

Working to Learn*

Drew Fudenberg[†] George Georgiadis[‡] Luis Rayo[§]

July 28, 2020

Abstract

We study the joint determination of wages, effort, and training in “apprenticeships” where novices must work in order to learn. We introduce the idea of learning-by-doing as an inequality constraint, which allows masters to strategically slow training down. Every Pareto-efficient contract has an initial phase where the novice learns as fast as technologically feasible, followed by a phase where their master constrains how fast they learn. This latter phase mitigates the novice’s commitment problem, and thus lets the novice consume more than they produce early on in the relationship. Our model also has novel implications for optimal regulation.

*We thank Robert Akerlof, Daniel Barron, Gonzalo Cisternas, Michael Powell, Ori Shelef, and conference participants at CMID20 for helpful conversations, the National Science Foundation grant SES 1643517 for financial support, and Miguel Talamas for valuable research assistance.

[†]Department of Economics, MIT

[‡]Kellogg School of Management, Northwestern University

[§]Kellogg School of Management, Northwestern University

1 Introduction

Careers in a wide range of industries, such as medicine, academia, professional services, culinary arts, investment banking, and the traditional trades, frequently begin with a lengthy “apprenticeship” stage where novices gain knowledge from their masters while working hard and receiving relatively low wages. We propose that these apprenticeships are shaped by cognitive constraints that bound the speed at which novices can learn, combined with the novices’ desire to smooth their consumption, their initial lack of money, and their inability to commit not to leave once they are trained.

In our model, a principal offers an agent an apprenticeship consisting of time paths of knowledge transfer, wages, and effort, subject to the constraint that the agent can walk away at any time, and subject to a learning constraint that bounds how quickly they can learn. Previous work on learning-by-doing has followed Arrow (1962) in modeling the learning constraint as an equality: workers or firms learn as quickly as their effort or production level allows. In contrast, we model it as an inequality constraint, to allow the masters to hold back knowledge even when their novices are working hard.¹

To gain better insight into the forces that shape the apprenticeship, our analysis solves for the whole family of incentive-compatible, Pareto-optimal contracts given the agent’s initial knowledge. This also facilitates comparative statics based on the agent’s bargaining power, as measured by their outside option at the time the principal offers to hire them.

We show that every Pareto-optimal contract has two phases. In the first one, the agent learns as fast as their learning-by-doing constraint allows given their effort level, while earning rents in the sense that they are more than compensated for the economic cost of working for the principal. Then in the

¹Masters may have an incentive to hold back knowledge to prolong the apprenticeship while extracting rents from their novices (e.g. Smith, 1776, Chapter 10). Moreover, it seems plausible that without active participation by the master, a novice’s ability to learn will be limited.

second phase, the principal only allows the agent to learn as quickly as is consistent with the agent being willing to remain in the apprenticeship; here the principal keeps all rents. This solution follows from a natural assumption on the value of the knowledge and the agent's learning ability, namely that the agent is willing to work for the minimum subsistence wage if they are taught as quickly as their effort allows them to learn.

The nature and length of these phases vary significantly with the agent's outside option. When this outside option is low, phase 1 is relatively short and prescribes low wages, while phase 2, which is relatively long, offers an increasing wage path that converges to the agent's steady state (post graduation) earnings. As the agent's outside option improves, phase 1 grows and prescribes higher wages, while phase 2 becomes shorter, but never disappears completely. This is because when the agent's outside option is high, phase 1 pays them more than they produce, thus placing them in "debt." Phase 2 then allows the principal to gradually collect on this debt, despite the agent's lack of commitment power, through the promise of just enough additional training to prevent the agent from walking away. The agent prefers this apprenticeship over a shorter one with only phase 1 because it allows for better consumption smoothing.²

Throughout the apprenticeship, effort is distorted above the static first best (i.e., the first-best effort when there is no learning). One reason is that higher effort allows the agent to learn faster; this force is only relevant in the first phase. A second reason is that increased effort transfers rents to the principal; this force is present in both phases, but gradually vanishes as the apprenticeship nears its end.

We discuss some salient features of real-life apprenticeships in Section 2. These features will motivate our analysis and, in particular, will highlight the need for a model of dynamic knowledge transfers that is richer than existing

²If the agent had no reason to smooth consumption, they would prefer a shorter apprenticeship in which they are always trained at the maximum rate.

ones.

Related work. Our work builds on Garicano and Rayo (2017) and Fudenberg and Rayo (2019) (henceforth GR and FR), where players exchange work for knowledge. What distinguishes our model is the agent’s learning-by-doing constraint, which is absent in those papers. This constraint lowers the apprentice’s productivity at the start of their career, which given the agent’s desire for early consumption, means that the principal may incur initial losses that need to be recouped. Because the agent lacks commitment power, the principal recoups these early losses by intentionally slowing down training later on. Thus, unlike in GR and FR, the model predicts lengthy apprenticeships with strategically slow training even when the agent has all of the bargaining power. This richer model better fits the real-life practices noted in Section 2, and also offers new insights for optimal regulation.

The many papers on human capital accumulation (e.g. Ben-Porath, 1967, Rosen, 1972, Weiss, 1972) summarized and synthesized in Killingsworth (1982) all assume that the agent chooses the time paths of effort, wages, and learning to maximize their utility given some technological constraints. These models cannot explain inefficiently long training periods, and imply that regulation of wage or effort paths can only lower welfare.³

Thomas and Worrall (1994) and Albuquerque and Hopenhayn (2004), like us, study a contracting problem where the agent’s outside option and productivity increase gradually over time. Both the assumptions and conclusions of these papers are quite different: In these earlier papers payment can only be enforced when the principal is able to directly punish the agent,⁴ there is no excess effort, and there is no reason for consumption smoothing, so the agent gets no “wages” until the steady-state is reached. Moreover,

³In Ben Porath (1962) and Killingsworth (1982) the agent’s earnings equal their productivity; in our model, in contrast, earnings can be both above and below that level, depending on the stage of the apprenticeship and the agent’s bargaining power.

⁴In our setting, one such punishment would be if the principal could forbid the agent from using the production technology. In this case the principal would never want to strategically slow the agent’s training.

there is no analog to our learning constraint, and hence the solution involves a single type of regime.

Kolb and Madsen (2020) consider the design of careers in environments where the agent might be a saboteur. As in our model, the agent goes through various stages while the stakes of the relationship gradually increase. Unlike in our model, though, these stages serve as a dynamic screening device meant to weed out disloyal agents.

There is a large literature where transfers of general human capital are only possible because of market frictions (e.g. Katz and Ziderman, 1990, Acemoglu, 1997, Acemoglu and Pischke, 1998, and Malcomson, Maw, and McCormick, 2003); these papers all assume that training is instantaneous. There is also a large literature on effort distortions that arise when the agent’s productivity is unknown (e.g. Akerlof, 1976, Landers, Rebitzer, and Taylor, 1996, Holmström, 1999, Dewatripont, Jewitt, and Tirole, 1999, Board and Meyer-ter-Vehn, 2013, Barlevy and Neal, 2019, Bonatti and Hörner, 2017, and Cisternas, 2018); in our model such distortions arise because the agent is liquidity constrained.

2 Empirical motivation

Real-life training relationships, including formal apprenticeships, tend to be rather complex. While these relationships are, in essence, a work-for-knowledge exchange, they frequently consist of a bundle of interrelated practices. These include:

1. *Distinct phases.* Apprenticeships are often criticized for taking excessively long, with masters strategically slowing down training while profiting from their novices (e.g. Smith, 1776).⁵ Yet, a novice’s training need not be uniformly slow. Ph.D. programs, for instance, frequently begin with an in-

⁵Indeed, across time and throughout a wide range of industries, novices find themselves performing menial tasks for their masters—rather than being fully devoted to acquiring more advanced skills. GR and FR provide a variety of examples.

tense pre-candidacy instruction phase where students spend nearly all their time learning (and receive abundant input from faculty). Then students enter a post-candidacy phase with more work. Here they might be assigned tasks that benefit their university but can easily distract them from learning, such as grading or performing menial laboratory work.⁶ Formal apprenticeships (e.g. in skilled trades) also frequently include both classroom phases with focused, practical learning—for instance, in a boot camp or at an affiliated college—and phases where novices work as they learn (e.g. Stockman, 2019). These work phases allow novices to master new skills, but they may also entail a degree of grunt work that slows down training.

2. *Financial support.* From the standpoint of novices, one of the most attractive features of being an apprentice is the ability to receive income or other forms of financial support while they learn. Ph.D. students, for instance, may receive a stipend sufficient for living even before they do any work. Similarly, “German apprenticeships generally offer a living wage for two or three years while students learn and work alongside experienced employees” (Hackman, 2018, WSJ); and in the U.S., according to the Department of Labor, “From their first day of work, apprentices receive a paycheck” averaging \$15/hour to begin.⁷ Early in the apprenticeship, when the novice’s productivity is likely to be low, such stipends and guaranteed wages are a potential source of losses for the master—especially if the novice is primarily devoted to learning.⁸

⁶Ph.D. programs have a shared history with other forms of apprenticeship. As noted by Adam Smith, “to have studied seven years under a master properly qualified was necessary...to become a master, teacher, or doctor (words anciently synonymous) in the liberal arts, and to have scholars or apprentices (words likewise originally synonymous) to study under him” (Smith, 1776, Ch. 10).

⁷See www.dol.gov/apprenticeship/toolkit/toolkitfaq.htm (accessed 7/7/20).

⁸Employers may also face a variety of additional costs. Overall costs for German company’s “range from \$25,000 per apprentice to more than \$80,000,” and might be even higher in the U.S. “where firms will have to build programs from scratch, pay school tuition...and in many cases funnel money into local high schools and community colleges to transform them into effective training partners....the Siemens USA plant in Charlotte...reportedly spends some \$170,000 per apprentice” (Jacoby, 2014).

3. *Growing wages.* As novices gain skills their earnings typically grow. They presumably prefer that wages increase holding the starting wage fixed, but many of them might prefer flatter wages with the same present value, which suggests that there are some incentive-based roadblocks to consumption smoothing. In formal U.S. apprenticeships, the paycheck of novices is “guaranteed to increase as their training progresses,”⁹ and employers in Germany “must grant apprentices reasonable remuneration [and] remuneration increases with progressive vocational training, at least annually.”¹⁰ In some cases, these gains are pronounced. For instance, many craft apprentices in Ireland can expect their earnings to more than double throughout their approximately 4-year apprenticeships.¹¹

4. *High effort.* Many novices, whether in traditional apprenticeships or in the early stages of high-skilled careers, encounter heavy workloads with long hours.¹² While hard work may accelerate learning, it may also serve as a way to extract rents from the novice.

The optimal design of apprenticeships has been the subject of much debate and, in many cases, heavy regulation. In countries like Germany and Switzerland, where apprenticeships have a long and successful history, training programs are jointly designed by companies, trade associations, and state and federal authorities (Wyman, 2017). In the U.S., in contrast, where apprenticeships have recently attracted considerable interest, the government is seeking to expand participation by taking a less active role and instead letting trade associations, nonprofits, schools, and labor unions set standards themselves (Morath, 2019). Yet, critics fear that this approach will leave novices unprotected.

⁹See www.dol.gov/apprenticeship/toolkit/toolkitfaq.htm (accessed 7/7/20).

¹⁰See Section 17 of the Vocational Training Act, www.gesetze-im-internet.de/bbig_2005/ (accessed 7/12/20).

¹¹See www.apprenticeship.ie/en/apprentice/craft/Pages/ApprenticeInfo.aspx (accessed 7/8/20).

¹²See FR for a collection of examples ranging from cooks to doctors to industrial-revolution-era apprentices.

This debate highlights the need for a better theoretical understanding of the problem. Our model, while stylized, will help explain why the above practices may arise, and clarify their impact on the well-being of novices, in a way that past work does not.¹³ Our model is also able to offer new insights into optimal regulation.

3 Model

Technology and physical constraints. A principal wishes to employ and train an agent (each of whom will be referred to as “they”). Both players are infinitely-lived and discount the future at rate $r > 0$. Time is continuous. The agent is endowed with knowledge $\underline{X} \geq 0$ and the principal is endowed with knowledge $\bar{X} > \underline{X}$. The only way the agent can raise their knowledge level is by means of knowledge transfers from the principal, and the agent’s knowledge can never decrease.

At time t the agent possesses knowledge $X_t \geq \underline{X}$ and can use this knowledge whether or not they work for the principal. If they work for the principal, they exert flow effort $a_t \in [0, \bar{a}]$ at cost $d(a)$ and produce flow output $f(X_t) + a_t$. If instead the agent works on their own, they can either use this same production technology and keep all output (but not gain additional knowledge), or select an altogether different occupation that does not make use of X_t . The agent also has access to a bank account that pays interest r , but is liquidity constrained: They have a zero balance at time 0 and can never hold a negative balance. The agent’s flow consumption level is c_t , which we assume cannot fall below a minimum subsistence level $\underline{c} \geq 0$, and the agent’s flow utility is $u(c_t) - d(a_t)$. Variables in bold (such as **a** and **c**) will denote time paths.

¹³GR and FR predict that throughout the apprenticeship training is inefficiently slow and wages are lower than output. Moreover, unlike the present model, they suggest that apprenticeships will be short when there is competition between masters, as there is no need to slow down training to smooth the novice’s consumption.

Assumption 1. f , u , and d are twice continuously differentiable with bounded first and second derivatives, and satisfy:

1. $f'(X) > 0$ and $f''(X) < 0$ for all $X \geq \underline{X}$.
2. $u'(c) > 0$ and $u''(c) < 0$ for all $c \geq \underline{c}$.
3. $d'(0) = 0$ and $d''(a) > 0$ for all $a \geq 0$.

