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[†]

28 July 2008

Abstract

I combine an infinite-horizon version of Ben-Porath's (1967) model of human-capital accumulation with a vintage structure as in Chari & Hopenhayn (1991). Different skill levels inside a vintage are complementary in production. Vintage-specific human capital is accumulated based on workers' optimal strategies and is lost when the technology is phased out by an endogenous firm decision. I establish equivalence between competitive equilibrium and a planner's problem. It is shown that returns to skill are highest in young vintages. Accelerated technological change shortens the life cycle of a technology and speeds up obsolescence; the premium on tenure rises because more workers are concentrated in young technologies with high skill premia. A calibration exercise comparing two steady states shows that the model quantitatively accounts for the changes in the experience premium, earnings dispersion and earnings turbulence in German data.

Keywords: Vintage human capital, age-earnings profiles, partial differential equations
JEL codes: J01, E24, C63

1 Introduction

The neo-classical model of human-capital accumulation (Ben-Porath, 1967) is very successful in modeling the structure of age-earnings profile in the

*This paper is part of my dissertation thesis at New York University. I thank Boyan Jovanovic, Jess Benhabib, Jonathan Eaton, Hugo Hopenhayn, Greg Kaplan, Silvana Melitsko, Gianluca Violante as well as participants at various seminars for discussion and comments.

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

cross section.[1] However, there have been prominent changes in age-earnings profiles in industrialized countries over the past decades (see Gottschalk & Smeeding, 1997, for a review). This paper will argue that the neo-classical framework fails to provide a coherent theory for these changes and provides an alternative explanation based upon the notion of vintage-specific human capital.

In a companion paper (Kredler, 2008), I document the changes in age-earnings profiles for the case of Germany in the years 1975-2001 in the *IAB Employment Sample* at the *Institute for Employment Research* (IAB), a large administrative data set of German workers provided by the German Federal Employment Agency (BA). The following two trends are highlighted: First, the experience premium (defined as the difference in median earnings between workers in their 40s and workers in their 20s) has increased from 13 to 33 log points over the sample period; similar (or higher) numbers are found when restricting the sample to educational subgroups. Second, earnings turbulence (the variance in the growth rate of individual workers' earnings) and earnings dispersion (the cross-sectional variance in earnings) have increased in lockstep with a rising mobility of workers between jobs.[2] In a linearized version of a Ben-Porath-type model, Kredler (2008) shows that implausibly high changes in parameters are necessary to reconcile the neo-classical human-capital theory with the observed increase in the experience premium.

Models like Huggett et al.'s (2006), which enriches Ben-Porath's model with individual-specific learning ability and shocks to human capital, are most likely to inherit these properties. There is a better chance to match the observed hike in the experience premium for frameworks in which human capital comes in different types; these different types can command different experience premia, and the composition of types in the population may change over time. Guvenen & Kuruscu's (2007) model, for example, is in this vein. Indeed, their model generates an increase in the experience premium over recent decades. This rise is driven by an exogenous change in the marginal rate of technical substitution (MRTS) between raw labor and skilled labor, skilled labor being analogous to human capital in the Ben-Porath framework.

The model that I propose differs from Guvenen & Kuruscu's (2007) in the following respects: First, it generates a rise in the experience premium

[1] Recent examples include Guvenen & Kuruscu (2007) and Huggett, Ventura & Yaron (2006).

[2] See the boxplots in figures 5 and 6 for a quick overview of the changes in the cross-sectional distribution of earnings and earnings growth rates by age in Germany.

without changing the MRTS between skilled and unskilled in production, relying solely upon the endogenous forces of skill accumulation by workers and technology choice by firms. Second, it does so in a way that is intimately linked to technological progress and earnings turbulence; it thus has predictions on a richer set of observables.

An entirely different class of models that are 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 differ in the following ways from Burdett & Coles's (2003): First, in Burdett & Coles (2003) changes in *employer* are vital for the determination of wage profiles, whereas in my setting, changes in the *technology* a worker is employed in are crucial. Second, Burdett & Coles (2003) predict that workers' wage profiles never cross; in my setting, this crossing (often termed "overtaking" in the literature) is necessary to make workers indifferent between entering different technologies.

In modeling the environment, I depart from the observation that developments similar to those in Germany have occurred in most other industrialized countries. In a survey of labor-market studies conducted in the 1990s, Gottschalk & Smeeding (1997) and Gottschalk & Joyce (1998) report that the experience premium had increased in six countries over the 1980s, remained constant in one and that the evidence was inconclusive in three countries.[3] They also cite evidence that earnings dispersion has increased in almost all countries.

Led by the observation that the empirical phenomena found in the German data are shared by many other countries, I seek a technological explanation to account for the changes in age-earnings profiles. Specifically, this paper shows that an acceleration in the pace of technological progress –as has been measured over the last decades by Cummins & Violante (2002) based upon work by Gordon (1990)– can generate a steepening of age-earnings profiles alongside an increase in dispersion and turbulence.[4]

[3] Gottschalk & Smeeding (1997) cite a study according to which the experience premium did not increase in Germany; since the sample period of this study is far shorter than the one considered in my data, I discard this study from the above count. For the U.S., the evidence on these phenomena is mixed: There is a consensus in the literature that earnings dispersion has increased over recent decades, but for the experience premium the case is less clear. Gottschalk & Smeeding (1997) find an increase in the experience premium, but studies like Heckman, Lochner & Todd (2003) find no important changes over the last decades.

