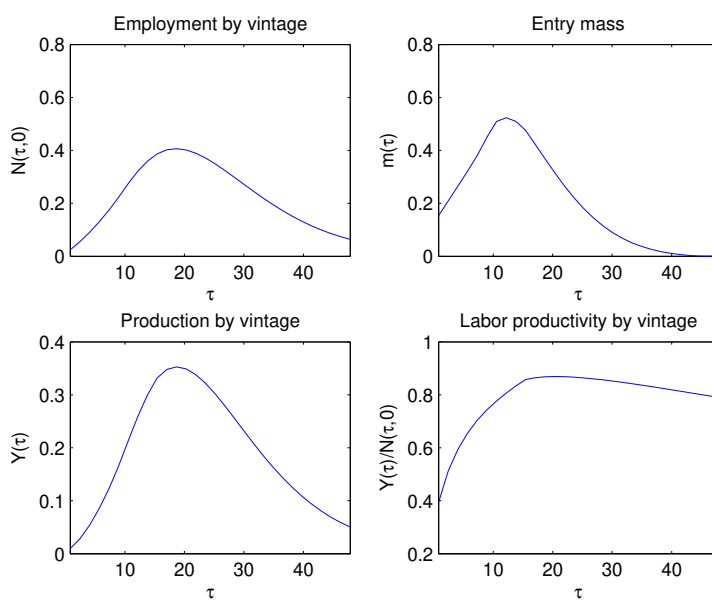
Figure 3: Equilibrium (I)

Figure 4: Equilibrium (II)

very unbalanced mix of labor inputs; marginal returns to the different skill levels are far from equalized since high-skill labor is very scarce. In older vintages, human-capital accumulation leads to gradual equalization of skill returns. However, these returns wear off over time and the negative TFP effect eventually dominates, as section 3.6 pointed out.

4.2 Tenure-earnings profiles

Figure 1 shows tenure-earnings profiles, following a cohort of vintage entrants over time as they accumulate human capital in their respective technology. The curve that extends farthest to the right refers to workers who enter the frontier technology; the shorter the curves become, the later the respective workers enter the vintage.

To understand the forces at work in these profiles, it is useful to decompose the growth of log-wages into its different components. Consider infinitesimal changes in log-wage along a career $\{h(t), \tau(t)\}$ using (18):

$$\left. \frac{d\ln w}{dt} \right|_{\tau(t), h(t)} = \dot{h}(\tau, h)\frac{f'(h)}{f(h)} + (1-\rho)\left(\frac{\partial \ln \tilde{Y}(\tau)}{\partial \tau} - \frac{d\ln n[\tau(t), h(t)]}{dt} \right). \quad (19)$$

The three terms on the right-hand side have a clear economic interpretation: I term them (from left to right) the *experience effect*, the *organization-capital effect* and the *obsolescence effect*. The latter two can be combined into a *relative-supply effect*.

The *experience effect* captures returns from learning. This is the only effect present when skills are perfect substitutes ($\rho = 1$). In figure 3 we see that human-capital accumulation $\dot{h}$ is always positive but decreasing over all careers. Since the learning function f was assumed to be concave decreases over a career, this means that the experience effect is always positive but wears off over the course of a career.

The innovation in the model presented here with respect to the models in the literature lies in the terms that are switched on when lowering ρ below one. These are stemming from relative factor supply. The second term involving production $\tilde{Y}(\tau)$ is always increasing; it represents the gains from joint learning and the equalization of factor returns, which grow as the vintage ages and its skill mix improves. I call this term the *organization-capital effect*.

Finally, the *obsolescence effect* (which involves the density n evaluated along the career path) is key for understanding why earnings profiles are decreasing for most workers towards the end of their careers. As a vintage

ages and more agents enter it, once-scarce skills become more abundant as workers from the lower ranks press up in the skill hierarchy. A real-world example for this might be an HTML-programmer whose skills commanded high returns when the Internet was in its infancy but saw his wages dwindle as more and more other programmers learned HTML and his knowledge became less scarce. Proposition 2.15 shows that – under any parameter constellation –, for some skilled workers in young vintages this effect indeed overrides the other two and leads to earnings losses.