For any given knowledge level X , let $a^*(X)$ denote the (unique) solution to $\max_{a \in [0, \bar{a}]} [u(f(X) + a) - d(a)]$, which represents the agent's myopically optimal effort level when consuming all output, and denote the agent's corresponding utility by $\eta(X) := u(f(X) + a^*(X)) - d(a^*(X))$. We assume $a^*(\bar{X}) < \bar{a}$ and normalize $u'(f(\bar{X}) + a^*(\bar{X})) = 1$.

The agent is unable to learn except when working for the principal. The speed $z_t := \dot{X}_t$ at which the principal can train the agent at time t is bounded by the *learning constraint*

$$z_t \leq L(X_t, a_t), \tag{1}$$

as well as the constraints $0 \leq z_t$ and $X_t \leq \bar{X}$.¹⁴ Note that here we frame learning by doing as an inequality constraint as opposed to the equality used in Arrow (1962) and subsequent work. This is because we assume that even when the agent works they will not learn without guidance from the principal, who may have a strategic reason to slow the agent's learning.¹⁵

Assumption 2. $L(X, a)$ is additively separable, strictly positive, weakly increasing, weakly concave, and twice differentiable with bounded first and second derivatives.

¹⁴The learning constraint could reflect a bound on the agent's learning ability, the principal's teaching ability, or both.

¹⁵As far as we are aware in all past work on learning by doing there is no reason for learning to take place inefficiently slowly, so it is assumed that learning takes place as quickly as possible given other variables.

This assumption implies that the agent can be fully trained in finite time. We use additive separability to give a simple sufficient condition for the uniqueness of the optimal contract.¹⁶

Assumption 3. For all $X \in [\underline{X}, \bar{X}]$ and all $a \in [0, \bar{a}]$,

$$\underbrace{L(X, a) \eta'(X) / r}_{\text{value of maximum knowledge gain}} > \underbrace{\eta(X) - [u(\underline{c}) - d(a)]}_{\text{economic cost}}.$$

The left-hand side of the inequality is the agent's instantaneous gain from being trained at the maximum rate, and the right-hand side is the opportunity cost of working for the principal when earning the minimum subsistence wage. Assumption 3 therefore says that if over a small period of time the principal teaches the agent as much as they can possibly absorb, while paying them only the minimum subsistence wage, then the agent earns rents. This assumption greatly simplifies the structure of the optimal contract.

Apprenticeship contracts. The principal employs (and trains) the agent between time 0 and a terminal time $T \leq \bar{T}$, where \bar{T} is an exogenous upper bound of say 200 years.¹⁷ After time T , the agent works on their own, using knowledge X_T . At the start of the relationship, the principal commits to a contract $\mathcal{C} := \langle T, (z_t, w_t, a_t)_{t=0}^T \rangle$, which consists of a terminal time T and time paths of knowledge transfers z_t , wages w_t , and effort a_t .

We assume that throughout the duration of the contract the principal controls the agent's savings, so consumption c_t equals wages w_t .¹⁸ We adopt the convention that T is the earliest time t such that $X_t = X_T$; that is, the agent "graduates" as soon as the knowledge transfer has ended. From time T onward the agent enjoys flow payoff $\eta(X_T)$.

The principal's and agent's continuation payoffs from any time t onward,

¹⁶Alternate sufficient conditions involve restrictions on the third partials of L .

¹⁷Provided \bar{T} is sufficiently large the constraint $T \leq \bar{T}$ will not bind.

¹⁸As we shall see, the principal will specify a non-decreasing wage path, and therefore this assumption is without loss.

provided the agent remains with the principal until time T , are given by:

$$\begin{aligned}\Pi_t &= \int_t^T e^{-r(\tau-t)} [f(X_\tau) + a_\tau - w_\tau] d\tau, \\ V_t &= \int_t^T e^{-r(\tau-t)} [u(w_\tau) - d(a_\tau)] d\tau + e^{-r(T-t)} \eta(X_T) / r.\end{aligned}$$

Let $v_t := rV_t$ denote the agent's continuation value measured in flow terms.

The agent can walk away from the principal at any time before T and receive flow payoff $\eta(X_t)$ in perpetuity. The agent can also reject the principal's contract altogether and obtain flow utility \underline{v} from the alternative occupation in perpetuity. Consequently, the principal is bound by the *participation constraints*

$$v_t \geq \eta(X_t) \text{ for all } t \leq T, \quad (2)$$

$$v_0 \geq \underline{v}. \quad (3)$$

We call the first constraint the *ongoing participation constraint* and the second one the *initial participation constraint*.¹⁹ Notice that absent the learning constraint, the principal would wish to raise knowledge at each time to the level where the ongoing participation constraint binds, as this would maximize the agent's productivity.

4 Benchmark: The first-best

Here we consider a simple benchmark scenario where two of the central frictions are removed. First, we allow the agent to learn on their own, without any assistance from the principal, subject only to the learning constraint (1). Second, we allow the agent to commit to any output and wage paths they desire, which means the agent can use the bank to borrow and save. In this

¹⁹In principle the agent could also move to the alternative occupation at $t > 0$, but as long as the initial participation constraint is met they will never choose to do so.

benchmark, which we call the “first-best,” the principal plays no role.

Formally, we solve

$$\max_{z, w, a} \int_0^{\infty} e^{-rt} [u(w_t) - d(a_t)] dt \quad (I)$$

s.t. (1)

$$z_t \geq 0, X_0 = \underline{X}, X_t \leq \bar{X}$$

$$a_t \in [0, \bar{a}]$$

$$\int_0^{\infty} e^{-rt} w_t dt \leq \int_0^{\infty} e^{-rt} [f(X_t) + a_t] dt. \quad (4)$$

The objective here is the agent’s payoff, and both the ongoing and initial participation constraints (2) and (3) are omitted. Constraint (4) indicates that the present value of wages cannot exceed the present value of output, as required by the agent’s bank.²⁰

Theorem 1. *The first-best contract is unique. In this contract, at every $t \in [0, T)$, the learning constraint binds, the agent earns a constant wage w^{**} , and there is a strictly positive function D_t such that their effort path satisfies*

$$d'(a_t) = \min \{d'(\bar{a}), u'(w^{**})(1 + D_t)\}. \quad (5)$$

*At time T the agent graduates with knowledge \bar{X} , and from that time onward consumes w^{**} and exerts the constant effort a^{**} given by*

$$d'(a^{**}) = \min \{d'(\bar{a}), u'(w^{**})\}. \quad (6)$$

Proof. See the Online Appendix. □

The first thing to note is that the agent uses their commitment power

²⁰We have also omitted the constraint $w_t \geq \underline{c}$ as we shall assume that the agent is sufficiently productive as to secure at least this level of consumption.

to fully smooth their consumption across time. Moreover, because faster learning leads to higher output and greater consumption (by relaxing the financial constraint (4)), the agent raises their knowledge as quickly as the learning constraint allows, until fully trained.

What remains is to characterize the optimal effort path. If effort did not impact learning, the agent’s ideal effort would equate the marginal cost of effort with the marginal utility of consumption. When the agent is still learning, they distort effort upward in proportion to the term D_t in equation (5). This term is given by

$$D_t = L_a(X_t, a_t) \int_t^T e^{-r(s-t)} f'(X_s) e^{\int_t^s L_X(X_\tau, a_\tau) d\tau} ds. \quad (7)$$

In this equation, L_a captures the fact that greater effort leads to faster learning. The integral captures the fact that faster learning today raises output tomorrow—which happens both directly (per the term $f'(X_s)$) and also via the compounding impact of knowledge on future learning (per the second exponential inside the integral).²¹

As we shall see next, once we reintroduce the original frictions, the optimal contract will preserve some but not all of these features.

5 Main result

Here we return to the optimal contracting problem where all constraints are present. As we show in Theorem 2, the agent’s lack of commitment will cause a variety of distortions relative to the first best. These distortions, moreover, will be magnified by the principal’s desire to extract rents.

²¹After time T the agent is done learning and faces a static problem, so the effort equation simplifies to (6).

The principal's problem is:

$$\begin{aligned} \max_{\underline{c}} \quad & \int_0^T e^{-rt} [f(X_t) + a_t - w_t] dt & \text{(II)} \\ \text{s.t.} \quad & (1), (2), (3) \\ & z_t \geq 0, X_0 = \underline{X}, X_t \leq \bar{X} \\ & a_t \in [0, \bar{a}], w_t \geq \underline{c}, T \leq \bar{T}. \end{aligned}$$

Note that varying the agent's outside payoff \underline{v} traces the Pareto-efficient implementable payoffs via the initial participation constraint (3) (i.e. $v_0 \geq \underline{v}$).²²

We solve this model under two further assumptions. The first is that when the agent learns as fast as the learning constraint allows given their effort, the present value of output they produce is less than the present value of the output they would have produced in the same period of time had they been working with knowledge \bar{X} and exerting effort $a^*(\bar{X})$.

Assumption 4. *For every time path of effort \mathbf{a} ,*

$$\int_0^T e^{-rt} [f(\hat{X}_t(\mathbf{a})) + a_t] dt < \int_0^T e^{-rt} [f(\bar{X}) + a^*(\bar{X})] dt,$$

where $\hat{X}_t(\mathbf{a})$ is the knowledge path when the agent learns at rate $L(X_t, a_t)$ and exerts effort \mathbf{a} until fully trained.

This assumption is met if knowledge is sufficiently valuable relative to effort, in the sense that high effort cannot make up for low knowledge.²³

Our next assumption imposes some parametric restrictions.

²²A key challenge from a technical standpoint is that this problem is linear in z_t , so the optimal trajectory of z_t cannot be determined using a first-order approach. Instead, we must conjecture a solution and verify that it is optimal.

²³This assumption implies, in particular, that any arrangement where the agent earns steady-state wages $f(\bar{X}) + a^*(\bar{X})$ from the beginning would cause the principal losses.

Assumption 5. *The parameters of the model are such that:*

1. $\underline{v} > \eta(\underline{X})$.
2. *There is a feasible contract with positive training (i.e., $X_T > \underline{X}$) where the principal makes a non-negative profit.*
3. *When the principal is indifferent they choose to train the agent.*

Part 1 of this assumption says that the agent's initial knowledge is sufficiently low that without training they would be more productive in their alternative occupation.²⁴ Part 2 holds whenever \underline{v} and \underline{c} are sufficiently small.

We begin by introducing some notation. Given a fixed time path of knowledge, let

$$m_t := \frac{f'(X_t)}{\eta'(X_t)/r} \text{ and } S_t := -1 + \int_t^T m_s ds.$$

The ratio m_t , which can also be written as $r/u'(f(X_t) + a^*(X_t))$, measures the marginal impact of knowledge on output relative to its impact on the agent's outside option. The term S_t represents the slope of the player's payoff frontier in a world with no learning constraints so the agent's ongoing participation constraint always binds, and where the agent is fully trained by time T .

To understand this payoff frontier, observe that, from the standpoint of time t , gifting the agent 1 extra util upon graduation costs the principal 1 (since at that time the agent's marginal utility becomes 1), but it also increases the agent's continuation payoff by 1 at all times before T , which allows the principal to raise the agent's productivity by m_s at all times s before T , so the principal recoups $\int_t^T m_s ds$. Notice that as we move backward in time, S_t falls in absolute terms, i.e., the payoff frontier becomes flatter. This is because the greater the time that remains in the apprenticeship, the larger the loss of output due to the agent's low level of knowledge.

²⁴This assumption guarantees that the learning constraint binds for some interval of time.

Theorem 2 shows that, regardless of the agent’s bargaining power (as represented by their initial outside option v) the optimal contract consists of two phases. The first one resembles the first-best contract in that the agent is trained as fast as the learning constraint allows, exerts high effort, and earns flat but less than first-best wages. This phase ends before the agent is fully trained. In the second phase, training becomes artificially slow (in particular, the agent is trained just quickly enough to meet their ongoing participation constraint), effort remains distorted above the static first best, wages grow, and the principal keeps all rents for themselves. The agent graduates at the completion of this phase.

Theorem 2. *The unique profit-maximizing contract consists of two learning phases, separated by a time $\theta \in (0, T)$:*

1. *Phase 1 (“technologically-restricted learning”). In this phase the learning constraint binds. Moreover, the agent receives the constant wage w^1 given by*

$$u'(w^1) = \min \{u'(\underline{c}), 1/|S_\theta|\},$$

and is assigned the effort path

$$d'(a_t) = \min \{d'(\bar{a}), (1 + D_t)/|S_\theta|\},$$

*where $D_\theta = 0$ and $D_t > 0$ for all $t < \theta$.*²⁵

2. *Phase 2 (“principal-restricted learning”). In this phase the learning constraint is slack. Moreover, the agent earns zero rents, is offered the non-decreasing wage path*

$$u'(w_t) = \min \{u'(\underline{c}), 1/|S_t|\},$$

²⁵Specifically, $D_t = L_a(X_t, a_t) \int_t^\theta e^{-r(s-t)} f'(X_s) e^{\int_t^s L_X(X_\tau, a_\tau) d\tau} ds$, which is analogous to (7).

and is assigned the non-increasing effort path

$$d'(a_t) = \min \{d'(\bar{a}), 1/|S_t|\}.$$

At time T the agent graduates with knowledge \bar{X} , and from that time onward exerts first-best effort $a^*(\bar{X})$ and consumes the corresponding output.