[4] Violante (2002) shows that accelerated technological change can lead to an increase in residual earnings dispersion when human capital is vintage-specific. However, in his

I augment the neo-classical model of human-capital accumulation with a vintage structure as in Chari & Hopenhayn (1991). Human capital is tied to a technology and is lost when the technology is phased out. Unlike in Chari & Hopenhayn's (1991) model, human-capital accumulation is endogenous and the possibly infinite lives of individuals allow for rich patterns in age-earnings profiles. In each vintage, different levels of human capital are complementary inputs to a production function. I show equivalence between the market equilibrium and the planner's solution, which implies uniqueness of equilibrium.

In the model, the premium on technology-specific skills is highest in young technologies. This result is driven by the relatively larger scarcity of skill in young technologies. As technologies age, this premium shrinks and eventually vanishes entirely.[5] When the speed of technological progress increases, firms switch to the frontier vintage faster. More workers are then employed in new technologies with a high experience premium, which increases the average experience premium. Also, workers lose vintage-specific human capital more often, which induces higher turbulence and dispersion in earnings profiles.

I prove that faster churning of technologies (which occurs in the wake of a technological acceleration) can be accompanied by a hike in the experience premium only if high- and low-human-capital inputs are complementary in production. Under substitutability, in fact, the shortened life span of a technology decreases incentives for human-capital accumulation and flattens the productivity- and thus the age-earnings profiles.

The model's parameters are calibrated to match key moments in the German data. A comparison between two steady states shows that the model can quantitatively account for the secular changes in age-earnings profiles.

The remainder of the paper is organized as follows: Section 2 presents the model and analyzes the competitive equilibrium. Section 3 shows that the model's equilibrium is equivalent to the solution of a planner's problem. Section 4 presents the calibration results. Section 5 concludes and discusses potential further applications of the framework.

framework human-capital accumulation follows a learning curve and is thus exogenous.

[5]The continuous-time setting allows me to prove that the wage structure in the last vintage is completely flat.

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: Newer vintages are more productive holding input ratios equal and the production function is complementary in its inputs. Specifically, I impose

$$Y(t,s) = e^{\gamma s} \tilde{Y}\left[n(t,s,\cdot)\right]$$

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 $\tilde{w}[n]$ exists everywhere, is continuous in n and $\tilde{w}[n](h) > 0$ for all h, n
- Weak concavity: $\tilde{Y}\left[\lambda n + (1-\lambda)n'\right] \geq \lambda \tilde{Y}[n] + (1-\lambda)\tilde{Y}[n']$ for all $0 \leq \lambda \leq 1$.

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 $w(t, s, h) = e^{\gamma s} \tilde{w}(h)[n(t, s, \cdot)]$.

2.2 Workers

There is a continuum of agents that has 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 $\{s(t), h(t)\}_{0 \le t < \infty}$, 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. $s(t) \le t$, and that $s(t)$ be a measurable function in t.[6]

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 $s(t) = \bar{t}$ for all $a < \bar{t} < b$.[7] There is no cost of switching between vintages. I will refer to a *career segment* $l'(t)$ (or short *career*) as the maximal open interval $(l'_0(t), l'_1(t))$ around an instant t that is entirely spent in one vintage.[8] 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

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

[7] 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 and may be relaxed; in equilibrium, workers would not want to return to vintages they have once left.

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

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 and $\dot{h}$ denotes the time derivative of h and c is a function with the following properties:[9]

- Costless demotion: $c[\dot{h}] = 0$ if $\dot{h} \leq 0$.
- Convexity: $c'(\dot{h})$ is a continuous, strictly increasing function on $(0, \infty)$.
- Jumps prohibitively expensive: $\lim_{\Delta t \to \infty} c[\Delta h / \Delta t] \Delta t = \infty$

The last two assumptions imply $\lim_{\dot{h} \to \infty} c'(\dot{h}) = \infty$. 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 $h(t) = 0$.

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 with $h(t) = 0$ for all $t > 0$.

To summarize, the agent's criterion for a given life l_u is

$$v(l_u) = \int_u^\infty e^{-(\beta+\delta)(t-u)} \Big[w\big(t, s_u(t), h_u(t)\big) - e^{\gamma s_u(t)} c\big[\dot{h}_u(t)\big] \Big] dt,$$

where it is understood that $\dot{h} = 0$ on non-segments and $u \geq 0$ indicates the agent's birth date. The value function is defined as $V(t, s, h) = \sup_{l_t(t)=(s,h)} v(l_t)$. 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.

[9] 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 putting hours into the model.