An interesting feature of the wage profiles generated by the model is that they have heterogeneous slopes and curvature. In Ben-Porath-type models, heterogeneity in shape is usually attained by assuming heterogeneous learning ability, see for example Guvenen & Kuruscu (2007) and Huggett, Ventura & Yaron (2006). In contrast to these models, agents here are ex-ante equal and all heterogeneity is ex-post and endogenous. In fact, heterogeneity in earnings profiles is essential in order to give workers the incentives to enter *all* existing vintages and to ensure that all vintages have an efficient skill mix.

Another topic from the labor literature addressed by the model is "overtaking". Hause (1981) defines overtaking as the fact that two wage profiles with different slope but the same present value have to intersect at a certain point. The model has precise predictions on when this overtaking point occurs for different pairs of agents in the economy. In the calibrated model, overtaking happens throughout the first decade of workers' careers, as is evident from figure 1.

4.3 Comparative statics

When varying parameters with respect to the baseline calibration, steady-state characteristics change as follows:

An increase in γ creates incentives to abandon old vintages earlier since frontier TFP is increasing faster. For example, an increase from 0.45% to 0.9% cuts T^* (the age of the oldest vintage) from 49 to 34 years, which of course lowers the fraction of workers employed in old technologies. Tenure premia rise in young vintages but fall in older ones. The difference between median wages in young vintages and old vintages, however, decreases with respect to the baseline calibration.

Changes in ρ have ambiguous effects on T^*. There seem to be two fundamental effects that work in different directions. A first (partial-equilibrium) effect is always positive: The higher ρ, the less important complementarities between skills in determining production and wages; this takes away the

only advantage that older technologies have over younger ones – their more balanced skill mix –, so it makes sense to scrap vintages earlier. A second (general-equilibrium) effect works in the opposite direction: The more important complementarities, the higher the returns for highly-skilled workers in young technologies — this leads more entrants to choose young technologies and to acquire skill faster, which leads to a balanced skill mix inside a vintage earlier and encourages early scrapping. Indeed, in the parameter region around the preferred calibration, the second effect is stronger: Lowering ρ from 0.71 to 0.60 lowers the optimal scrapping time T^* from 49 to 44 years. Tenure and skill premia in young vintages increase as expected: Stronger complementarities mean that scarce factors reap higher rewards. However, tenure and skill premia decrease in over-20-year-old vintages.

An increase of $\bar{c}$ (the cost of skill accumulation) unambiguously leads to a shortening of the vintage horizon T^*: Doubling $\bar{c}$ from 0.75 to 1.50 increases T^* from 49 to 65 years, for example. It takes agents longer to accumulate skills, which leads to longer vintage lifetime. It turns out that changes in $\bar{c}$ do not have systematic effects on tenure premia. This is probably due to the fact that there are two competing forces at play: On the one hand, slower skill accumulation exacerbates skill scarcity and hence increases skill premia; on the other hand it also slows down workers' rise in the skill ladder and thus leads to lower skill differences between workers a given difference in tenure.

4.4   More evidence

Note that the identification strategy pursued so far has relied on the assumption that younger establishments operate with technology of a newer vintage. However, the model still has strong implications for the data even when this assumption is dropped.

To see this, suppose that a certain industry has plants that are predominantly of new vintages. Then the model makes us expect higher tenure premia and lower mean wages in this industry than in an industry mainly comprised of old-vintage plants. The same should be true for occupations or even establishments: Occupations/Establishments which make heavier use of new vintages should display higher tenure premia and lower mean wages. Now, note from figure 4 that young vintages grow strongly in both employment and output, whereas older vintages are contracting; so industries/occupations/establishments (IND/OCC/EST) with predominantly new vintages grow faster according to the model. To summarize, we would expect high tenure premia and low mean wages in fast-growing IND/OCC/EST.