Proof. See Appendix A. □

Figure 1 helps us understand this result. It depicts two different Pareto frontiers. The higher one, which we call the *unconstrained frontier*, represents the world with no learning constraints (i.e., where knowledge can be instantly raised). At any point along this frontier, the agent's ongoing participation constraint binds. The lower frontier, which we call the *constrained frontier*, is the one encountered by the players when the learning constraint is present. This frontier lies below the unconstrained one because at time 0 knowledge cannot be instantly raised to the point where the agent's ongoing participation constraint binds.²⁶

The contract begins at a point along the constrained frontier—either at the peak of the frontier if \underline{v}/r is to the left of this peak, or at the point where the agent receives exactly \underline{v}/r . (In the figure, points **a**, **b**, and **c** indicate three possible starting points.) Because the unconstrained frontier lies above the constrained one, the principal wishes to reach this higher frontier without delay, and hence trains the agent as fast as they can learn. Along the way the principal allows the agent to earn rents because the principal's priority is to boost the agent's productivity. Once the higher frontier is reached, which occurs at time θ , the contract enters the next phase where the principal extracts all additional rents. As this phase moves forward, the players' continuation payoffs gradually move along the higher frontier until the agent's

²⁶Recall that $\underline{v} > \eta(\underline{X})$ and the initial participation constraint requires that $v_0 \geq \underline{v}$. Therefore, the fact that knowledge cannot be instantly raised implies that $v_0 > \eta(X_0)$.

training is complete. At each instant, the slope of this frontier represents the (shadow) cost of paying the agent with knowledge (i.e. the loss in profits from raising the agent’s outside option). Modulo the exogenous bounds on wages and effort, the optimal contract equates both the marginal utility of wages and the marginal cost of effort to this shadow cost.

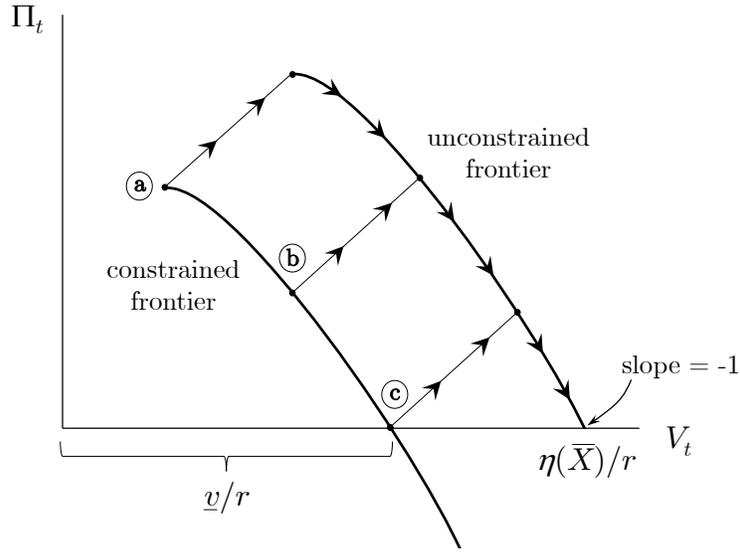


Figure 1: Pareto frontiers and evolution of payoffs

The path traced by the players’ continuation payoffs connects the two frontiers at points of equal slope. This is a form of smooth pasting between the two phases that guarantees that wages and effort do not jump at time θ . To understand this smooth pasting, notice that at time θ the learning constraint no longer binds, so the principal can increase the agent’s continuation payoff at θ by a small amount dV by accelerating knowledge transfer from θ to $\theta + dt$, while leaving all other aspects of the contract unchanged. This slightly shortens phase 2, raises the agent’s overall payoff V_0 by $e^{-r\theta}dV$ and changes the principal’s overall payoff Π_0 by $e^{-r\theta}S_\theta dV$, so $d\Pi_0/dV_0 = S_\theta$.

As the agent’s initial outside option increases, the contract begins farther

and farther to the right along the constrained frontier. This means that more training occurs in phase 1 and less in phase 2. It also means higher wages and lower effort distortions, and thus more rents for the agent. Notably, phase 2 has positive length even in the most preferred contract for the agent (i.e., the zero-profit contract starting at point \mathbf{c}). This is because the principal can use phase 2, where they extract the most rent, to collect on “debt” incurred by the agent during phase 1. As a result, early in the apprenticeship, when the agent’s productivity is still low, they can consume more than they produce and hence better smooth their consumption. The principal collects on this debt once the agent is more productive by holding on to them for an artificially long time, paying them less than they produce, and promising a training rate just high enough to keep them from leaving. In this way, phase 2 allows the players to work around the agent’s commitment problem. Because a two-phase arrangement allows for better consumption smoothing, the agent prefers it over a shorter apprenticeship with faster training and only a phase 1.²⁷

Throughout phase 2, the agent’s ongoing participation constraint binds and they are trained at the *zero-rent rate*²⁸

$$z_t = \frac{\eta(X_t) - [u(w_t) - d(a_t)]}{\eta'(X_t)/r}. \quad (8)$$

The numerator is the (instantaneous) utility loss incurred by the agent when working for the principal rather than on their own and consuming all output; the denominator is the agent’s utility gain per unit of knowledge they acquire from the principal. The zero-rent rate therefore guarantees that the agent is indifferent between staying and leaving at each moment in time, so the principal extracts all gains from further training. As time goes by and the agent becomes more productive, it becomes increasingly expensive to pay

²⁷Phase 2 would also arise if the agent had linear utility provided the minimum consumption \underline{c} was sufficiently large.

²⁸This training rate is a generalized version of the training rates in GR and FR.

them with additional knowledge (i.e., the shadow cost $|S_t|$ grows), so the agent is paid higher wages and is overworked to a lesser degree. Toward the end of phase 2, $|S_t|$ converges to 1. Hence, effort converges to its steady-state level $a^*(\bar{X})$, wages converge to the steady-state output $f(\bar{X}) + a^*(\bar{X})$, and the training rate converges to zero. As a result, neither effort, wages, nor training have jumps at time T .

In practice, apprentices might experience a discreet jump in wages at the time of graduation. This possibility can be accounted for via a simple extension of our model where the principal receives a prize as soon as the agent graduates (for example, because they can start training a new agent). In this case, since the principal is in more of a rush to complete the apprenticeship, the training rate (8) no longer converges to zero toward the end of the apprenticeship. Accordingly, wages remain strictly below (and effort strictly above) their post-graduation levels.

The apprenticeships our model predicts have features that are roughly in line with the real-life practices discussed in Section 2. Phase 1 in our model, for instance, represents a period of intense learning, including classroom-based learning (e.g. during Ph.D. courses or boot camps for new employees), where productivity is relatively low, and phase 2 represents a stage where novices learn below their potential (e.g. because they devote time to work they can already do well) while at the same time producing valuable output for their masters.²⁹ The increasing wage path, wages potentially higher than output at first, and high effort also seem to mirror real-life practices.

While some of these practices might in principle seem abusive to the

²⁹In some real-life apprenticeships, stages of work and study alternate with each other. Our analysis suggests that all phases where learning is carried out at the maximum rate, whether on-the-job or in a laboratory/classroom setting, should be front-loaded. Ph.D. programs seem to follow this idea. The same is true, for example, of the Vermont HITEC two-stage apprenticeship model, where at first “apprentices are immersed in the field of study for nine hours per day, five days a week [with classrooms] typically set up at the employer’s facility,” and then “move into the job setting full-time to apply these technical competencies on a daily basis” (see Vermont HITEC Program Case Study, www.dol.gov/apprenticeship/toolkit.htm, accessed 7/14/20).

agent, our analysis shows that they might actually be beneficial. With this in mind, we turn to the problem of optimal regulation.

6 Regulation

Apprenticeships are a frequent target of regulation. Sometimes, regulators intervene in nearly all aspects of the apprenticeship (e.g. wages, duration, curriculum, and sometimes even location), as occurs for instance in the German and Swiss dual-education models.³⁰ Other times, they seem especially concerned with specific aspects of the relationship, such as the ACGME restricting the hours of medical residents in the U.S.³¹

Here we show how a planner can induce the principal to select the most preferred contract for the agent, which is also the contract that maximizes the players' combined payoffs subject to the principal not making losses. To do so, as we shall see, the planner must regulate several aspects of the apprenticeship at once.

To begin, let \mathcal{C}^* denote the agent's ideal contract among those described in Theorem 2 (that is, the contract in which the principal makes 0 profits), and let $(\mathbf{w}^*, \mathbf{a}^*, \mathbf{X}^*)$ denote the agent's lifetime wages, effort, and knowledge paths when trained under this contract. Proposition 1 shows that if the planner can exercise control over the effort and wage paths (through upper bounds on effort and lower bounds on wages), they can implement the agent-optimal contract without any need to exercise control over the knowledge path. This is good news for the planner, as caps on effort (i.e. hours worked) and floors on wages are presumably much easier to enforce than restrictions on the rate of knowledge transfer, as the agent's knowledge can be difficult

³⁰See, e.g., Wyman (2017) and the German Vocational Training Act (www.gesetze-im-internet.de/bbig_2005/, accessed 7/12/20).

³¹These restrictions include both maximum weekly hours (80 on average) as well as limits on hours worked straight (24 in some cases). See www.acgme.org/Portals/0/PDFs/dh-ComparisonTable2003v2011.pdf (accessed 7/12/20).

to monitor by an external regulator.

Proposition 1. *Suppose the planner restricts the effort path to be no higher than \mathbf{a}^* and the wage path to be no lower than \mathbf{w}^* for the duration of the apprenticeship. Then \mathcal{C}^* is the unique profit-maximizing contract.*

Proof. See Appendix B. □

To understand this result, notice that a contract that specifies any other effort and wage paths, while also satisfying the planner’s bounds, would lead to a strictly higher payoff for the agent and hence cause the principal losses. Moreover, because the knowledge path \mathbf{X}^* maximizes the agent’s output given \mathbf{a}^* and \mathbf{w}^* , while also meeting the learning and ongoing participation constraints, the principal will opt for this path.

The planner need not worry about capping the overall length of the contract, even though the principal might in principle be tempted to hold on to the agent for too long. This is because, once the principal is forced to pay the right wages and limit effort, it is in their interest to quickly train the agent in order to raise productivity and make up for these wages.³²

An even better outcome is achieved when the planner helps remedy the agent’s commitment problem. Let \mathcal{C}^{**} denote the agent-optimal (first-best) contract when the agent has commitment power, as characterized in Theorem 1, and let $(\mathbf{w}^{**}, \mathbf{a}^{**}, \mathbf{X}^{**})$ denote the agent’s lifetime wages, effort, and knowledge paths when trained under this contract. Recall that wages are constant throughout the agent’s lifetime, the agent learns as quickly as feasible until fully trained, and the present value of the agent’s lifetime wages equals the present value of output.

Proposition 2. *Suppose the planner grants the principal the power to retain the agent for as long as the principal wishes, so the players no longer face an*

³²In contrast, the regulations proposed by FR, which do not involve wages, require caps on length. GR considers wage regulations, but restricts to a constant minimum wage, and does not consider effort distortions. While these interventions help, they can be improved.

*ongoing participation constraint. Suppose further that the planner restricts the effort path to be no higher than \mathbf{a}^{**} and the wage path to be no lower than \mathbf{w}^{**} for the duration of the apprenticeship. Then \mathcal{C}^{**} is the unique profit-maximizing contract.*³³

Proof. See Appendix B. □

This proposition suggests that certification requirements and non-compete clauses, which are frequently observed, should be accompanied by restrictions on the wage and effort paths. Only then can we guarantee that the agent benefits from them. The intuition is similar to that of Proposition 1, save for the principal no longer needing to worry about the ongoing participation constraint.

7 Conclusion

We have studied the problem of training an apprentice who must work in order to learn. To do so, we introduced the idea of learning-by-doing as an inequality constraint, as this allows the master to strategically slow training down. Perhaps paradoxically, slow training expands the players' payoff frontier, as it allows the principal to capture rents despite the agent's inability to commit to make payments.

In the novice's most preferred contract, the learning constraint causes the novice to initially produce less than they are paid, so they accumulate "debt." The slow-training phase then allows the master to gradually collect on this debt. Because this arrangement allows for better consumption smoothing, the novice prefers it over a shorter apprenticeship with only the first phase.

³³Note that, unlike \mathcal{C}^* , this contract is infinitely long because the agent must forever service the debt they accumulate during training. However, the principal can exit the relationship as soon as training is over by selling the agent's debt to a third party for a lump-sum payment of $\int_T^\infty e^{-r(t-T)}[f(\bar{X}) + a_t^{**} - w_t^{**}]dt$.

Our model helps rationalize why real-life training relationships, including formal apprenticeships, consist of a bundle of interrelated practices, including distinct phases and imperfect consumption smoothing. It also suggests optimal regulation based on the idea that by restricting both the effort and wage paths, the social planner can induce the master to train the novice at the ideal rate.

We have abstracted from the possibility that the master learns about the novice's intrinsic ability during the apprenticeship, which would generate yet richer predictions. We leave this for future work.

References

- Acemoglu, D., 1997. Training and innovation in an imperfect labour market. *Review of Economic Studies*, 64, pp.445-464.
- Acemoglu, D. and Pischke, J.S., 1998. Why do firms train? Theory and evidence. *Quarterly Journal of Economics*, 113, pp.79-119.
- Akerlof, G., 1976. The economics of caste and of the rat race and other woeful tales. *Quarterly Journal of Economics*, 90, pp.599-617.
- Albuquerque, R. and Hopenhayn, H.A., 2004. Optimal lending contracts and firm dynamics. *Review of Economic Studies*, 71, pp.285-315.
- Arrow, K., 1962. The economic consequences of learning by doing. *Review of Economic Studies*, 29, pp.155-173.
- Barlevy, G. and Neal, D., 2019. Allocating effort and talent in professional labor markets. *Journal of Labor Economics*, 37, pp.187-246.
- Ben-Porath, Y., 1967. The production of human capital and the life cycle of earnings. *Journal of Political Economy*, 75, pp.352-365.
- Board, S. and Meyer-ter-Vehn, M., 2013. Reputation for quality. *Econometrica*, 81, pp.2381-2462.
- Bonatti, A. and Hörner, J., 2017. Career concerns with exponential learning. *Theoretical Economics*, 12, pp.425-475.
- Cisternas, G., 2018. Career concerns and the nature of skills. *American Economic Journal: Microeconomics*, 10, pp.152-89.
- Dewatripont, M., Jewitt, I. and Tirole, J., 1999. The economics of career concerns, part II: Application to missions and accountability of government agencies. *Review of Economic Studies*, 66, pp.199-217.