2.3 Stationary equilibrium

I will limit the discussion to densities n which have a collection of neighborhoods in $X \equiv [0,\infty) \times [0,\infty) \times [0,1]$ as their support S_n and which are continuous and differentiable on these neighborhoods.[10]

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 this rate: $V(t,s,h) = e^{\gamma t}\bar{V}(\tau,h)$. 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 t,[11]

$$\int_B e^{-\delta(t-u)} I\left[\left(\tau_u(t), h_u(t)\right) \in B\right] d\mu(l) = \int_B n(\tau,h),$$

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

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

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:

[10] This specification allows for densities that drop precipitously down to zero when a vintage dies — which is exactly what happens in equilibrium. Also, note that feasibility requires that the neighborhoods be connected to points with $h = 0$ or $t = 0$.

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

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) *We have $V(\tau, 0) = W$ 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 $(\tau, h) \in \bar{S_n}$.*[12]

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 $0 \leq \tau \leq T$:

Lemma 2.4. (Finite support of technologies) *There is $T < \infty$ 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 growth of TFP is not.

Proof. Since there exists τ such that $w(\tau, 0) > 0$, 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

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

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. □

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 something 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) *For the end of any career segment l_0, it must be that 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 possible choice 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 ratios since 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 $(\tau, h) \in S_n$.

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 $T^* \geq \tau_1 > \tau'$. □

Lemma 2.7 together with 2.3 shows that the closure of the support of n must be a rectangle $[T_0, T^*] \times [0, 1]$ if the Inada condition 2.1 holds. Section 2.5 shows that there cannot be holes in the τ-direction either when labor inputs are perfect substitutes. Arguments in sections 2.5 and 2.11 will then 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 the partial derivatives V_τ and V_h are denoted by subscripts. The equation says the following: If we know the value function for a given τ for all experience levels h, we can get the value a bit left of this τ by letting the agent choose the optimal slope $\dot{h}$. This slope is contingent on the slope of the value function 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 gets from moving up in the hierarchy (the term inside the max-operator) and another term which is the difference between the current wages and the flow value of $V(\tau,h)$ under the discount factor $\beta+\delta-\gamma$, which is modified for economic growth.

The first-order condition (FOC) corresponding to the HJB (2) is

$$c'\left[\dot{h}(\tau,h)\right] = V_h(\tau,h), \quad (3)$$

where a unique solution for $\dot{h}$ is assured by the assumptions on c whenever $V_h > 0$. Since c is convex, the FOC implies that the greater the value differentials in the hierarchy, the faster is 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 Hamiltonian (the term comprised in 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 usually be non-linear in V_h.[13]

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 a career. Differentiate (2) with respect to h and use the envelope condition (3) to obtain

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

where I have parameterized an agent's career in time by $d\tau = dt$ and $dh = \dot{h}dt$ and the dependence of the various functions on (τ,h) is suppressed for the sake of clarity. We can solve this as an ordinary differential equation

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

in t along an agent's optimal career path to see that the marginal value of human capital equals the discounted integral of marginal wage gains over a career:

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

where T is the end of the career segment and it is clear that $V_h(T, h) = 0$ since $V(T, h) = W$ for all h, i.e. the value of learning 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 and thus human-capital accumulation should be decreasing over segments.

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

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

where the notation $\dot{h}_h = \frac{\partial \dot{h}}{\partial h}$ is used. This PDE says that when one follows the path of an agent, the density thins out at the death rate plus the divergence of the promotion policies $\dot{h}_h$ inside the hierarchy. Section 3.1 contains a derivation of this equation.[14] For a given boundary condition $n(\tau, 0)$ on $\tau \in [T_0, T^*]$, we may solve this PDE through the support of n to obtain the density everywhere.

To summarize, he HJB (2) with its boundary conditions characterizes workers' optimal strategies given wages. Equation (6) tells us how the resulting decisions by workers from (3) translate into a density n once we know how new-borns enter vintages.[15] 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.

2.5 Subcase: Perfect substitutes

We will now 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 given by $w(\tau, h) = e^{-\gamma\tau} f(h)$. We will assume a standard learning curve and require that $f'(h) > 0$ and $f''(h) < 0$.

[14] 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.

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

In this case, 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 the equilibrium density n.[16]

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

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 is strictly dominated by choosing the same career in $s(t_1) = t_1$ and hence not optimal. Obviously, the same holds true for choosing $s(t) > t$ and $h(t) = 0$ on non-segments of positive measure. $\square$

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

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

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 worker with a longer horizon. The following lemma (proven in appendix A.1.3) will be sufficient to prove the formalizations 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:*

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

We will now study comparing optimal careers in worlds with different values for technological growth γ with a view on the experience premium that we would observe. Since equation (7) shows that technological growth is inessential for the intertemporal incentives for human-capital accumulation

[16]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 and the latter in a weak sense.

while inside a vintage, we can use the above lemma to compare optimal careers for different values of γ. As the experience premium, we define the ratio of the wage of an experience-t worker in a vintage to the wage of a career starter, which in a stationary context is $p_\gamma(t) = e^{-\gamma t} w[h(t)]/w[h(0)]$ and where the optimal path h of course depends on γ.