Table 2 shows the results of a regression of cross-sectional earnings on worker (IND/OCC/EST-)tenure, IND/OCC/EST growth and quadratic and interaction terms of the two. The interaction terms are all positive, as predicted by the model: Faster growing IND/OCC/EST have higher tenure premia. The coefficients on growth (g) have the expected negative sign for IND and EST, but a positive sign for OCC: Faster-growing IND/OCC/EST have lower mean wages. So overall, the regression results support the model's predictions.[29]

Table 2: Regressions on growth and earnings structure

Category	N	$const$	ten	$ten^2/100$	g	$g^2/100$	$ten \times g$
Industry	208,986	4.930	.055	-.141	**-.113**	.065	**.023**
Occupation	209,721	4.923	.060	-.164	**.189**	-.221	**.027**
Establishment	152,383	5.043	.0386	-.098	**-.005**	.007	**.006**

Censored regression (top-coded data) of log earnings on a second-order polynomial in IND/OCC/EST tenure and IND/OCC/EST growth, both referring to the respective category. IND/OCC/EST-growth g is calculated as $\ln M(2000) - \ln M(1995)$, where M is the number of workers in the respective category. Tenure and earnings data are from 2000. All coefficients – except for the one on g for EST – are significantly different from 0 at the 1%-level.

The magnitude of the effects is considerable: An IND (OCC, EST) that grows by 1% on a yearly basis has a 5-year tenure premium that is 0.58 (0.67, 0.14) percentage points higher than that of a stagnant one. In an IND (OCC, EST) that grows by one standard deviation faster than a stagnant one, the 5-year tenure premium is 2.38 (2.42, 1.92) percentage points higher than in stagnant one.

5 Conclusions

This paper has studied a model of vintage-human-capital accumulation that matches key facts on the tenure-earnings distribution in a German data set. It provides a promising avenue for understanding the systematic variation in the earnings structure across establishments, industries and occupations.

---

[29]These empirical patterns are in line with Michelacci & Quadrini's (2004) results from Finnish matched-employer-employee data. They find that in fast-growing firms returns to tenure are highest and starting wages are lowest . Their model explains this phenomenon by financial constraints that are especially severe for fast-growing firms, inducing firms to "borrow" from their workers by offering back-loaded tenure-earnings profiles.

In the following, some potential applications of the framework are briefly discussed.

A first proposed application is a macroeconomic one: The model relates the rate of embodied technological growth to the earnings structure, both at the industry and the economy-wide level. Previous versions of the paper had focused on this point, arguing that the steepening of age-earnings profiles and the concomitant rise in cross-sectional and time-series variance of earnings in many industrialized countries over the last decades could have been caused by a technological acceleration.

Second, the model tells us how the extant skill structure across industries influences sector choice and human-capital-accumulation decisions of young people. An implication is that young Germans are put to productive use entering the relatively old car industry and complementing their skilled ancestors, whereas young Indians are doing well concentrating in new industries (becoming programmers, for example). Since there is a lack of established industries with a large stock of existing human capital in India, young Indians do not have the same incentives as their German counterparts to enter old sectors. The model implies that the described sector-choice patterns are efficient given the extant skill structure in each country.

A third aspect worth mentioning, which has only been touched upon in the previous discussion, is the productivity profile of a vintage over time (see the lower-right panel of figure 4). It displays the typical back-loaded shape that is often posited in an ad-hoc fashion for organization capital (see Atkeson & Kehoe, 2005, for example). In fact, the model presented here can be construed as a micro-foundation for the way an organization increases its productivity over time and how it shares these productivity gains among its members.

Finally, one could study the riskiness of human capital and technology choice for workers by introducing a stochastic component into the framework.

References

Atkeson, A. & Kehoe, P. J. (2005), 'Modeling and measuring organization capital', *Journal of Political Economy* **113**, 1026–1053.

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

Burdett, K. & Coles, M. (2003), 'Equilibrium wage-tenure contracts', *Econometrica* **71**(5), 1377–1404.