- Fudenberg, D. and Rayo, L., 2019. Training and effort dynamics in apprenticeship. *American Economic Review*, 109, pp.3780-3812.
- Garicano, L. and Rayo, L., 2017. Relational knowledge transfers. *American Economic Review*, 107, pp.2695-2730.
- Hackman, M. 2018. Looking for an Alternative to College? U.S. Studies German Apprenticeships. *Wall Street Journal*. Dec. 31.
- Hartl, R.F., Sethi, S.P. and Vickson, R.G., 1995. A survey of the maximum principles for optimal control problems with state constraints. *SIAM Review*, 37, pp.181-218.
- Holmström, B., 1999. Managerial incentive problems: A dynamic perspective. *Review of Economic Studies*, 66, pp.169-182.
- Jacoby, T., 2014. Why Germany is so much better at training its workers. *The Atlantic*, Oct. 16.
- Katz, E. and Ziderman, A., 1990. Investment in general training: the role of information and labour mobility. *Economic Journal*, 100, pp.1147-1158.
- Killingsworth, M.R., 1982. "Learning by doing" and "investment in training": A synthesis of two "rival" models of the life cycle. *Review of Economic Studies*, 49, pp.263-271.
- Kolb, A. and Madsen, E., 2020. Leaks, Sabotage, and Information Design. Working Paper.
- Landers, R.M., Rebitzer, J.B. and Taylor, L.J., 1996. Rat race redux: Adverse selection in the determination of work hours in law firms. *American Economic Review*, 86, pp.329-348.
- Malcomson, J.M., Maw, J.W. and McCormick, B., 2003. General training by firms, apprentice contracts, and public policy. *European Economic Review*, 47, pp.197-227.

- Morath, E., 2019. Trump Administration Proposes New Type of Apprenticeship. *Wall Street Journal*, June, 24.
- Rosen, S., 1972. Learning and experience in the labor market. *Journal of Human Resources*, 7, pp.326-342.
- Seierstad, A. and Sydsaeter, K., 1986. Optimal control theory with economic applications. Elsevier North-Holland, Inc.
- Smith, A., 1776. The wealth of nations. New York: The Modern Library.
- Stockman F., 2019. Want a White-Collar Career Without College Debt? Become an Apprentice. *New York Times*, Dec. 10.
- Thomas, J. and Worrall, T., 1994. Foreign direct investment and the risk of expropriation. *Review of Economic Studies*, 61, pp.81-108.
- Weiss, Y., 1972. The risk element in occupational and educational choices. *Journal of Political Economy*, 80, pp.1203-1213.
- Wyman, N., 2017. Jobs Now! Learning from the Swiss Apprenticeship Model. *Forbes*, Oct. 20.

A Proof of Theorem 2

The proof is organized in four steps. First, we consider a relaxation of (II) that omits the constraint $z_t \geq 0$, and we introduce a Lagrange multiplier for the constraint $v_0 \geq \underline{v}$. Lemma 1 shows that there exists an optimal solution to this relaxed problem in which z_t , w_t , and a_t are as given in Theorem 2, and it characterizes the trajectories of X_t and v_t , as well as the thresholds θ and T . Second, Lemma 2 shows that this solution is in fact unique. Third, we turn to solving the original problem, (II). We show in Lemma 3 that the solution of the relaxed problem satisfies the omitted constraint $z_t \geq 0$. Finally, we complete the proof by showing that there exists a Lagrange multiplier for the constraint $v_0 \geq \underline{v}$ such that the corresponding solution of the relaxed problem uniquely solves (II).

For a fixed $\omega \geq 0$, consider the following optimal control problem:

$$S(\omega) = \max r \int_0^{\bar{T}} e^{-rt} [f(X_t) + a_t - w_t] dt + \omega v_0 \quad (9)$$

$$\text{s.t. } \dot{X}_t = z_t \quad (10)$$

$$\dot{v}_t = r [v_t - u(w_t) + d(a_t)] \quad (11)$$

$$v_0 \text{ free and } v_T = \eta(X_T)$$

$$z_t \leq L(X_t, a_t) \quad (12)$$

$$v_t \geq \eta(X_t) \quad (13)$$

$$X_0 = \underline{X}, X_t \leq \bar{X} \quad (14)$$

$$a_t \in [0, \bar{a}], w_t \geq \underline{c}.$$

This problem is a relaxation of (II) as we have omitted the constraint $z_t \geq 0$ and replaced the constraint $v_0 \geq \underline{v}$ with the assumption that the principal maximizes the weighted sum of their own and the agent's payoff with weights one and ω , respectively. To obtain (11), we differentiated the agent's continuation payoff with respect to t and used that $v_t = rV_t$. Note that we

multiplied the integral in the objective by r . We also fixed the horizon to be equal to \bar{T} . This will turn out to be without loss of generality, because after the agent's knowledge reaches \bar{X} , they earn their output and the principal's continuation payoff is zero.

We say that a five-tuple $(X_t, v_t, z_t, w_t, a_t)$ is *admissible* if the functions X_t and v_t are piecewise continuously differentiable, and w_t , z_t , and a_t are piecewise continuous functions that satisfy the constraints in (9). Define the functions

$$m(X) := \frac{f'(X)}{\eta'(X)/r} \quad \text{and} \quad \phi(X, w, a) := \frac{\eta(X) - u(w) + d(a)}{\eta'(X)/r}.$$

The following lemma characterizes one optimal solution for this problem.

Lemma 1. *There are $\theta \geq 0$, $T > \theta$, and functions $(X_t, v_t, z_t, w_t, a_t)$ that solve the optimal control problem given in (9) such that:*

- (i) z_t , w_t , and a_t satisfy the expressions given in Theorem 2, where the function $D_t := L_a(X_t, a_t)\mu_t$ and μ_t is defined in (iv) below.
- (ii) On $t \in [\theta, T]$, the function $S_t = -1 + r \int_t^T m(X_s)ds$ and the agent's continuation payoff v_t satisfy the system of ordinary differential equations

$$\begin{bmatrix} \dot{S}_t \\ \dot{v}_t \end{bmatrix} = \begin{bmatrix} -m(\eta^{-1}(v_t)) \\ r[v_t - u(w_t) + d(a_t)] \end{bmatrix}$$

subject to the initial conditions $S_T = -1$ and $v_T = \eta(\bar{X})$.

- (iii) $\theta = \min \{t \geq 0 : S_t = -\omega \text{ or } \eta^{-1}(v_t) = \underline{X}\}$, and $X_t = X_\theta + \int_\theta^T \phi(X_s, w_s, a_s)ds$ with $X_T = \bar{X}$.

- (iv) If $\theta > 0$, then for $t \in [0, \theta]$, v_t , X_t , and the function μ_t uniquely solve the system of ordinary differential equations

$$\begin{bmatrix} \dot{X}_t \\ \dot{v}_t \\ \dot{\mu}_t \end{bmatrix} = \begin{bmatrix} L(X_t, a_t) \\ r(v_t - u(w_t) + d(a_t)) \\ -f'(X_t) + \mu_t(r - L_X(X_t, a_t)) \end{bmatrix}$$

such that $X_0 = \underline{X}$, the initial values X_θ and v_θ are determined from (iii), and $\mu_\theta = 0$.

Proof of Lemma 1.

Define the Hamiltonian

$$\mathcal{H} := re^{-rt} [f(X_t) + a_t - w_t] + p_t^X z_t + p_t^v r [v_t - u(w_t) + d(a_t)],$$

where p_t^X and p_t^v are the co-state variables associated with the state variable X_t and v_t , respectively, and the Lagrangian

$$\begin{aligned} \mathcal{L} := & re^{-rt} [f(X_t) + a_t - w_t] + p_t^X z_t + p_t^v r [v_t - u(w_t) + d(a_t)] \\ & + q_t^L r [L(X_t, a_t) - z_t] + q_t^\eta r [v_t - \eta(X_t)] + q_t^{\bar{X}} (\bar{X} - X_t), \end{aligned}$$

where q_t^L , q_t^η , and $q_t^{\bar{X}}$ are multipliers associated with the agent's learning constraint, their ongoing participation constraint, and the constraint that their knowledge level X_t does not exceed \bar{X} , respectively.³⁴

This problem is a special case of the one considered in Section 6.7 of Seierstad and Sydsaeter (1986), and Theorem 6.13 provides sufficient conditions for a solution to be optimal. To be specific, an admissible five-tuple $(X_t, v_t, z_t, w_t, a_t)$ solves (9) if there exists piecewise continuously differentiable functions p_t^X and p_t^v , and piecewise continuous functions q_t^L , q_t^η , and $q_t^{\bar{X}}$ such that the following conditions are satisfied:

(C.1) The control variables (z_t, w_t, a_t) maximize the Lagrangian \mathcal{L} .

(C.2) The trajectory of the co-state variable p_t^X and p_t^v is governed by

³⁴For convenience, we have multiplied both sides of the first two inequality constraints by r . Doing so is without loss of generality.

the adjoint equation

$$\dot{p}_t^X = -\frac{\partial \mathcal{L}}{\partial X_t} = -re^{-rt}f'(X_t) - q_t^L r L_X(X_t, a_t) + q_t^\eta r \eta'(X_t) + q_t^{\bar{X}}, \text{ and} \quad (15)$$

$$\dot{p}_t^v = -\frac{\partial \mathcal{L}}{\partial v_t} = -p_t^v r - q_t^\eta r, \quad (16)$$

respectively.

(C.3) The functions q_t^L , q_t^η , and $q_t^{\bar{X}}$ satisfy the complementary slackness conditions

$$\begin{aligned} q_t^L &\geq 0 \quad (= \text{if } z_t < L(X_t, a_t)), \\ q_t^\eta &\geq 0 \quad (= \text{if } v_t > \eta(X_t)), \text{ and} \\ q_t^{\bar{X}} &\geq 0 \quad (= \text{if } X_t < \bar{X}). \end{aligned}$$

(C.4) The transversality condition

$$p_0^v \leq -\omega \quad (= \text{if } v_0 > \eta(\underline{X})) \quad (17)$$

is satisfied

(C.5) The Hamiltonian is concave in the state and the control variables for each t , and the right-hand-side of the equality constraints is quasi-concave in the state and the control variables.

To complete the proof, it suffices to show there are constants θ and T , and continuously differentiable functions p_t^X and p_t^v , and piecewise continuous functions q_t^L , q_t^η , and $q_t^{\bar{X}}$ such that the trajectories of $(X_t, v_t, z_t, w_t, a_t)$ satisfy conditions (i)-(iv) of Lemma 1, and these functions together with $(p_t^X, p_t^v, q_t^L, q_t^\eta, q_t^{\bar{X}})$ satisfy conditions (C.1-5).

Let us begin with (C.5). Since $f(X)$ is strictly concave, $\eta(X)$ is strictly increasing, and $L(X, a)$ is additively separable and concave in each of its arguments, this condition is satisfied as long as $p_t^v \leq 0$ for all t .

Next, consider (C.1). Differentiating the Lagrangian with respect to each control variable, we obtain the following expressions:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial z} &= p_t^X - r q_t^L, \\ \frac{\partial \mathcal{L}}{\partial w} &= r [-e^{-rt} - p_t^v u'(w)], \\ \frac{\partial \mathcal{L}}{\partial a} &= r [e^{-rt} + p_t^v d'(a_t) + q_t^L L_a(X_t, a_t)].\end{aligned}$$

We want to show that given the trajectories of X_t , w_t , and a_t , either $z_t = \phi(X_t, w_t, a_t)$, or $z_t = L(X_t, a_t)$. Since both ϕ and L are finite-valued, it must be the case that $p_t^X = r q_t^L$ for all t . After taking into consideration that $w_t \geq \underline{c}$ and $a_t \leq \bar{a}$, it follows from the above expressions that the optimal wage satisfies $u'(w_t) = \min \{u'(\underline{c}), -e^{-rt}/p_t^v\}$, and the optimal effort is implicitly defined by the equation $d'(a_t) = \min \{d'(\bar{a}), -[q_t^L L_a(X_t, a_t) + e^{-rt}]/p_t^v\}$. Because $d'(0) = 0$, $d'' > 0$, and $L_{aa} \leq 0$, as long as $p_t^v < 0$, there exists a unique a_t that satisfies this equation.