Proposition 2.10. (Shorter horizon lowers experience premium) *Again take $\tilde{Y} = \tilde{Y}$ with $\rho = 1$. Suppose that $\gamma' \neq \gamma$ but fix all other parameters. Then $T^{*'} > T^*$ implies $p_{\gamma'}(t) > p_\gamma(t)$ for any $0 < t \leq T^*$.*

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. □

This result shows that it is impossible that the life span of technologies shortens –which I will refer to as the economy being more turbulent– and simultaneously we see an increase in the experience premium, as has arguably been the case in reality. Figure 1 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.

The following sections will show that this result need not hold when different human-capital levels are complementary in production 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 2.11. (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

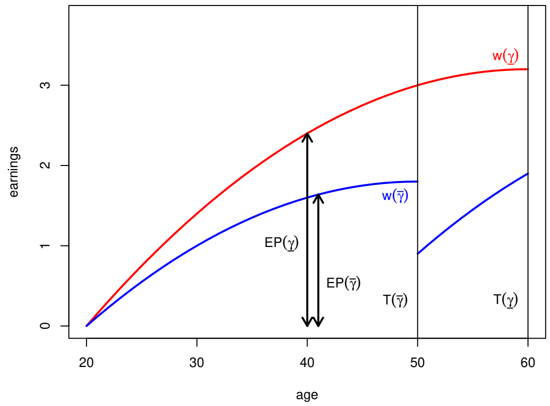

Figure 1: Horizon effect under substitutability

$w(\tau, 0) > w(T^*, 0)$ 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. $\square$

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. When labor inputs are complementary, some workers are also needed in low-skill tasks in the oldest technologies. In order for both entry options being equally attractive, entry wages in young technologies have to be lower there than in old technologies.

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

Lemma 2.12. (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* $\tau \in (0, T^*)$.

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 under our assumptions on $c(\cdot)$. 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 earn outlandish returns by the Inada condition.

Collecting the previous results yields:

Corollary 2.13. (Wage Compression) *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 2.11) and $w(\epsilon, 1) \to \infty$ (see lemma 2.12). The second statement follows from lemma 2.13. $\square$

Intuitively, the wage structure is compressed because experience becomes less scarce over the lifecycle of a vintage; it is easier to acquire skills over a long time than to master a technology that was barely invented. Wage compression under complementarity again opens the possibility that higher turbulence (a shortening of the vintage horizon T^*) can occur alongside an increase in the experience premium — since the wage structure is steeper in young technologies, a shortening of the vintage horizon can send more workers into steep earnings paths, increasing the average experience premium. Section 4 will present calibrations for which this is indeed the case.

Another consequence of the discussion above is the following:

Corollary 2.14. (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 2.12: 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 2.11 and 2.6. $\square$

This result is remarkable since these wage losses occur without assuming that human capital depreciates — an assumption often invoked in Ben-Porath-type models to obtain downward-bending wage profiles for old workers. Agents do not lose any of their skills in the model presented here; the reason for the wage losses is that the relative price for some skills falls over time, a phenomenon which is usually referred to as *obsolescence*.

Also, 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 here (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 output to supply one unit to each surviving member of a cohort born at t and 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

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

where $C(t)$ denotes the aggregate cost of human-capital accumulation at t. First, we will derive an expression for $C(t)$ under 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 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 quantity must decrease at the death rate δ when we evaluate it along an agent's path (always staying in a fixed vintage s). A first-order approximation following a career line $[h(t), \tau(t)]$ yields:

$$N_t(t,s,h) + \dot{h}(t,s,h)N_h(t,s,h) = -\delta N(t,s,h), \tag{8}$$

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 saw for the competitive equilibrium in equation (6).

Re-arranging equation (8) gives us an expression for the career slope $\dot{h}$ the planner should choose given that she wants to implement a given n:

$$\dot{h}(t,s,h) = \frac{N_t(t,s,h) + \delta N(t,s,h)}{n(t,s,h)}. \tag{9}$$

In order to aggregate costs over all agents, we have to weigh the cost of $\dot{h}$ at each point by the mass of agents at the respective point and obtain $C(t) = \int_{s,h} n(t,s,h)c[\dot{h}(t,s,h)]$.[17]

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 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 a variable $u(t,s,h) \equiv n_t(t,s,h)$ into the problem and connect it to the functions n, N and N_t with equality constraints. The Lagrangian is then the following object:[18]

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

[17] This expression is derived in a more formal manner in Kredler (2008).