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

Drews, N., Hamann, S., Köhler, M., Krug, G., Wübbeke, C. & Autorengemeinschaft 'ITM-Benutzerhandbücher' (2006), 'Variablen der schwach anonymisierten version der iab-beschäftigten-stichprobe 1975-2001', *Benutzerhandbuch des Forschungsdatenzentrum des Instituts für Arbeitsmarkt- und Berufsforschung* .

Garicano, L. & Rossi-Hansberg, E. (2008), 'Organizing growth', *Working Paper* .

Guvenen, F. & Kuruscu, B. (2007), 'A quantitative analysis of the evolution of the U.S. wage distribution: 1970-2000', *Working Paper* .

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

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

Kushner, H. & Dupuis, P. (1992), *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer Verlag, Berlin and New York.

Lucas, R. E. (1988), 'On the mechanics of economic development', *Journal of Monetary Economics* **22**(1), 3.

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

Michelacci, C. & Quadrini, V. (2004), Financial markets and wages, 2004 Meeting Papers 116, Society for Economic Dynamics.

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

Prescott, E. C. & Boyd, J. H. (1987), 'Dynamic coalitions: Engines of growth', *The American Economic Review* **77**(2), 63–67.

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

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

A   Additional proofs

A.1   Proofs on the competitive equilibrium

A.1.1   Bounded resources

Lemma A.1. (Bounded resources) *There is a uniform bound $\bar{y} < \infty$ on $\tilde{Y}(n)$, $\int n \leq 1$. Thus, resources in the economy are bounded for each fixed t.*

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

A.1.2   Proof of lemma 2.3: All jobs filled in producing vintage

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

A.1.3   Proof of lemma 2.9: Paths cross at most once ($\rho = 1$)

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

$$h(t) - h(s) \geq g(t) - g(s) \Rightarrow \int_s^t \dot{h}(u)du \geq \int_s^t \dot{g}(u)du, \qquad (20)$$

i.e. h must grow by at least as much as g over the interval to end up above. By the assumption on the wage function, $w_h(h)$ is a decreasing function in h. Using the

FOC (3), this implies that for all $s < u < t$, we have

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

since by assumption $\dot{h}(t) < \dot{g}(t)$ and $w_h[h(v)] \leq w_h[g(v)]$ point-wise; the inequality follows from c' being increasing. This again implies $\dot{h}(u) < \dot{g}(u)$ for all u, which in turn contradicts (20). $\square$

A.1.4 Proof of proposition 2.11: Faster growth shortens careers and lowers tenure premia ($\rho = 1$)

Proof. Let $W(\gamma)$ be the value of being an inexperienced worker at time 0 given vintage productivity growth γ. His problem is then to choose the switching time T when to leave the vintage to maximize

$$V(\gamma, T) = J(T) + e^{-\rho T} W(\gamma),$$

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

$$V_T(\gamma, T) = J'(T) - (\rho - \gamma)e^{-(\rho - \gamma)T} W(\gamma) \tag{21}$$
$$V_{TT}(\gamma, T) = J''(T) + (\rho - \gamma)^2 e^{-(\rho - \gamma)T} W(\gamma), \tag{22}$$

The FOC for the optimal career length $T^*(\gamma)$ is $V_T(\gamma, T^*(\gamma)) = 0$, the SOC is $V_{TT}(\gamma, T^*(\gamma)) < 0$.

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

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

where the envelope condition $W_T(\gamma, T^*(\gamma)) = 0$ is used.

Now, re-state the first-order condition for $T^*(\gamma)$ from (21):

$$J'\big(T^*(\gamma)\big) = (\rho - \gamma)W^*(\gamma)e^{-(\rho - \gamma)T^*(\gamma)}$$

Take the total derivative of this equation with respect to γ and use (23) to obtain

$$\frac{dT^*}{d\gamma} \underbrace{V_{TT}\big(\gamma, T^*(\gamma)\big)}_{<0 \text{ by SOC (22)}} = \underbrace{(\rho-\gamma)e^{-(\rho-\gamma)T^*(\gamma)}W^*(\gamma)}_{>0} \left[\underbrace{\frac{T^*(\gamma)}{1-e^{-(\rho-\gamma)T^*(\gamma)}} - \frac{1}{\rho-\gamma}}_{\equiv \Phi_\gamma(T^*)} \right]$$