We now fix an arbitrary $T \leq \bar{T}$ and a $\theta \in (0, T)$, and we characterize the variables $(p_t^X, p_t^v, q_t^L, q_t^\eta, q_t^{\bar{X}})$ such that (C.1-4) are satisfied. We aim to characterize an optimal solution in which the learning constraint (12) is slack for $t > \theta$, the ongoing participation constraint (13) is slack for $t < \theta$, and the knowledge constraint (14) is slack for $t < T$. Therefore, by the complementary slackness conditions in (C.3), the following must be true:

$$q_t^L \begin{cases} \geq 0 & \text{if } t \leq \theta \\ = 0 & \text{if } t > \theta \end{cases}, \quad q_t^\eta \begin{cases} = 0 & \text{if } t < \theta \\ \geq 0 & \text{if } t \geq \theta \end{cases} \quad \text{and} \quad q_t^{\bar{X}} \begin{cases} = 0 & \text{if } t < T \\ \geq 0 & \text{if } t \geq T. \end{cases} \quad (18)$$

We now characterize the trajectories of the co-state variables by solving

the corresponding adjoint equations. Solving (16) yields

$$p_t^v = -e^{-rt} \left(-p_{\bar{T}}^v e^{r\bar{T}} - r \int_t^{\bar{T}} e^{rs} q_s^\eta ds \right)$$

for some $p_{\bar{T}}^v$ that remains to be determined.

For $t \in [0, T)$, since $p_t^X = r q_t^L$ and $q_t^{\bar{X}} = 0$, (15) can be rewritten as

$$\dot{p}_t^X = -r e^{-rt} f'(X_t) - p_t^X L_X(X_t, a_t) + q_t^\eta r \eta'(X_t),$$

and this ODE admits the following solution:

$$p_t^X = e^{-\int_0^t L_X(X_s, a_s) ds} \left[p_0^X - r \int_0^t (e^{-rs} f'(X_s) - q_s^\eta \eta'(X_s)) e^{\int_0^s L_X(X_\tau, a_\tau) d\tau} ds \right],$$

where p_0^X is an initial value which we determine next. Recall that for $t > \theta$, the learning constraint is slack, so by (18) we have $q_t^L = 0$. This implies that $p_t^X = 0$ for all $t > \theta$. The continuity of p_t^X implies that $p_\theta^X = 0$, and therefore

$$p_0^X = r \int_0^\theta e^{-rs} f'(X_s) e^{\int_0^s L_X(X_\tau, a_\tau) d\tau} ds,$$

where we have used from (18) that $q_t^\eta = 0$ for all $t < \theta$. Because $p_t^X = 0$ for all $t \in (\theta, T)$, $e^{-rs} f'(X_s) - q_s^\eta \eta'(X_s) = 0$, or equivalently,

$$q_t^\eta = e^{-rt} m(X_t) / r \text{ for all } t \in (\theta, T),$$

and recall that by definition, $m(X) = r f'(X) / \eta'(X)$.

For $t \in [T, \bar{T}]$, we must have $p_t^X = 0$ (since $q_t^L = 0$). Since $p_{\bar{T}}^X = 0$, it suffices that $\dot{p}_t^X = 0$ for all $t > T$, or equivalently using (15),

$$q_t^{\bar{X}} = r e^{-rt} f'(\bar{X}) - r q_t^\eta \eta'(\bar{X}).$$

Let us guess that for all $t > T$, $q_t^\eta = 0$ and $p_t^v = -e^{-rt}$.³⁵ Then we have the following expressions for $(p_t^X, p_t^v, q_t^L, q_t^\eta, q_t^{\bar{X}})$:

$$p_t^X = r q_t^L = \begin{cases} r e^{-\int_0^t L_X(X_s, a_s) ds} \int_t^\theta e^{-rs} f'(X_s) e^{\int_0^s L_X(X_\tau, a_\tau) d\tau} ds & \text{if } t \leq \theta \\ 0 & \text{if } t > \theta, \end{cases}$$

$$p_t^v = \begin{cases} -e^{-rt} \left[1 - \int_\theta^T m(X_s) ds \right] & \text{if } t \leq \theta \\ -e^{-rt} \left[1 - \int_t^T m(X_s) ds \right] & \text{if } \theta < t \leq T \\ -e^{-rt} & \text{if } t > T, \end{cases} \quad (19)$$

$$q_t^\eta = \begin{cases} 0 & \text{if } t < \theta \\ e^{-rt} m(X_t) / r & \text{if } \theta < t < T \\ 0 & \text{if } t > T, \text{ and} \end{cases} \quad (20)$$

$$q_t^{\bar{X}} = \begin{cases} 0 & \text{if } t < T \\ r e^{-rt} f'(\bar{X}) & \text{if } t > T. \end{cases}$$

Using the above expressions and that $d'(a(\bar{X})) = u'(f(\bar{X}) + a(\bar{X})) = 1$, the optimal wage and effort satisfy

$$u'(w_t) = \begin{cases} \min \left\{ u'(\underline{c}), \frac{1}{1 - \int_\theta^T m(X_s) ds} \right\} & \text{if } t < \theta \\ \min \left\{ u'(\underline{c}), \frac{1}{1 - \int_t^T m(X_s) ds} \right\} & \text{if } \theta < t < T \\ 1 & \text{if } t > T, \text{ and} \end{cases} \quad (21)$$

$$d'(a_t) = \begin{cases} \min \left\{ d'(\bar{a}), \frac{e^{rt} L_a(X_t, a_t) q_t^L + 1}{1 - \int_\theta^T m(X_s) ds} \right\} & \text{if } t < \theta \\ \min \left\{ d'(\bar{a}), \frac{1}{1 - \int_t^T m(X_s) ds} \right\} & \text{if } \theta < t < T \\ \min \{ d'(\bar{a}), 1 \} & \text{if } t > T. \end{cases} \quad (22)$$

³⁵Since these conditions are sufficient for an optimum, it suffices to show that a solution given this guess exists.

respectively. Notice that $q_t^L \geq 0$, and by assumption, $d'(0) = 0$, $d'' > 0$ and $L_{aa} \leq 0$. So for $t < \theta$, there exists a unique $a_t \in [0, \bar{a}]$ that satisfies the above implicit equation. Finally, because the learning constraint binds for $t < \theta$, while the ongoing participation constraint binds for $t \geq \theta$, the training rate

$$z_t = \begin{cases} L(X_t, a_t) & \text{for } t \in (0, \theta) \\ \phi(X_t, w_t, a_t) & \text{for } t \in [\theta, T) \\ 0 & \text{for } t \in [T, \bar{T}). \end{cases}$$

So far, we have fixed an arbitrary T and $\theta < T$, and characterized the functions $(z_t, w_t, a_t, p_t^X, p_t^v, q_t^L, q_t^\eta, q_t^{\bar{X}})$ such that conditions (C.1-4) are satisfied, and we argued that (C.5) is satisfied by assumption. Notice that these are functions of X_t , which evolves according to $\dot{X}_t = z_t$ subject to the boundary conditions $X_0 = \underline{X}$ and $X_T = \bar{X}$. Moreover, the agent's continuation value v_t must satisfy $v_t > \eta(X_t)$ for all $t < \theta$ and $v_\theta = \eta(X_\theta)$. (By the construction of $\phi(X, w, a)$, $v_t = \eta(X_t)$ for all $t > \theta$.) A priori, there is no guarantee that there exists a T and a θ such that the conditions pertaining to X_t and v_t are satisfied. We now show that this is indeed the case.

First, we will determine the trajectory of v_t and hence that of X_t (since $v_t = \eta(X_t)$) during Phase 2, that is, during the interval $[\theta, T]$. In doing so, we will pin down the duration of this interval (*i.e.*, $T - \theta$). Then we will turn to Phase 1.

Let us fix some arbitrary T . Since $v_t = \eta(X_t)$ on this interval and $\eta(X)$ is strictly increasing, it will be convenient to define the function $\xi(y) := m(\eta^{-1}(y))$ and recall that $m(X) = r f'(X)/\eta'(X)$. Recall that $S_t = -1 + \int_t^T m(X_s) ds$, which can be rewritten in differential form as $\dot{S}_t = -\xi(v_t)$ with $S_T = -1$.

Notice from (21) and (22) that for each $t \in [\theta, T]$, the agent's wage, w_t , and effort, a_t , depends solely on S_t . In particular, it satisfies $u'(w_t) = \min\{u'(\underline{c}), -1/S_t\}$ and $d'(a_t) = \min\{d'(\bar{a}), -1/S_t\}$, respectively. During

Phase 2, the trajectories of S_t and v_t satisfy the following system of ODE:

$$\begin{bmatrix} \dot{S}_t \\ \dot{v}_t \end{bmatrix} = G(S_t, v_t) := \begin{bmatrix} -\xi(v_t) \\ r[v_t - u(w(S_t)) + d(a(S_t))] \end{bmatrix} \quad (23)$$

subject to the initial conditions $S_T = -1$ and $v_T = \eta(\bar{X})$. Because u , d , f , and η have bounded first and second derivatives by assumption, G has bounded partial derivatives and hence it is Lipschitz continuous. Therefore, by the Picard–Lindelof theorem, this system has a unique solution. It immediately follows that $X_t = \eta^{-1}(v_t)$ during the interval $[\theta, T]$.

We now explain how to determine the threshold θ , and hence the duration of Phase 2 using the above solution (and a given T). To do so, we will use the transversality condition (C.4), which using (19) can equivalently be rewritten as $S_\theta \leq -\omega$ ($' ='$ if $v_0 > \eta(\underline{X})$). This condition implies that either (I) $S_\theta \leq -\omega$ and $v_\theta = \eta(\underline{X})$, or (II) $S_\theta = -\omega$ and $v_\theta > \eta(\underline{X})$. In Case (I), Phase 1 is non-existent, $v_\theta = \eta(\underline{X})$, and hence θ is the first time that v_t hits $\eta(\underline{X})$. In Case (II), Phase 1 has strictly positive duration, $v_\theta > \eta(\underline{X})$, and hence θ is the first time that S_t hits $-\omega$. Thus, given a solution to the system of ODE (23), we define

$$\theta := \min \{t : S_t = -\omega \text{ or } \eta^{-1}(v_t) = \underline{X}\}.$$

That is, starting at T and moving backward in time, θ is the first time that S_t hits $-\omega$ or $\eta^{-1}(v_t) = X_t$ hits \underline{X} , whichever occurs first. Because $\dot{S}_t = -\xi(v_t) < 0$ for all t , such θ exists and it is unique (for given T). Let us consider the two cases mentioned above separately:

Case I: If X_t hits \underline{X} first, Phase 1 has zero length. Because the above system of ODE is autonomous, that is, it doesn't explicitly depend on time, without loss of generality, we can shift time by replacing t with $\tilde{t} = T - t$ so that $X_0 = \underline{X}$. Then Phase 2 starts at $\theta = 0$, and the agent's level of knowledge reaches \bar{X} at $\tilde{t} = T - \theta$. In this case, the characterization of a

solution for (9) is complete.

Case II: If S_t hits $-\omega$ first, this procedure determines (i) the duration of Phase 2, which equals $T - \theta$, (ii) the agent's continuation payoff at the beginning of Phase 2, denoted v_θ^* , and (iii) the knowledge level $X_\theta = \eta^{-1}(v_\theta)$. We characterize the duration of Phase 1 and the trajectory of the state and control variables next.

We now characterize the duration of Phase 1 for the case in which S_t hits $-\omega$ first in the procedure described above. It will be convenient to define $\mu_t := e^{rt} q_t^L$. Using (15), and that $p_t^X = r q_t^L$ and $q_t^\eta = q_t^{\bar{X}} = 0$ during Phase 1, we obtain the following expression for the trajectory of μ_t :

$$\dot{\mu}_t = -f'(X_t) + \mu_t (r - L_X(X_t, a_t)),$$

and notice that this ODE is autonomous. The trajectory of X_t , v_t , and μ_t satisfies the following system of ODE:

$$\begin{bmatrix} \dot{X}_t \\ \dot{v}_t \\ \dot{\mu}_t \end{bmatrix} = H(X_t, v_t, \mu_t) := \begin{bmatrix} L(X_t, a_t) \\ r(v_t - u(w_0) + d(a_t)) \\ -f'(X_t) + \mu_t (r - L_X(X_t, a_t)) \end{bmatrix}, \quad (24)$$

subject to the initial conditions $X_\theta = \eta^{-1}(v_\theta^*)$, $v_\theta = v_\theta^*$, and $\mu_\theta = 0$, where v_θ^* was determined in the analysis of Phase 2 above, and $\mu_\theta = 0$ follows from the fact that $q_\theta^L = 0$. The wage $w_0 = \max\{c, u'^{-1}(1/\omega)\}$, and effort a_t is implicitly defined by the equation $d'(a_t) = \min\{d'(\bar{a}), [L_a(X_t, a_t)\mu_t + 1]/\omega\}$. Because by assumption, f'' , L_X , L_{XX} , and L_{aa} are bounded, $L_{Xa} = 0$, d'' is strictly positive, and $L_{aa} \leq 0$, H has bounded partial derivatives and hence it is Lipschitz continuous. Therefore, by the Picard–Lindelof theorem, this system of ODE has a unique solution.

Define t_0 to be the first time such that $X_{t_0} = \underline{X}$. Such t_0 exists and it is unique since $L(X, a) > 0$ for all X and a . Then $\theta - t_0$ is the duration of Phase

1, and the agent's initial payoff is v_{t_0} . Finally, because this system of ODE is autonomous, we can without loss of generality, shift time so that $X_0 = \underline{X}$. To be specific, we can replace t with $\tilde{t} = t - t_0$. Then the above solution continues to satisfy (24) with $X_0 = \underline{X}$, $X_{\theta-t_0} = \eta^{-1}(v_\theta^*)$, $v_{\theta-t_0} = v_\theta^*$, and $\mu_{\theta-t_0} = 0$, where v_θ^* denotes the agent's continuation value at the beginning of Phase 2, which was characterized in the last step. Therefore, in the new time-space, Phase 1 ends at $\tilde{t} = \theta - t_0$, and Phase 2 ends at $\tilde{t} = \theta^* + (T - \theta)$. By assumption, \bar{T} is sufficiently large such that $\tilde{t} < \bar{T}$.