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

where we scale the Lagrange multipliers by $e^{-(\beta-\gamma)t}$ to make them stationary. The set of constraints linked to the multipliers ν is taken from equation (9). 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 equals the marginal cost of human-capital accumulation and that ν and ξ are closely linked to η:

$$\begin{aligned}\eta(t,s,h) &= e^{-\gamma\tau}c'\big[\dot{h}(t,s,h)\big] \\ \nu(t,s,h) &= e^{-\gamma\tau}c'\big[\dot{h}(t,s,h)\big]n(t,s,h) \\ \xi(t,s,h) &= \delta e^{-\gamma\tau}c'\big[\dot{h}(t,s,h)\big]\end{aligned} \tag{10}$$

Using the above, the FOC with respect to $n(t,s,h)$ is

$$\begin{aligned}\lambda(t,s,h) =& w(t,s,h) - e^{-\gamma\tau}c\big[\dot{h}(t,s,h)\big] + e^{-\gamma\tau}\dot{h}(t,s,h)c'\big[\dot{h}(t,s,h)\big] - \\ &- \mu(t) - \delta\int_0^h \eta(t,s,\tilde{h})d\tilde{h}.\end{aligned} \tag{11}$$

where we recognize in the terms involving $c(\cdot)$ the Hamiltonian in the value function (2). The last remaining derivative we have to evaluate is 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):

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

At a stationary solution, we require that the density fulfill $n(t,s,h) = \bar{n}(\tau,h)$, and 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, where again we drop the bar-notation in the following. When plugging the expressions for the Lagrange multipliers (10) and (11) into (12) and imposing stationarity, we obtain

$$\begin{aligned}\int_\tau^T e^{-(\beta-\gamma)(\tilde{\tau}-\tau)}\Big[w(\tilde{\tau},h) - e^{-\gamma\tilde{\tau}}c\big[\dot{h}(\tilde{\tau},h)\big] + \dot{h}(\tilde{\tau},h)c'\big[\dot{h}(\tilde{\tau},h)\big] - \mu - \\ -\delta\int_0^h \eta(\tilde{\tau},\tilde{h})d\tilde{h}\Big]d\tilde{\tau} = \int_0^h e^{-\gamma\tau}c'\big[\dot{h}(\tau,\tilde{h})\big]d\tilde{h} \equiv \Lambda(\tau,h)\end{aligned} \tag{13}$$

We will now see that $\Lambda(\tau,h)$ is an "excess-value function", i.e. 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 untrained agent.

Directly from (13), 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 (13) 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 (13) in both directions to see how this excess-value function evolves on the interior:

$$\Lambda_h(\tau, h) = e^{-\gamma\tau} c'\big[\dot{h}(\tau, h)\big] \tag{14}$$

$$-\Lambda_\tau(\tau, h) = w(\tau, h) - e^{-\gamma\tau} c\big[\dot{h}\big] + e^{-\gamma\tau}\dot{h}c'\big[\dot{h}\big] - \mu - (\beta + \delta - \gamma)\Lambda(\tau, h) \tag{15}$$

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

$$-V_\tau(\tau, h) = w(\tau, h) - e^{-\gamma\tau} c(\dot{h}) + \dot{h}V_h - (\beta + \delta - \gamma)V(\tau, h) \tag{16}$$

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

3.3 Uniqueness

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

Proposition 3.1. (Solution to planner's problem is unique) *If $Y[n_{t,s}(\cdot)]$ is strictly convex in $n_{t,s}(h)$*[19]*, $J(n)$ is strictly convex in n and there is a unique density $n(t, s, h)$ that maximizes $J(n)$*

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

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 of t and s by the convexity assumption on Y, which implies the desired result. □

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 convex, 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 does not seem to be a problem computationally, but could not be established formally without making an equicontinuity assumption on the function space for n.[20]

3.4 Varying T

So far, we had fixed the maximal vintage age T and imposed this on the planner; we will now be concerned with varying T and finding the optimal T^* when $\tilde{Y}$ is strictly convex. By the fundamental convexity 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 $T^* < \infty$.

The first point to note is that for $T > T^*$, the problem 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 $J[n^*(T^*)] >$

[20] In order to reap the benefits of compactness, we can 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 $n_t \to 0$ everywhere 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.). If we further assume that n is pointwise bounded –which is unproblematic–, equicontinuity allows us to employ the Arzela-Ascoli theorem which says 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.

$J[n^*(T)]$, by convexity 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 may be discontinuous at this point.

The second point to note is that for $T < T^*$, simulations show that usually 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 improvement through human-capital accumulation. It is hard to formally show, though, that this *must* be the case.

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 (16) 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 HJB for the agent (2) and the corresponding optimal policy (3) in competitive equilibrium have their exact counterparts in equations (16) and (15) for the planner's problem. Equation (13) follows by integrating from the boundary over τ and h, which in turn is equivalent to (12). Since the first-order conditions (10) and (11) can be used to define the Lagrange multipliers, equation (12) 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.[21] However, there can be at most one stationary point for a given T since J is a convex 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 $T > T^*$. $\square$

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.[22]

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 and the decomposition of J into

[21] 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 Jacobian being zero in $\mathbb{R}^n$.

[22] In 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.

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 just observing productivity in dying vintages and average productivity in the economy.

4 Calibration

For the calibration, I choose the CES aggregator defined in (1) for the production function and the quadratic specification $c(\dot{h}) = \bar{c}\max\{\dot{h}, 0\}^2/2$ for the cost of human-capital accumulation. 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:[23]

$$w(\tau, h) = e^{-\gamma\tau} f(h) \left(\frac{\tilde{Y}(\tau)}{n(\tau, h)} \right)^{1-\rho} \tag{17}$$

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. Kredler (2008) provides the complete documentation.