(23)

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

The derivative of Φ_γ in T^* is

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

Note that in the numerator, $1 + (\rho-\gamma)T^*$ is nothing but the first-order Taylor expansion of the function $e^{(\rho-\gamma)T^*}$ in T^* around 0, which always below the function itself since the exponential function is convex. This implies that the fraction in the numerator is always smaller than one, implying that Φ_γ is an increasing function all the way from zero to infinity. This in turn implies $\Phi_\gamma > 0$ (recall that $\lim_{T^*\to 0} \Phi(T^*) = 0$), which tells us we have $dT^*/d\gamma < 0$ for all $\gamma > 0$.[30]

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

Finally, since $T^*(\gamma)$ is a decreasing function it follows from proposition 2.10 that $p_\gamma(t)$ is decreasing in γ for any fixed $t > 0$. $\qquad\square$

A.1.5 Proof of lemma 2.13: Wage explosion for the skilled in young technologies

Proof. First, I show that the cost $C(\Delta h, \Delta t)$ of accumulating human capital Δh in a time interval Δt goes to infinity for fixed Δh when letting $\Delta t \to 0$. By Jensen's inequality, the minimal cost of accumulating Δh within Δt is by setting a constant $\dot{h} = \Delta h/\Delta t$ throughout Δt. Then $C(\Delta h, \Delta t) \geq c(\Delta h/\Delta t)\Delta t \to \infty$ if $\Delta t \to 0$ by our assumption on $c(\cdot)$.

By lemma 2.7, S_n must be a rectangle $(T_0, T^*) \times [0,1]$. Now, suppose there was no singularity for w in the upper left corner and $w(T_0, 1) < \infty$. Then, by continuity of w, for each ϵ there is a ball $B_\delta(T_0, 1)$ in which wages deviate not more than ϵ from $w(T_0, 1)$. So the parts of any career segment contained in $B_\delta(T_0, 1)$ yield bounded wage payments. But we can definitely find a sequence of careers for which learning costs inside the ball exceed any bound. To see this, set $\Delta h = \delta$, take a sequence $\Delta\tau \to T_0$ and note that the cost of reaching $(T_0 + \Delta\tau, 1)$ inside B must go to infinity. Note also that a positive measure of workers must take such paths since no region is empty by lemma 2.3. But then, those workers cannot behave

---

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

optimally and should change their h-path through the ball B, which is inconsistent with equilibrium.

This also implies that $T_0 = 0$. If this was not the case, then workers with careers in $B_\delta(T_0, 1)$ should reach those by choosing flatter careers entering at $\tau = 0$, which would imply that those careers could achieve unbounded value by the above argumentation. This contradicts $W < \infty$. $\qquad\blacksquare$

A.2 Planner's cost-minimizing promotion strategy

Lemma A.2. (No-crossing measure is optimal) *For a given density $n(t, s, h)$, it is optimal for the planner not to let career paths cross when implementing the density. This means that the planner makes workers follow paths $h(t+u)$ for any given t, any vintage s and any $u \in (0, T-t)$ such that $N[t+u, s, h(t+u)] = \exp(-\delta u) N[t, s, h(t)]$.*

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

Now, we will construct a lower bound C_k^* on this cost for a fixed iteration k in the algorithm. Note that it is enough to consider the task of moving workers between t_i and t_{i+1} for each point in time. It does not matter how we combine these path segments sequentially later, any such combination must obviously yield the same value.

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

Also, notice that there always exists a policy which does not make any worker flows cross: First, fill the uppermost interval at $t = 1$ with the uppermost workers from $t = 0$; proceed by filling the second interval with the uppermost workers left at

$t = 0$ after the first step, and so forth. It is also clear that any process that does not follow these rules must make some workers cross and that any such process can be rendered into the proposed no-crossing algorithm by a finite number of improving operations; this shows that the no-crossing mechanism is optimal for a fixed k.