To summarize, we have shown that there exists an admissible five-tuple $(X_t, v_t, z_t, w_t, a_t)$ and thresholds T and θ such that the sufficient conditions (C.1-5) are satisfied. Moreover, this five-tuple and the variables S_t and μ_t satisfy conditions (i)-(iv) of Lemma 1. Specifically, during Phase 1, which lasts from $t = 0$ until θ , the agent's wage is constant and satisfies $u'(w) = \min \{u'(\underline{c}), 1/\omega\}$. Moreover, their effort satisfies $d'(a_t) = \min \{d'(\bar{a}), (1 + D_t)/\omega\}$, where $D_t := L_a(X_t, a_t)\mu_t$, and their training rate is $z_t = L(X_t, a_t)$.³⁶ During Phase 2, which lasts from $t = \theta$ until T , the wage, effort and training rates are $u'(w_t) = \min \{u'(\underline{c}), 1/|S_t|\}$, $d'(a_t) = \min \{d'(\bar{a}), 1/|S_t|\}$, and $z_t = \phi(X_t, w_t, a_t)$. that $S_t < 0$. After T , the agent's knowledge stays constant at \bar{X} , their effort satisfies $d'(a_t) = \min \{d'(\bar{a}), 1\}$, and they earn $w_t = f(\bar{X}) + a_t$, while the principal earns zero (since $f(X_t) + a_t - w_t = 0$ for all $t \geq T$). This completes the proof. \square

We have characterized *one* optimal solution for the relaxed problem given in (9). The following lemma shows that this solution is in fact unique.

Lemma 2. *Consider the optimal control problem given in (9) for a fixed $\omega \geq 0$. This problem has a unique solution.*

Proof of Lemma 2.

³⁶Note that whenever $\theta > 0$, $\omega = -S_\theta$, and so the expressions for the optimal wage and effort are the same as in Theorem 2.

This proof is organized in two steps. First, using Corollary 8.2 of Hartl et al. (1995), we establish uniqueness of the optimal trajectories of the state variables X_t and v_t . Then we show that this implies uniqueness of the optimal trajectories of the control variables.

Hartl et al. (1995) analyze a problem that is similar to (9), except that (i) there is a terminal value function in the objective, whereas we have the initial value function $-\omega v_0$, and (ii) they assume that the initial values of the state variables are fixed, whereas v_0 is free in (9).

We now explain how to modify our relaxed problem, (9) so that it is a special case of the one considered in Hartl et al. (1995), thus allowing us to apply their Corollary 8.2. First, we reverse time in (9) so that time “starts” at \bar{T} and “ends” at 0, and hence the term $-\omega v_0$ becomes a terminal value function. Formally, we use the transformation $\tilde{t} = \bar{T} - t$, and solve the problem in \tilde{t} -space. Second, in every optimal solution to (9), as well as the transformed problem, $X_{\bar{T}} = \bar{X}$.³⁷ Therefore, we fix the “initial” values of the state variables $X_{\bar{T}} = \bar{X}$ and $v_{\bar{T}} = \eta(\bar{X})$, and impose the condition that $X_0 = \underline{X}$ meanwhile v_0 is free (which is permitted in the formulation of Hartl et al. (1995)). The sufficiency conditions given in Theorem 8.2 of Hartl et al. (1995) are identical to Conditions (C.1-5), and hence an optimal solution takes the same form, except that the requirement that $f(X)$ is concave (from

³⁷To see why, towards a contradiction, suppose that there exists another optimal contract C' with $X'_{\bar{T}} < \bar{X}$. Consider a modified version of the relaxed problem given in (9) where \bar{X} is replaced with $X'_{\bar{T}}$; i.e. the principal is endowed with knowledge $X'_{\bar{T}}$ instead of \bar{X} . Because C' is feasible for this modified problem, any solution C'' to this problem must achieve a payoff no less than that of C' . By Lemma 1, one such solution is a contract such that for some $T < \bar{T}$, $X_T = X'_{\bar{T}}$ and the principal earns zero payoff for all $t \geq T$. Now consider extending this contract such that during $(T, T + \Delta]$ for some $\Delta > 0$ sufficiently small, the agent is paid the minimum wage \underline{c} , exerts the maximum effort \bar{a} , and is trained at the zero-rent rate $\phi(X_t, \underline{c}, \bar{a})$. This modified contract is feasible for (9) and because $f(X_t) + \bar{a} > \underline{c}$ for all $t \in (T, T + \Delta]$, it strictly increases the principal’s objective. Therefore, the principal’s objective is strictly higher under this modified contract than under C' , which is a contradiction.

C.5) is replaced by the condition that the function

$$\mathcal{H}^0(X, v, p^X, p^v, t) := \max_{L(X_t, a_t) \geq z_t} re^{-rt} [f(X_t) + a_t - w_t] + p_t^X z_t + p_t^v r [v_t - u(w_t) + d(a_t)]$$

is concave in X and v for any given p^X, p^v and t . If in addition, $\mathcal{H}^0(X, v, p^X, p^v, t)$ is strictly concave in X and v for any given p^X, p^v and t , then by Corollary 8.2, the optimal trajectory of the state variables, X_t and v_t is unique. We will first show that this is indeed the case, and then argue that the trajectories of the control variables z_t, w_t , and a_t are also unique.

Let us fix $t, X_t, v_t, p^X \geq 0$, and $p^v \leq 0$, and evaluate $\mathcal{H}^0(X, v, p^X, p^v, t)$. For this purpose, we write the Lagrangian for the static problem:

$$\begin{aligned} \tilde{L}(\kappa) &= \max_{z, w, a} re^{-rt} [f(X) + a - w] + p^X z + p^v r [v - u(w) + d(a)] + \kappa r [L(X, a) - z] \\ &\quad \text{s.t. } w \geq \underline{c} \quad \text{and} \quad a \leq \bar{a} \end{aligned}$$

where $\kappa \geq 0$ is a dual multiplier. This problem is convex, and for any κ , the optimal controls satisfy³⁸

$$z = \begin{cases} -\infty & \text{if } p^X < \kappa r \\ \in \mathbb{R} & \text{if } p^X = \kappa r \\ \infty & \text{if } p^X > \kappa r \end{cases}, \quad \begin{aligned} u'(w) &= \min \{u'(\underline{c}), -e^{-rt}/p^v\}, \text{ and} \\ d'(a) &= \min \{d'(\bar{a}), -[e^{-rt} + \kappa L_a(X, a)]/p^v\}. \end{aligned}$$

Moreover, strong duality is satisfied, so $\mathcal{H}^0(X, v, p^X, p^v, t) = \min_{\kappa \geq 0} \tilde{L}(\kappa)$. We argue that the Lagrangian-minimizing $\kappa = p^X/r$. That is because for any $\kappa < p^X/r$ ($\kappa > p^X/r$), $L(\kappa)$ can be made ∞ by setting $z = \infty$ ($z = -\infty$).

Noting that $L(X, a)$ is additively separable in X and a by assumption, and the optimal control variables, z, w, a are all independent of X and v . We

³⁸The expressions for w and a assume that $p^v < 0$. If $p^v = 0$, then $w = \underline{c}$ and $a = \bar{a}$ is optimal.

now compute the Hessian of $\mathcal{H}^0(X, v, p^X, p^v, t)$,

$$\begin{bmatrix} \partial^2 H^0 / \partial X^2 & \partial^2 H^0 / \partial X \partial v \\ \partial^2 H^0 / \partial v \partial X & \partial^2 H^0 / \partial v^2 \end{bmatrix} = \begin{bmatrix} re^{-rt} f''(X) + \lambda r L_{XX}(X, a) & 0 \\ 0 & 0 \end{bmatrix},$$

and observe that it is negative semidefinite since $f''(X) < 0$ and $L_{XX}(X, a) \leq 0$ for all X and a by assumption. Therefore, $\mathcal{H}^0(X, v, p^X, p^v, t)$ is strictly concave in X and v , and hence the optimal trajectory of the state variables, X_t and v_t is unique.

Since $\dot{X}_t = z_t$, this immediately implies that the trajectory of z_t is also unique. We now show that the trajectories of w_t and a_t are unique as well. Towards this goal, let us define $k_t := u(w_t) - d(a_t)$, and note that it is also unique as $\dot{v}_t = r(v_t - k_t)$ and v_t are unique. Define θ such that $z_t = L(X_t, a_t)$ for all $t < \theta$, and $z_t < L(X_t, a_t)$ for all $t > \theta$. Such θ is uniquely determined since X_t and z_t are unique.

First, consider $t < \theta$. Because $L(X, a)$ is strictly increasing in a , the trajectory of a_t and hence that of $w_t = u^{-1}(k_t + d(a_t))$ is also unique on $[0, \theta]$. Next, consider $t > \theta$. Any optimal solution must satisfy the first-order conditions $u'(w_t) = \min\{u'(\underline{c}), -e^{-rt}/p_t^v\}$ and $d'(a_t) = \min\{d'(\bar{a}), -[q_t^L L_a(X_t, a_t) + e^{-rt}]/p_t^v\}$ are satisfied, and $q_t^L = 0$ for all $t > \theta$ (see, for example, Theorem 6.15 in Seierstad and Sydsaeter (1986)). Observe that either $k_t = u(\underline{c}) - d(\bar{a})$, or k_t is a strictly increasing function of p_t^v . Since k_t is unique, then so is p_t^v on the domain such that $k_t > u(\underline{c}) - d(\bar{a})$. Therefore, w_t and a_t are also unique for such t . \square

Recall that in relaxing the original problem, we omitted the constraint $z_t \geq 0$. The next lemma shows that this constraint is in fact satisfied in the solution given in Lemma 1.

Lemma 3. *Consider the optimal control problem given in (9) for a fixed $\omega \geq 0$. In the unique solution characterized in Lemma 1, the training rate $z_t \geq 0$ for all t .*

Proof of Lemma 3.

Clearly, $z_t = L(X_t, a_t) \geq 0$ for all $t < \theta$ since $L(X, a) > 0$ for all X and a by assumption. For $t \geq \theta$, we have $z_t = \phi(X_t, w_t, a_t) = \dot{v}_t/\eta'(X_t)$, where we have used the fact that for such t , the ongoing participation constraint binds so $v_t = \eta(X_t)$, and that $\dot{v}_t = r[v_t - u(w_t) + d(a_t)]$. Since $\eta'(X) > 0$ for all X , it suffices to show that $\dot{v}_t \geq 0$ for all $t \geq \theta$. Towards a contradiction, suppose that there exists some $t' \geq \theta$ such that $\dot{v}_{t'} < 0$. This implies that for $dt > 0$ sufficiently small, $v_{t'+dt} < v_{t'}$. Recall that $u'(w_t) = \min\{u'(\underline{c}), -1/S_t\}$, $d'(a_t) = \min\{d'(\bar{a}), -1/S_t\}$, and $\dot{S}_t = -\xi(v_t) < 0$, implying that $-u(w_t) + d(a_t)$ is weakly decreasing in t . Therefore,

$$\dot{v}_{t'+dt} = r[v_{t'+dt} - u(w_{t'+dt}) + d(a_{t'+dt})] < r[v_{t'} - u(w_{t'}) + d(a_{t'})] = \dot{v}_{t'} < 0.$$

By induction, it follows that $\dot{v}_t < 0$ and hence $\dot{X}_t < 0$ for all $t > t'$. This however, contradicts the fact that $X_{\bar{T}} = \bar{X}$ and $X_t \leq \bar{X}$ for all t . Therefore, we conclude that such t' cannot exist, and hence $z_t \propto \dot{v}_t \geq 0$ for all t . \square

To complete the proof of Theorem 2, we will show that for an appropriately chosen $\omega \geq 0$, the solution of the relaxed problem (9) solves the original problem (II). Let us denote the contract which solves (9) for given ω by $C(\omega) = \{X_t, v_t, z_t, w_t, a_t\}$ with corresponding ex-ante payoffs $\pi_0^*(\omega)$ and $v_0^*(\omega)$ for the principal and the agent, respectively. Thus, $S(\omega) = \pi_0^*(\omega) + \omega v_0^*(\omega)$. We will show that the contract $C(\omega)$ for the smallest ω such that $v_0^*(\omega) \geq \underline{v}$ uniquely solves the original problem (II).

First, we claim that $v_0^*(\omega)$ is strictly increasing in ω , while $\pi_0^*(\omega)$ is strictly decreasing in ω . To see why the first claim is true, because $C(\omega)$ uniquely solves (9), for any pair ω, ω' we have

$$\begin{aligned} \pi_0^*(\omega') + \omega' v_0^*(\omega') &> \pi_0^*(\omega) + \omega' v_0^*(\omega), \text{ and} \\ \pi_0^*(\omega) + \omega v_0^*(\omega) &> \pi_0^*(\omega') + \omega v_0^*(\omega'). \end{aligned}$$

Therefore,

$$\omega' [v_0^*(\omega') - v_0^*(\omega)] > \pi_0^*(\omega) - \pi_0^*(\omega') > \omega [v_0^*(\omega') - v_0^*(\omega)], \quad (25)$$

implying that $v_0^*(\omega') - v_0^*(\omega) > 0$ if and only if $\omega' > \omega$. It follows from (25) that for any ω and $\omega' > \omega$, $\pi_0^*(\omega) - \pi_0^*(\omega') > 0$, which implies the second claim.