The model is calibrated to yearly data. The death rate δ is set to 0.025 to match an expected labor-market participation of 40 years for a 20-year old[24]; β is set to 0.015 to obtain a standard yearly discount rate of $\beta + \delta = 0.04$. A parsimonious functional form is assumed for returns to experience:

[23]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.

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

$f(h) = A + ah$, where $A = 1$ is chosen as a normalization.[25] The remaining parameters $\bar{c} = 5$, $\gamma = 0.008$ and $a = 1$ are chosen to match the experience premium, the dispersion of earnings and the peak of the median age-earnings profile.

4.1 Equilibrium properties

Figure 2 is a summary of the results. The career lines in the upper-left panel show that 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 for old vintages skill scarcity vanishes as more and more workers press up in the skill hierarchy from below.

Figure 3 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.[26] As apparent in figure 2, 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 die at a faster rate than entrants replace them.

The lower-right panel in figure 3 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 like Atkeson & Kehoe (2005). Young vintages are unproductive because they have a 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, joint learning leads to gradual equalization of skill returns. However, these returns wear off over time and the negative TFP effect eventually dominates, as predicted in section 3.6.

[25] To see that this is indeed a normalization, consider a model where the parameters A, a and $\bar{c}$ are each divided by A. Since production is CRS, output and wages are all divided by A for a fixed allocation n. The same is true for the cost of human-capital accumulation. Thus agents will choose the same quantities in both allocations.

[26] A straightforward calculation shows that the entry density is $m(\tau) = n(\tau, 0)/\dot{h}(\tau, 0)$.

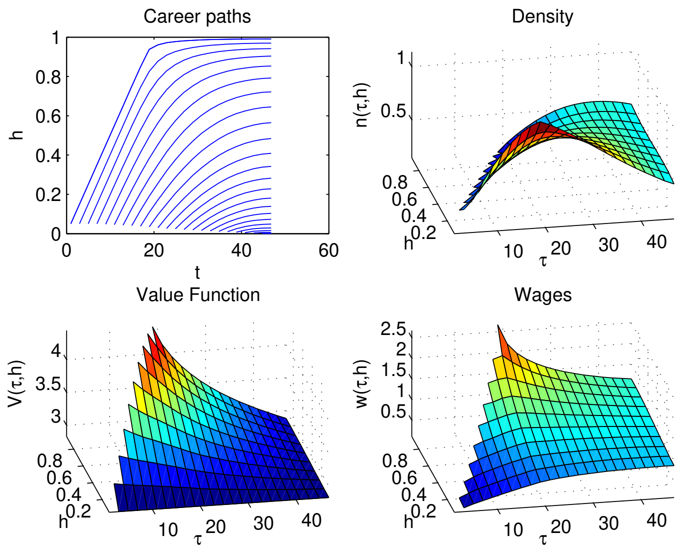

Figure 2: Equilibrium (I)

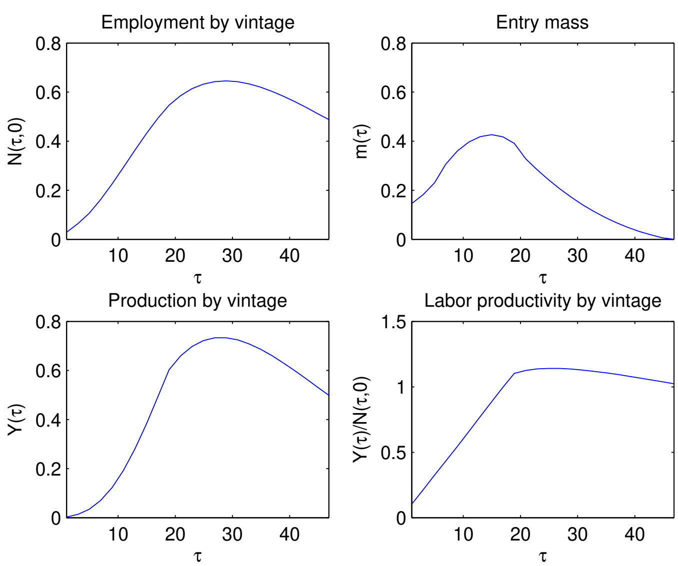

Figure 3: Equilibrium (II)

4.2 Age-earnings profiles

Figure 4 shows age-earnings profiles, following a cohort of labor-market entrants over time as they accumulate human capital in their respective vintage. 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.

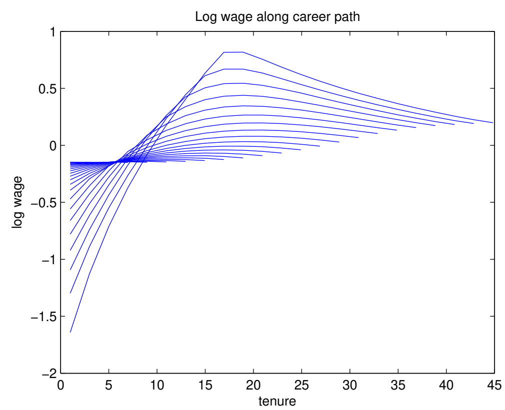

Figure 4: Age-earnings profiles over career

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 (17):

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

The three terms have a clear economic intuition: 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; it is the only effect that is present when skills are perfect substitutes ($\rho = 1$). In the calibration illustrated in figure 2, 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 and $\dot{h}$ decreases over a career, this means that this effect is always positive but wears off over the 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. 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, skills that were once scarce become more abundant as workers from the lower ranks are acquiring higher skills. 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.