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

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

A.3 Existence of solution to the planner's problem

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

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

B Computational algorithm

The following method discretizes the state space into a finite number of vintages and a finite number of rungs in the skill ladder. The algorithm can be interpreted as introducing a random element to skill accumulation (see Kushner & Dupuis, 1992, on the approximation of continuous-time models by discrete-time Markov chains). The algorithm is not only useful compute an approximation to the equilibrium, but also to form some intuition about the value function, agent's paths and other objects of the (continuous) model. The death probability δ is set to zero to simplify

the exposition; of course, all arguments presented here also apply to the case $\delta > 0$. Throughout, we only consider stationary allocations, i.e. variables depend only on (τ, h) but not on t.

First, construct a discrete grid on the rectangle ($0 \leq \tau \leq T, 0 \leq h \leq 1$ as follows: Divide the vintages into S sub-intervals (of equal size $\Delta\tau$) and the experience levels into h sub-intervals (of equal size Δh). The center points of these intervals are denoted by $\{\tau_i\}_{i=1}^T$ and $\{h_j\}_{j=1}^S$.

To approximate skill approximation choices $\dot{h}$, we linearly interpolate the value function between adjacent cells. If the grid is such that workers climb less than Δh over a time interval of $\Delta\tau$ in all cells, than linear interpolation is equivalent to the following "stochastic careers": Set the probability $p(\tau_i, h_j)$ that the agent moves one box up (from h_j in vintage τ_i to h_{j+1} in vintage τ_{i+1}, that is) such that the expected slope of his career equals $\dot{h}(\tau_i, h_j)$, but that it does not exceed one:

$$p(\tau_i, h_j) = \min\left\{\dot{h}(\tau_i, h_j)\frac{\Delta\tau}{\Delta h}, 1\right\}$$

This means that in order to be able to replicate very steep slopes in this fashion, we need to make the slope $\Delta h/\Delta\tau$ become successively greater as k grows. I will make the following limiting argument: If we have an infinite sequence of discrete approximations as described above, choose the number of grid points as $S_k = kS_0$ and $H_k = k^{3/2}H_0$ (the reason for this choice will become clear later). Now, since the number of grid points for the hierarchy grows faster than the number of grid points for vintages, the maximal possible slope $\Delta h_k/\Delta\tau_k$ will grow to infinity, so any slope $\dot{h}$ will be covered from some k on, and all points in the upper-left corner of the rectangle will be reached by some mass from some k on. Of course, for each given grid size there still might be some cells in which the bound 1 is reached.

Consider now how the density of workers evolves on the grid:

$$n(\tau_{i+1}, h_j) = \big[1 - p(\tau_i, h_j)\big]n(\tau_i, h_j) + p(\tau_i, h_{j-1})n(\tau_i, h_{j-1}).$$

Now, introduce the (upward-) difference operators $\Delta_h f(\tau_i, h_j) = f(\tau_i, h_{j+1}) - f(\tau_i, h_j)$ and $\Delta_\tau f(\tau_i, h_j) = f(\tau_{i+1}, h_j) - f(\tau_i, h_j)$ for arbitrary functions $f(\cdot, \cdot)$. Then we can re-write the above as

$$\Delta_\tau n(\tau, h) = -\Delta_h\Big[n(\tau, h-1)p(\tau, h-1)\Big] = -n(\tau, h-1)\Delta_h p(\tau, h-1)$$
$$- p(\tau, h-1)\Delta_h n(\tau, h-1) - \Delta_h n(\tau, h-1)\Delta_h p(\tau, h-1).$$

Note that the last term on the right-hand side will become small compared to the others when the grid becomes very fine. In the limit, the equation becomes equivalent to the mass-transport PDE (6) that describes the behavior of the density n.

Production in a vintage $\tilde{Y}$ (excluding the TFP term $e^{-\gamma\tau_i}$) is calculated as

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

where again the function f is evaluated in the middle of the corresponding box (τ_i, h_j). This expression converges to $\tilde{Y}(n(\tau_i, \cdot))$ (under mild conditions) for a given function $n(\cdot)$ as $\delta h \to 0$.