Next, we show that $v_0^*(\infty) := \lim_{\omega \rightarrow \infty} v_0^*(\omega) > \underline{v}$. Note first that as $\omega \rightarrow \infty$, the wages prescribed by $C(\omega)$ go to infinity. This implies $\pi_0^*(\infty) = -\infty$, and hence $\pi_0^*(\omega) < 0$ for all large ω . Now suppose towards a contradiction that $v_0^*(\infty) := \lim_{\omega \rightarrow \infty} v_0^*(\omega) \leq \underline{v}$. This implies that there exists a large ω' such that $\pi_0^*(\omega') < 0$ and $v_0^*(\omega') \leq \underline{v}$, which in turn implies that there is no feasible contract $C(\omega')$ for the relaxed problem such that $\pi_0^*(\omega') \geq 0$ and $v_0^*(\omega') \geq \underline{v}$. This leads to a contradiction because, by assumption, \underline{v} is sufficiently small such that the principal can fully train the agent while meeting the initial participation constraint and obtaining non-negative profits.

Moreover, because the trajectories of w_t and a_t , which together determine the agent's payoff, vary continuously with ω, θ, T , and the latter two variables vary continuously with ω , $\pi_0^*(\omega)$ and $v_0^*(\omega)$ are continuous in ω .

Let us denote

$$\omega^* = \inf \{ \omega \geq 0 : v_0^*(\omega) \geq \underline{v} \}.$$

We will now show that the solution of (9) corresponding to $\omega = \omega^*$ uniquely solves (II). There are two cases to consider:

Case 1: $v_0^*(0) \geq \underline{v}$. In this case $C(\omega^*) = C(0)$ uniquely solves the original problem because it uniquely maximizes profits in the relaxed version of the original problem where the agent's initial participation constraint is ignored, and yet it is satisfied by $C(0)$. Uniqueness follows directly from Lemma 2.

Case 2: $v_0^*(0) < \underline{v}$. Suppose contrary to the claim, that $C(\omega^*)$ does not solve the original problem. Then, because $C(\omega^*)$ is feasible for the original problem, there must exist another feasible contract C' leading to payoffs

$\pi'_0 \geq \pi_0^*(\omega^*)$ and $v'_0 \geq v_0^*(\omega^*) = \underline{v}$, so

$$\pi'_0 + \omega^* v'_0 \geq \pi_0^*(\omega^*) + \omega^* v_0^*(\omega^*).$$

But since C' is also feasible for the auxiliary problem, this contradicts the fact that $C(\omega^*)$ uniquely solves the auxiliary problem. This completes the proof.

B Proofs of Propositions 1 and 2

Proof of Proposition 1.

We will show that C^* is the unique solution to the problem that is identical to (II) with the addition of the planner's constraints, that is,

$$\begin{aligned} \max_{\mathcal{C}} \quad & \int_0^T e^{-rt} [f(X_t) + a_t - w_t] dt & \text{(III)} \\ \text{s.t.} \quad & (1), (2), (3) \\ & z_t \geq 0, X_0 = \underline{X}, X_t \leq \bar{X} \\ & a_t \in [0, \bar{a}], w_t \geq \underline{c}, T \leq T_{\max} \\ & a_t \leq a_t^* \quad \text{and} \quad w_t \geq w_t^* \quad \text{for all } t. \end{aligned}$$

Let v_0^* denote the agent's initial payoff under the contract C^* . By Theorem 2, C^* uniquely solves problem (II) with $\underline{v} = v_0^*$, and by definition, the corresponding wage and effort path satisfies the planner's constraints. Let $\widehat{\Pi}(\underline{v})$ denote the principal's expected payoff at $t = 0$ evaluated under the contract which uniquely solves problem (II) given the agent's initial outside option \underline{v} .

Towards a contradiction, suppose there exists a contract C' that solves (III), yet it does not coincide with C^* almost everywhere. If C' prescribes a wage or effort path that differs from \mathbf{w}^* or \mathbf{a}^* , respectively, this contract

must give the agent a strictly higher initial payoff, that is, $v'_0 > v_0^*$. Since (II) is a relaxation of (III), the principal's payoff must be weakly smaller than $\widehat{\Pi}(v'_0)$. It follows from the last step of the proof of Theorem 2 that $\widehat{\Pi}(\underline{v})$ is strictly decreasing in \underline{v} , and so $\widehat{\Pi}(v'_0) < \widehat{\Pi}(v_0^*)$. Therefore, the principal's payoff under \mathcal{C}' must be strictly smaller than $\widehat{\Pi}(v_0^*)$, contradicting the premise that \mathcal{C}' solves (III).

It remains to show that the principal will select a knowledge path that, for the duration of the apprenticeship, coincides with \mathbf{X}^* . To see why, observe that given the effort and wage paths, any profit-maximizing knowledge path must maximize the agent's output subject to the learning and ongoing participation constraints. Per Theorem 2, \mathbf{X}^* grows at rate $L(X_t^*, a_t^*)$ whenever the ongoing participation constraint is slack and at the zero-rent rate given in (8) whenever the ongoing participation constraint binds; hence, it uniquely satisfies this criterion. \square

Proof of Proposition 2.

We will show that the contract \mathcal{C}^{**} uniquely solves

$$\max_{\mathcal{C}} \int_0^{\infty} e^{-rt} [f(X_t) + a_t - w_t] dt \quad (\text{IV})$$

$$\text{s.t. } z_t \leq L(X_t, a_t)$$

$$v_0 = \int_0^{\infty} e^{-rt} [u(w_t) - d(a_t)] dt \geq \underline{v} \quad (26)$$

$$z_t \geq 0, \quad X_0 = \underline{X}, \quad X_t \leq \bar{X}$$

$$a_t \in [0, \bar{a}], \quad w_t \geq \underline{c}$$

$$a_t \leq a_t^{**} \quad \text{and} \quad w_t \geq w_t^{**} \quad \text{for all } t. \quad (27)$$

By construction, \mathcal{C}^{**} is feasible for (IV). Let v_0^{**} denote the agent's initial

payoff under this contract.³⁹ Because the planner's constraints (27) bind under \mathcal{C}^{**} for all t , any effort and wage path that satisfies (27) must give the agent a payoff at least as large as v_0^{**} . Given this observation, we will consider a relaxation of (IV) where we replace (27) and the agent's initial participation constraint $v_0 \geq \underline{v}$ with the constraint $v_0 \geq v_0^{**}$. Because \mathcal{C}^{**} is feasible for (IV), if it uniquely solves this relaxed problem, then it also solves (IV) uniquely.

Towards solving this (relaxed) problem, consider, for a fixed multiplier $\gamma \geq 0$, the following (doubly-relaxed) optimal control problem:

$$\begin{aligned}
S'(\gamma) = \max & \int_0^\infty e^{-rt} [f(X_t) + a_t - w_t] dt + \gamma \int_0^\infty e^{-rt} [u(w_t) - d(a_t)] dt \quad (28) \\
\text{s.t. } & \dot{X}_t = z_t \\
& z_t \leq L(X_t, a_t) \\
& X_0 = \underline{X}, \quad X_t \leq \bar{X} \\
& w_t \geq \underline{c}, \quad a_t \in [0, \bar{a}].
\end{aligned}$$

This problem is identical to (29) in the Online Appendix after substituting $\gamma = 1/\beta$, and Lemma 4 there shows that it has a unique solution. For every $\gamma \geq 0$, let \mathbf{X}^γ , \mathbf{a}^γ and \mathbf{w}^γ denote the knowledge, effort, and wage path that solves this problem, respectively, and define the corresponding principal's payoff $Y(\gamma) := \int_0^\infty e^{-rt} [f(X_t^\gamma) + a_t^\gamma - w_t^\gamma] dt$ and agent's payoff $V(\gamma) := \int_0^\infty e^{-rt} [u(w_t^\gamma) - d(a_t^\gamma)] dt$. By the last step of the proof of Theorem 1, $Y(\gamma)$ is strictly decreasing in γ , while $V(\gamma)$ is strictly increasing in γ . Moreover, there exists a unique γ^* such that $Y(\gamma^*) = 0$. By construction, $V(\gamma^*) = v_0^{**}$, and the solution corresponding to γ^* is the contract \mathcal{C}^{**} .

Finally, we argue that \mathcal{C}^{**} solves the relaxed problem where we replaced (27) and the agent's initial participation constraint $v_0 \geq \underline{v}$ with the constraint

³⁹By Assumption 5.2, $v_0^{**} \geq \underline{v}$

$v_0 \geq v_0^{**}$. Towards a contradiction, suppose there exists another solution with four-tuple $(\tilde{X}_t, \tilde{a}_t, \tilde{w}_t, \tilde{z}_t)$ such that the principal's payoff $\tilde{Y} \geq Y(\gamma^*)$ and the constraint $\tilde{v}_0 \geq v_0^{**}$ is satisfied. But then this implies that

$$\tilde{Y} + \gamma^* \int_0^{\infty} e^{-rt} [u(\tilde{w}_t) - d(\tilde{a}_t)] dt \geq S'(\gamma^*),$$

contradicting the fact that \mathcal{C}^{**} uniquely solves (28) when $\gamma = \gamma^*$. This completes the proof. \square

Working to Learn: Online Appendix

Here we prove Theorem 1.

The proof is organized in three steps. First, we consider a finite-horizon version of (I) in which we introduce a Lagrange multiplier for the credit-balance constraint, (4). Lemma 4 shows that for a sufficiently long horizon length, this problem admits a unique solution in which the wage, effort and training rate satisfy expressions similar to those given in Theorem 1. This lemma also characterizes the trajectory of X_t , and the duration of the learning phase, T . Second, we argue that this characterization is preserved as we take the length of the horizon to infinity. Finally, we complete the proof by showing that there exists a Lagrange multiplier for the constraint (4) such that the corresponding solution of the relaxed problem uniquely solves (I).

For fixed $\beta \geq 0$, consider the following optimal control problem:

$$S(\beta) = \max \int_0^\tau e^{-rt} [u(w_t) - d(a_t)] dt + \beta \int_0^\tau e^{-rt} [f(X_t) + a_t - w_t] dt \quad (29)$$

$$\begin{aligned} \text{s.t. } \dot{X}_t &= z_t \\ z_t &\leq L(X_t, a_t) \\ X_0 &= \underline{X}, X_t \leq \bar{X}, X_\tau \text{ free} \\ a_t &\in [0, \bar{a}]. \end{aligned} \quad (30)$$

for some finite but large τ .⁴⁰ We say that a four-tuple (X_t, z_t, w_t, a_t) is admissible if the function X_t is piecewise continuously differentiable, and w_t , z_t , and a_t are piecewise continuous functions that satisfy the constraints in (29).

The following lemma characterizes the optimal solution for this problem.

⁴⁰It suffices to set $\tau > \hat{T}$, where \hat{T} is the first time that the function X_t hits \bar{X} given that $X_0 = \underline{X}$, $\dot{X}_t = L(X_t, \hat{a})$ and \hat{a} satisfies $d'(\hat{a}) = \min \{d'(\bar{a}), \beta\}$.

Lemma 4. Fix a $\beta \geq 0$. There exists a unique $T > 0$ and functions (X_t, z_t, w_t, a_t) that solve the optimal control problem given in (I) such that:
(i) $u'(w_t) = \beta$, a_t satisfies $d'(a_t) = \min\{d'(\bar{a}), \beta[1 + \mu_t L_a(X_t, a_t)]\}$, and

$$z_t = \begin{cases} L(X_t, a_t) & \text{if } t < T \\ 0 & \text{if } t \geq T. \end{cases}$$

(ii) For $t \in [0, T]$, the functions X_t and μ_t satisfy the system of ODE

$$\begin{bmatrix} \dot{X}_t \\ \dot{\mu}_t \end{bmatrix} = \begin{bmatrix} L(X_t, a_t) \\ -f'(X_t) + [r - L_X(X_t, a_t)] \mu_t \end{bmatrix}$$

subject to the conditions $X_0 = \underline{X}$, $X_T = \bar{X}$, and $\mu_T = 0$.

(iii) For $t > T$, $X_t = \bar{X}$ and $\mu_t = 0$.

Proof of Lemma 4.

Define the Hamiltonian

$$\mathcal{H} := e^{-rt} [u(w_t) - d(a_t)] + \beta e^{-rt} [f(X_t) + a_t - w_t] + p_t^X z_t,$$

where p_t^X is the co-state variable associated with the state variable X_t , and the Lagrangian

$$\mathcal{L} := \mathcal{H} + q_t^L [L(X_t, a_t) - z_t] + q_t^{\bar{X}} (\bar{X} - X_t),$$

where q_t^L and $q_t^{\bar{X}}$ are the multipliers associated with the agent's learning constraint and the constraint that their knowledge level X_t does not exceed \bar{X} , respectively.

This problem is a special case of the one considered in Section 6.7 of Seierstad and Sydsaeter (1986), and Theorem 6.13 provides sufficient conditions for a solution to be optimal. To be specific, an admissible four-tuple

(X_t, w_t, a_t, z_t) solves (29) if there exists a piecewise continuously differentiable function p_t^X , and piecewise continuous functions q_t^L and $q_t^{\bar{X}}$ such that the following conditions are satisfied:

(C.1) The control variables (w_t, a_t, z_t) maximize the Lagrangian \mathcal{L} .

(C.2) The trajectory of the co-state variable p_t^X is governed by the adjoint equation

$$\dot{p}_t^X = -\frac{d\mathcal{L}}{dX} = -\beta e^{-rt} f'(X_t) - q_t^L L_X(X_t, a_t) + q_t^{\bar{X}}. \quad (31)$$

(C.3) The functions q_t^L and $q_t^{\bar{X}}$ satisfy the complementary slackness conditions

$$\begin{aligned} q_t^L &\geq 0 \quad (= \text{if } z_t < L(X_t, a_t)), \text{ and} \\ q_t^{\bar{X}} &\geq 0 \quad (= \text{if } X_t < \bar{X}). \end{aligned}$$

(C.4) The Hamiltonian is concave in the state and the control variables for each t , and the right-hand-side of the equality constraints is quasi-concave in the state and the control variables.