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 et al. (2006). In contrast to these models, agents are ex-ante equal here and all heterogeneity is endogenous. In fact, heterogeneity in earnings profiles is essential in order to give workers the incentives to enter *any* career.

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. This effect induces the strong negative correlation between earnings levels and earnings growth that is present in the data for young workers (see Kredler, 2008).

4.3 Varying γ

Figures 5 and 6 show calibrations (blue lines) for 1975-1980 and 1995-2000 and contrast them with the data (grey boxplots). The calibration parameters for 1975 are those described in section 4.1.[27] For the period 1995-2000,

[27] A constant was added to the growth rates in figure 6 to obtain the best possible fit in both the 1975-1980 and the 1995-2000 periods. As argued in section 1, fitting this quantity is beyond the scope of this paper. Also, to obtain predictions for the earnings profile over the entire life cycle, it is necessary to make an assumption as to what happens to workers who are laid off when exiting the last vintage T^*. Since in the model the assignment of laid-off workers to new careers is indeterminate, I assume that they are randomly spread over entry points according to the equilibrium entry density.

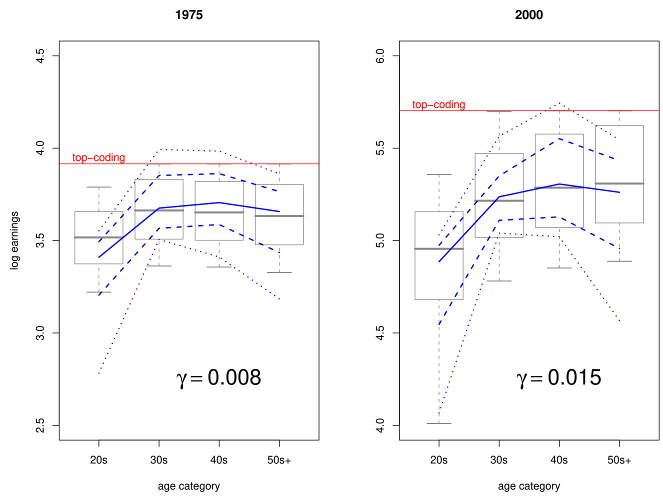

Figure 5: Calibration: Log earnings by age groups

Data: Boxplots in grey. Calibration: blue lines, representing 10-, 25-, 50-, 75- and 90-percentiles.

I increase γ to 0.015 in order to match the observed hike in the experience premium. This decreases the maximal vintage age T^* from 48 to 34 years.

Figure 5 shows that the model does a good job matching the age-earnings profiles in the cross section. The rise in the experience premium occurs due to the mechanism explained in section 2.4. Earnings dispersion increases for the young because a greater fraction of them enters new technologies in which very low entry wages are paid; it rises for the old because more of them face their skills becoming obsolete when vintages are phased out sooner. In figure 6, it is apparent that the model generates an increase in earnings turbulence but that it over-predicts both its level (especially for old workers) and its increase.[28]

[28] The high turbulence for the old is induced by the perpetual-youth assumption. Since a laid-off 50-year-old has the same prospects as a 20-year-old in the model, he is just as likely to enter a new vintage as a young worker. A vintage change entails large earnings losses and large subsequent earnings growth, explaining the patterns in the right panel in figure 6. In reality, of course, due to his shorter planning horizon, an old worker would rather settle for an older technology with a less back-loaded earnings profile. Also, it can be argued that the boxplots in figure 6 understate the true earnings risk faced by German workers since the data does not include workers going into unemployment; this is especially true for old workers for which large flows into unemployment have taken place in Germany in the late 1990s. In this sense, the extreme predictions for old workers'

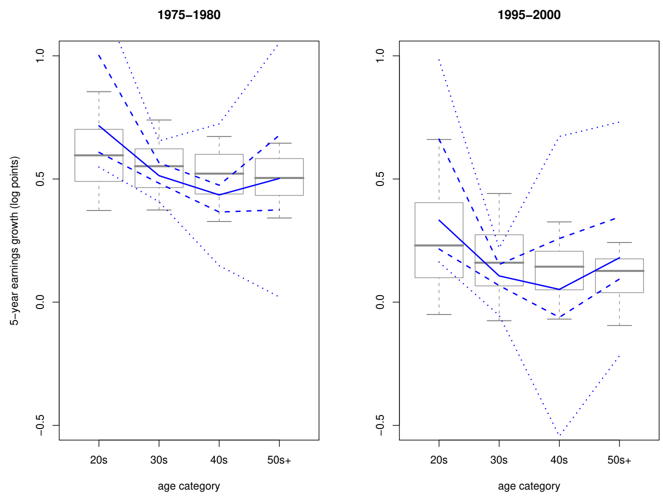

Figure 6: Calibration: Earnings growth by age groups

Data: Boxplots in grey. Calibration: blue lines, representing 10-, 25-, 50-, 75- and 90-percentiles.