The discrete counterpart for wages is

$$w(\tau_i, h_j) = \exp[-\gamma \tau_i] f_j \left(\frac{\tilde{Y}(\tau_i)}{n(\tau_i, h_j)} \right)^{1-\rho}. \tag{24}$$

Note that this gives the wage rate *per unit of time.* If we want to calculate the counterpart to wage payments over time a worker spends inside a box (τ_i, h_j), of course we have to multiply this wage rate by $\Delta \tau$.

The discrete counterpart of the value function is

$$V(\tau_i, h_j) = w(\tau_i, h_j)\Delta \tau + e^{-(\beta+\delta-\gamma)\Delta \tau} V(\tau_{i+1}, h_j) + \tag{25}$$
$$= \max_{\dot{h}} \left\{ -\frac{c}{2}\dot{h}^2 \Delta \tau + \dot{h}\underbrace{\frac{\Delta \tau}{\Delta h}}_{=p} e^{-(\beta+\delta-\gamma)\Delta \tau} \Delta_h V(\tau_{i+1}, h_j) \right\}.$$

Since agents only move upward in equilibrium, we take the upward-difference to approximate the h-derive of V in the spirit of upwind-differencing.

Solving for the optimal policy gives us

$$\dot{h}^*(\tau_i, h_j) = \frac{e^{-(\beta+\delta-\gamma)\Delta \tau}}{c} \frac{\Delta_h V(\tau_{i+1}, h_j)}{\Delta h}, \tag{26}$$

which converges to the optimal policy from the agent's first-order condition in the continuous case. Plugging back in, we obtain the Bellman equation

$$V(\tau_i, h_j) = w(\tau_i, h_j)\Delta \tau + e^{-(\beta+\delta-\gamma)\Delta \tau} V(\tau_{i+1}, h_j) + e^{-2(\beta+\delta-\gamma)\Delta \tau} \frac{1}{c} \left(\frac{\Delta_h V(\tau_{i+1}, h_j)}{\Delta h} \right)^2 \Delta \tau.$$

When dividing this equation by $\Delta \tau$ and taking the limit as $\Delta \tau \to 0$, we obtain the Hamilton-Jacobi-Bellman equation (HJB) (2) for the continuous case.

I solve the system for a given rectangle with length T using an algorithm that is inspired by how a real economy might converge to a steady state under adaptive expectations, assuming some inertia in agents' actions. Given a distribution of agents n_k (where k indexes the iterations of the algorithm) over the grid, one can calculate the resulting wages from (24). Using the fact that the marginal value of skill is zero when the vintage dies (i.e. $\Delta_h V^{(k)}(\tau_{T+1}, h_j = 0$ for all j), we can back out the value function recursively going from τ_T back to τ_1 using (25), which also yields optimal policies $\dot{h}^{*(k)}$ from (26).

As for the promotion efforts $\dot{h}$, we now mix some of the optimal policies into the existing ones: $\dot{h}^{(k+1)} = \alpha \dot{h}^k + (1-\alpha)\dot{h}^{*(k)}$. As for the entry decisions, I send more mass into the starting points with higher value and less mass into those with higher value. Since wages are inversely related to the density, this algorithm drives the

system towards an equilibrium if the tuning parameters are chosen right. Further work is required to prove that this algorithm is indeed a contraction.

To find T^*, the vintage horizon that is optimal from the planner's point of view, I vary T and find a density n_T by the algorithm above. I then choose T^* as the horizon T that maximizes the planner's criterion described in the beginning of section 3.

The complete *Matlab* code used in the calibration and more detailed documentation are available from the author upon request.