To complete the proof, it suffices to show there is a continuously differentiable function p_t^X and piecewise continuous functions q_t^L and $q_t^{\bar{X}}$ such that the trajectories of (X_t, w_t, a_t, z_t) satisfy conditions (i)-(iii) of Lemma 4, and these functions together with $(p_t^X, q_t^L, q_t^{\bar{X}})$ satisfy conditions (C.1-4).

Let us begin with (C.4). Since $f(X)$ is strictly concave and $L(X, a)$ is additively separable and concave in each of its arguments, this condition is satisfied.

Next, consider (C.1). Differentiating the Lagrangian with respect to each

control variable, we obtain the following expressions:

$$\begin{aligned}\frac{d\mathcal{L}}{dz} &= p_t^X - q_t^L, \\ \frac{d\mathcal{L}}{dw} &= e^{-rt} [u'(w_t) - \beta], \text{ and} \\ \frac{d\mathcal{L}}{da} &= -e^{-rt} d'(a_t) + \beta e^{-rt} + q_t^L L_a(X_t, a_t).\end{aligned}$$

We want to show that given the trajectories of X_t and a_t , $z_t = \{0, L(X_t, a_t)\}$ for all t . Since L is finitely-valued, it must be the case that $p_t^X = q_t^L$ for all t . After taking into consideration that $a_t \leq \bar{a}$, it follows from the above expressions that the optimal wage satisfies $u'(w_t) = \beta$ and the optimal effort is implicitly defined by the equation $d'(a_t) = \min \{d'(\bar{a}), q_t^L e^{rt} L_a(X_t, a_t) + \beta\}$. Because $d'(0) = 0$, $d'' > 0$, $q_t^L \geq 0$, and $L_{aa} \leq 0$, there exists a unique a_t that satisfies this equation.

Fix an arbitrary $T \leq \tau$. We will characterize the variables $(p_t^X, q_t^L, q_t^{\bar{X}})$ such that (C.1-3) are satisfied. We wish to characterize a solution in which the learning constraint binds if and only if $t < T$, and the constraint that $X_t \leq \bar{X}$ binds if and only if $t \geq T$. Therefore, by the complementary slackness conditions in (C.3), the following must be true:

$$q_t^L \begin{cases} \geq 0 & \text{if } t < T \\ = 0 & \text{if } t > T \end{cases} \quad \text{and} \quad q_t^{\bar{X}} \begin{cases} = 0 & \text{if } t < T \\ \geq 0 & \text{if } t > T. \end{cases} \quad (32)$$

We now characterize the trajectory of the co-state variable p_t^X . For $t \in [0, T)$, using that $p_t^X = q_t^L$ and $q_t^{\bar{X}} = 0$, (31) can be rewritten as

$$\dot{p}_t^X = -\beta e^{-rt} f'(X_t) - p_t^X L_X(X_t, a_t).$$

This ODE admits the solution

$$p_t^X = e^{-\int_0^t L_X(X_s, a_s) ds} \left[p_0^X - \beta \int_0^t e^{-rs} f'(X_s) e^{\int_0^s L_X(X_\tau, a_\tau) d\tau} ds \right],$$

where p_0^X is an initial value which we determine next. Recall that for $t \geq T$, the learning constraint is slack, and so by (32) we have $q_t^L = 0$. This implies that $p_t^X = 0$ for all $t \geq T$. The continuity of p_t^X implies that $p_T^X = 0$, and therefore,

$$p_0^X = \beta \int_0^T e^{-rs} f'(X_s) e^{\int_0^s L_X(X_\tau, a_\tau) d\tau} ds.$$

Because $p_t^X = 0$ for all $t > T$, it follows from (31) that

$$q_t^{\bar{X}} = \beta e^{-rt} f'(X_t) \text{ for } t > T.$$

Therefore, we have the following expressions for $(p_t^X, q_t^L, q_t^{\bar{X}})$:

$$p_t^X = q_t^L = \begin{cases} \beta \int_t^T e^{-rs} f'(X_s) e^{\int_t^s L_X(X_\tau, a_\tau) d\tau} ds & \text{if } t \leq T \\ 0 & \text{if } t > T, \text{ and} \end{cases}$$

$$q_t^{\bar{X}} = \begin{cases} 0 & \text{if } t \leq T \\ \beta e^{-rt} f'(X_t) & \text{if } t > T. \end{cases}$$

It will be convenient to define

$$\mu_t := \frac{q_t^L e^{rt}}{\beta} = \int_t^T e^{-r(s-t)} f'(X_s) e^{\int_t^s L_X(X_\tau, a_\tau) d\tau} ds,$$

which can be written in differential form as

$$\dot{\mu}_t = -f'(X_t) + [r - L_X(X_t, a_t)] \mu_t$$

for $t < T$, whereas $\mu_t = 0$ for all $t \geq T$. Using the definition of μ_t , the

first-order conditions with respect to w and a can be rewritten as

$$u'(w_t) = \beta, \text{ and} \tag{33}$$

$$d'(a_t) = \min \{d'(\bar{a}), \beta [1 + \mu_t L_a(X_t, a_t)]\}. \tag{34}$$

Finally, because the learning constraint binds for $t < T$, while the constraint that $X_t \leq \bar{X}$ binds for $t \geq T$, we have

$$z_t = \begin{cases} L(X_t, a_t) & \text{if } t < T \\ 0 & \text{if } t \geq T. \end{cases}$$

So far, we have fixed an arbitrary T and characterized the functions $(z_t, w_t, a_t, p_t^X, q_t^L, q_t^{\bar{X}})$ such that conditions (C.1-3) are satisfied, and we argued that (C.4) is satisfied by assumption. Notice that these are functions of X_t , which evolves according to $\dot{X}_t = z_t$, and must satisfy $X_0 = \underline{X}$ and $X_T = \bar{X}$. A priori, there is no guarantee that there exists a T such that the conditions pertaining to X_t are satisfied. We now show that this is indeed the case. To be specific, we will characterize the trajectories of X_t , μ_t , and a_t , and in doing so, we will pin down the duration of the training phase. During the learning phase, the trajectories of X_t and μ_t satisfy the following system of ODE:

$$\begin{bmatrix} \dot{X}_t \\ \dot{\mu}_t \end{bmatrix} = F(X_t, \mu_t) := \begin{bmatrix} L(X_t, a_t) \\ -f'(X_t) + [r - L_X(X_t, a_t)] \mu_t \end{bmatrix}, \tag{35}$$

where a_t is the unique solution of (34), and notice that it depends solely on X_t and μ_t .

Fix an arbitrary T , and consider this system of ODE subject to the initial value conditions $X_T = \bar{X}$ and $\mu_T = 0$. Because u , d , f , and L have bounded first and second derivatives by assumption, F has bounded partial derivatives and hence it is Lipschitz continuous. Therefore, by the Picard–

Lindelof theorem, this system has a unique solution. Define t_0 to be the first time such that $X_{t_0} = \underline{X}$. Such t_0 exists and it is unique since $L(X, a) > 0$ for all X and a . Because the above system of ODE is autonomous, without loss of generality, we can shift time by replacing t with $\tilde{t} = t - t_0$ so that for $\tilde{t} = 0$, $X_{\tilde{t}} = \underline{X}$. Thus, the training phase begins at $\tilde{t} = 0$ and ends that $\tilde{t} = T - t_0$.⁴¹ During this interval, the agent's learning constraint binds, and the trajectories of X_t and μ_t satisfy (35). The agent's wage and effort satisfies (33) and (34), respectively. This completes the proofs for parts (i)-(iii) of the lemma.

We now show that this solution is in fact unique. The problem given in (29) is a special case of the one studied by Hartl et al. (1995). We will apply their Corollary 8.2, which gives conditions such that trajectory of the state variable is unique. We will then argue that this implies the trajectories of (w_t, a_t, z_t) are also unique. The sufficient conditions given in Theorem 8.2 of Hartl et al. (1995) are identical to Conditions (C.1-4), except that the requirement that $f(X)$ is concave (from C.4) is replaced by the requirement that the function

$$\begin{aligned} \mathcal{H}^0(X, p^X, t) &:= \max e^{-rt} [u(w_t) - d(a_t)] + \beta e^{-rt} [f(X) + a_t - w_t] + p^X z_t \\ \text{s.t. } &w_t \geq \underline{c}, a_t \in [0, \bar{a}], z_t \leq L(X, a_t). \end{aligned}$$

is concave in X for any given p^X and t . If in addition, $\mathcal{H}^0(X, p^X, t)$ is strictly concave in X for any given p^X and t , then by Corollary 8.2, the optimal trajectory of the state variable X_t is unique. Notice that this is a static, convex program. By writing the Lagrangian and observing that strong duality is satisfied, it is easy to verify that $\mathcal{H}^0(X, p^X, t)$ is strictly concave in X using the facts that (a) $f(X)$ is strictly concave, (b) $L(X, a)$ is additively separable in X and a , and (c) $L(X, a)$ is concave in X . Therefore, the trajectory of X_t is unique. Since $z_t = \dot{X}_t$, the trajectory of z_t is also unique. The first-order

⁴¹The assumption that τ is sufficiently large ensures that $\tau > T - t_0$.

conditions, (33) and (34), which determine w_t and a_t are necessary; *i.e.*, they must be satisfied in any optimal contract. Since the optimal trajectory of w_t depends solely on β , its uniqueness follows. Turning to the effort path, because $L_X(X, a)$ does not depend on a , the trajectory of μ_t depends solely on X_t , and it is hence unique. From (34) observe that a_t is uniquely determined by X_t and μ_t , and hence its trajectory is also unique. This completes the proof. \square

Observe that after T , the agent receives no training, a constant wage, and since $\mu_t = 0$, exerts constant effort which is implicitly defined by the equation $d'(a_t) = \min \{d'(\bar{a}), \beta\}$. Moreover, the trajectories of (X_t, w_t, a_t, z_t) and the threshold T do not depend on the horizon τ , provided that it is sufficiently long; *i.e.*, $\tau > T$. Therefore, for any $\tau > T$ and any $\beta > 0$, the solution of (29) is identical. Therefore, this solution also uniquely solves (29) when $\tau = \infty$.

To complete the proof of Theorem 1, we now show that for an appropriately chosen $\beta \geq 0$, the solution of (29) solves the original problem, (I). For each β (and $\tau = \infty$), denote the solution of (29) for given β by $(X_t^\beta, w_t^\beta, a_t^\beta, z_t^\beta)$, and define the agent's discounted payoff $V(\beta) := \int_0^\infty e^{-rt} [u(w_t^\beta) - d(a_t^\beta)]$ and their credit balance $Y(\beta) := \int_0^\infty e^{-rt} [f(X_t^\beta) + a_t^\beta - w_t^\beta] dt$. Notice that $S(\beta) = V(\beta) + \beta Y(\beta)$. We will show that there exists a unique β^* such that $Y(\beta^*) = 0$ which solves (I).

First, we claim that $Y(\beta)$ is strictly increasing in β , while $V(\beta)$ is strictly decreasing in β . For any pair β and β' we have

$$\begin{aligned} V(\beta') + \beta' Y(\beta') &> V(\beta) + \beta' Y(\beta), \text{ and} \\ V(\beta) + \beta Y(\beta) &> V(\beta') + \beta Y(\beta'). \end{aligned}$$

Therefore,

$$\beta' [Y(\beta') - Y(\beta)] > V(\beta) - V(\beta') > \beta [Y(\beta') - Y(\beta)],$$

implying that $Y(\beta') > Y(\beta)$ if and only if $\beta' > \beta$, that is, $Y(\beta)$ is strictly increasing in β . The above inequality chain also implies that for any β and $\beta' > \beta$, $V(\beta) > V(\beta')$; *i.e.*, $V(\beta)$ is strictly decreasing in β .

We will now argue that $Y(\beta)$ single-crosses zero from below. First, consider the case when $\beta \rightarrow 0$. Because u is strictly increasing by assumption, $\lim_{\beta \rightarrow 0} w_t^\beta = \infty$, while the corresponding effort remains bounded (since $a_t^\beta \leq \bar{a}$). Therefore, the agent's credit balance $\lim_{\beta \rightarrow 0} Y(\beta) = -\infty$. On the other hand, for β sufficiently large, $w_t^\beta \simeq 0$ and $a_t^\beta = \bar{a}$ for all t . In this case, $Y(\beta) \simeq \int_0^\infty e^{-rt} [f(X_t) + \bar{a}] dt > 0$ where $\dot{X}_t = L(X_t, \bar{a}) \mathbb{I}_{\{X_t < \bar{X}\}}$. Moreover, because T and the trajectories of X_t , w_t , and a_t , which determine the credit balance $Y(\beta)$, vary continuously with β , $Y(\beta)$ is continuous in β . Therefore, there exists a unique β^* such that $Y(\beta^*) = 0$.

Finally, we argue that the solution to the relaxed problem (29) with $\beta = \beta^*$ solves the original problem, (I). Towards a contradiction, suppose there exists another solution with four-tuple $(\tilde{X}_t, \tilde{a}_t, \tilde{w}_t, \tilde{z}_t)$ such that the agent's payoff $\tilde{V} \geq V(\beta^*)$ and (4) is satisfied. But then this implies that

$$\tilde{V} + \beta^* \int_0^\infty e^{-rt} [f(\tilde{X}_t) + \tilde{a}_t - \tilde{w}_t] dt \geq S(\beta^*),$$

which contradicts the fact that $(X_t^{\beta^*}, a_t^{\beta^*}, w_t^{\beta^*}, z_t^{\beta^*})$ uniquely solves (29) when $\beta = \beta^*$. This completes the proof.