5 Conclusions

This paper has argued that accelerated technological change can explain both an increase in the experience premium and a rise in earnings turbulence and dispersion when human capital is technology-specific and different skill levels are complementary in production. In the model, the skill premium is highest in young technologies. Skills are scarcest in these technologies since fast accumulation of human capital is costly. In the wake of a technological acceleration, the life span of technologies shortens and more weight is put on new technologies; the average skill premium rises and so does the experience premium.

A further application of the proposed framework is the analysis of the effect that a varying demographic structure has on age-earnings profiles. The entry of a large cohort into the model economy would lead to a crowding of entry positions and thus a reduction in entry wages. This negative effect would likely persist as the cohort moves through the ranks of the skill hierarchy. The result would be a negative cohort effect as is often estimated in the labor literature.

earnings losses are less problematic.

Another obvious testing ground for the theoretical framework is suggested by its stark predictions on the experience-earnings structure conditional on the age of technology. An encouraging start is that in the German IABS data, the returns to tenure are highest for the youngest and fastest-growing establishments (see Kredler, 2008).[29] However, the model lacks a clear mapping from its vintages to establishments or firms, making it hard to derive precise predictions. An extension of the model in the style that Klette & Kortum (2004) build upon Grossman & Helpman's (1991) growth model would be necessary for a rigorous treatment of this issue.

A further point that has only been touched upon in the previous discussion is the productivity profile of a vintage over time (see figure 3). 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, a stochastic component could be introduced into the framework in order address the question of riskiness of human capital and technology choice for an individual.

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.

Cummins, J. G. & Violante, G. L. (2002), Investment-specific technical change in the US (1947-2000): Measurement and macroeconomic con-

[29] This is in line with Michelacci & Quadrini (2004) who find that returns to tenure are decreasing in a firm's age; 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.

sequences, Finance and Economics Discussion Series 2002-10, Board of Governors of the Federal Reserve System (U.S.).

Gordon, R. J. (1990), *The Measurement of Durable Goods Prices*, University of Chicago Press, Chicago.

Gottschalk, P. & Joyce, M. (1998), 'Cross-national differences in the rise in earnings inequality: Market and institutional factors', *Review of Economics and Statistics* **80**(4), 489–502.

Gottschalk, P. & Smeeding, T. M. (1997), 'Cross-national comparisons of earnings and income inequality', *Journal of Economic Literature* **35**(2), 633–687.

Grossman, G. M. & Helpman, E. (1991), 'Quality ladders in the theory of growth', *Review of Economic Studies* **58**(1), 43–61.

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.

Heckman, J., Lochner, L. & Todd, P. (2003), 'Fifty years of Mincer earnings regressions', *NBER Working Paper* .

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

Klette, T. J. & Kortum, S. (2004), 'Innovating firms and aggregate innovation', *Journal of Political Economy* **112**(5), 986–1018.

Kredler, M. (2008), 'Changes in German age-earnings profiles 1975-2001: An explanation(?)', *Working Paper* .

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.

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. Since $\tilde{Y}$ is weakly concave and takes finite values on the unit simplex $\Delta = \{n : \int n = 1\}$, it is impossible that there is a sequence n_k such that $\tilde{Y}(n_k)$ grows unbounded. To see this, take an interior point $n \in \Delta$ and observe that $\tilde{Y}(n)$ eventually lies below the ray between $(n_k, \tilde{Y}(n_k))$ and some point in a ball around n. □

A.1.2 Proof of lemma 2.3

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. □

A.1.3 Proof of lemma 2.9

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

$$h(t) - h(s) \geq g(t) - g(s) \Rightarrow \int_s^t \dot{h}(u)du \geq \int_s^t \dot{g}(u)du, \tag{18}$$

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

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

since by assumption $\dot{h}(t) < \dot{g}(t)$ and $w_h[h(v)] \leq w_h[g(v)]$ point-wise; the inequality follows from c' being increasing. This again implies $\dot{h}(u) < \dot{g}(u)$ for all u, which in turn contradicts (18). □

A.1.4 Proof of lemma 2.12

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 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 $W < \infty$. □

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 get grids $\{t_i^{(k)}\}_{i=1}^{\infty}$ and $\{s_i^{(k)}\}_{i=1}^{N_s}$. For human capital, cut 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 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 combinations 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, any combination must 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 for $s = 0$ 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 by a finite number of improving operations into the proposed no-crossing algorithm; 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, we obtain $C(\mu') = \lim_{k\to\infty} C_k(\mu') \geq \lim_{k\to\infty} C_k^* = C^*$.

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 had $N(0, 1, h)$ workers above himself in the career and survives until $t = 1$ will have $\exp(-1\delta)N(0, 1, h)$ workers above himself at $t = 1$, which is the second claim of the statement. The arguments above obviously generalize to any t, s and $t + u$. $\square$