C Data

This study uses the weakly anonymous IAB Employment Sample (years 1975-2001). Data access was provided via on-site use at the Research Data Center (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB) and remote data access.[31] The data set is a 2% random sample of all Germans covered by the mandatory public unemployment-insurance scheme.[32] Every individual holding a job that fell under this scheme for at least several weeks at any point of the period 1975-2001 was at the same 2% risk of being sampled. For every sampled individual, all available employment spells were collected and included in the data set. Available characteristics include pre-tax earnings, gender, age and 3-digit occupation code of the person as well as an identifier of the employer's establishment and a 3-digit industry classification of the establishment. As is common in the literature, I restrict the sample to males who work full time and are between 20 and 61 years old. For consistency reasons, only workers born in former West Germany are considered. The next section provides a more detailed description of the data and the exclusion restrictions.

The data used in this paper are the *Employment Samples* provided by the *Research Data Center* of the German unemployment office [33]. They were collected for administrative purposes by the mandatory unemployment-insurance (UI) system in Germany. From 1975 until 2001, spell data about the employment situation were collected at least once yearly from all German employees that were subject to contributions to the unemployment-insurance system. Among full-time employees, this excludes only the self-employed and *Beamte* (public-sector employees with life-time tenure). On the other hand, data are available for all unemployed workers who were paid benefits out of the UI system — the latter are not used in this paper.

---

[31]See Drews, Hamann, Köhler, Krug, Wübbeke & Autorengemeinschaft 'ITM-Benutzerhandbücher' (2006) for an excellent documentation of the IAB Employment Sample.

[32]The sample does not include tenured public-sector employees and the self-employed; these groups are not overwhelmingly large so that the data set can be seen as representative of the German labor market.

[33]In German, the data are succinctly called the: (IAB) *Beschäftigtenstichprobe*, provided by the *Forschungsdatenzentrum* (FDZ) of the *Bundesagentur für Arbeit* (BA) at the *Institut für Arbeitsmarkt- und Berufsforschung* (IAB).

The sampling design for the IABS is as follows: Every German individual who was subject to paying contributions to the UI system at any point between 1975 and 2001 was sampled at a uniform probability of 2%. Once an individual was sampled, all data from the UI system (work spells and unemployment spells) that could be matched to the individual were included in the sample. The data set consists of more than 12 million spells for more than 1 million individuals.

Individual-specific data include a person's age, sex and a measure of education[34]. For each spell, daily earnings, an establishment ID and the worker's 3-digit-level occupation classification are available. At the establishment level, some information is available that was obtained from aggregates over the original set of administrative data before the 2%-sample was drawn. These data include 3-digit-level industry classification, number of employees in the establishment in the respective year, and the first and last date between 1975 and 2001 in which the establishment hired a worker subject to UI contributions.

I only consider records of male individuals who are both older than 20 years and younger than 62 years in the beginning of the year under consideration. I only consider spells coming from the *BeH*, the database for work relationships – all spells stemming from *LeH*, the database for unemployment-insurance payments, are automatically excluded from the sample. Also, I consider only full-time employees (`stib`< 8). Furthermore, all spells that are marked as "geringfügige Beschäftigung" (tax-exempt part-time employment) are dropped.

Note that apprentices and interns are *included* in the sample. This is a deliberate choice; since these employees constitute arguably a considerable fraction of the labor force that has no job-specific skills yet, it would not be desirable to discard this information in a study on human-capital accumulation.

Also, I only consider individuals whose first employment is with an establishment located in former West Germany. This is done in order to ensure comparability of the results before and after 1990.[35] Furthermore, some quality checks are performed on the data: Spells of individuals for whom more than one full-time job is declared are discarded. Also, spells with unreasonably low daily earnings are deleted (below 7 Euros in 2000 Euros per day).

Earnings are adjusted for inflation using the consumer price index for West Germany provided by the *Bundesbank* (the German Central Bank).

Stata programs and documentation on how the moments in section 4 were obtained in detail are available from the author upon request.

---

[34]This education measure is only filled for employment contracts where education information is necessary to determine UI contributions or benefits, so its information content is limited.

[35]1990 is the year of German re-unification. Note that individuals born in the former East who have moved West before their first job are included. This should have no major bearing on the results since it is reasonable to assume that these individuals do not differ systematically from West Germans in terms of earnings potential — the education systems in both parts are of similar quality.