Working to Learn*

Drew Fudenberg† George Georgiadis‡ Luis Rayo§

July 20, 2021

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 has novel implications for optimal regulation.

*We thank Robert Akerlof, Daniel Barron, Gonzalo Cisternas, Michael Powell, Orie Shelef, and conference participants at CMID20 for helpful conversations, the National Science Foundation grant SES 1951056 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.

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 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.

We show that 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

[1]Masters 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.

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 a larger knowledge transfer, while phase 2 becomes shorter. Perhaps surprisingly, Phase 2 never disappears completely, even in the agent-optimal contract. 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.

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). Yet, a novice's training need not be uniformly slow. Ph.D. programs, for instance, frequently begin with an intense 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.[3] 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

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

[3]Ph.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).

college—and phases where novices work as they learn (e.g. Stockman, 2019). These work phases allow novices to master new skills, but may also entail a degree of grunt work that slows down their 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.[4] 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.[5]

3. *Growing wages.* As novices gain skills their earnings typical grow. 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.[6] The earnings of Ph.D. students might also grow over time, though less dramatically, as they take on more R.A. or T.A. assignments. While novices no doubt prefer that wages increase holding the starting wage fixed, many of them would likely prefer flatter wages with the same present value.

4. *High effort.* Many novices, whether in traditional apprenticeships or in the early stages of high-skilled careers, encounter heavy workloads with long hours (e.g. Landers et al., 1996, Landrigan et al., 2004, Barlevy and Neal, 2019). While hard work may accelerate learning, and serve as a screening device, it is also 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

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

[5]Employers 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).

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

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.

This debate highlights the need for a deeper theoretical understanding of the problem. Our model, while stylized, helps explain why the above practices may arise, and clarify their impact on the well-being of novices.

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. Our main innovation relative to that work is to incorporate the agent's learning-by-doing constraint. Making the model more realistic in this way introduces technical challenges that did not arise in GR and FR, because the optimal contract has two very different regimes, one where the learning constraint binds and one where it does not, and the learning constraint depends on the endogenous effort level. Thus to characterize the optimal contract we need to pin down the endogenous interface between these regimes, taking into account how it depends on the agent's effort and accumulated knowledge at the transition point. For this reason we cannot use a first-order approach to determine the speed of knowledge transfer as in FR; instead, we conjecture and verify a solution that involves the pasting of the two regimes. We also generalize the form of the agent's outside option: Instead of it being to use the same technology as with the principal, we allow for it to depend in a fairly general way on the agent's stock of knowledge when they leave.

From an applied perspective, the most important difference relative to GR and FR is that they predict a single apprenticeship phase with wages lower than output, and suggest that the apprenticeship will become vanishingly short as competition between masters increases. In contrast, our model explains why the apprentice can go through a phase where wages exceed output, and predicts that the apprenticeship will have non-zero length—and include a phase with artificially slow knowledge transfer—regardless of the extent of competition. In addition, the more general outside option in our model allows us to capture scenarios, such as the agent switching to a new technology, which are not possible under the more restrictive outside option in GR

and FR. The portion of our contract where the participation constraint binds (phase 2) is similar to the contracts that arise in GR and FR once the agent has received an initial knowledge gift. The main difference is that our phase 2 prescribes non-zero effort and wage paths, whereas GR and FR each focus on only one. Moreover, in our model the static first-best effort is decreasing (due to an income effect absent in FR), while in FR it is either constant or increasing.[7]

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.[8]

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, 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.[9] Moreover, 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.

Finally, 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

[7]While in most ways this paper is more general than GR and FR, it only has one kind of effort, so it cannot be used to explain the transition from menial to skilled work.

[8]In Ben Porath (1967) and Killingsworth (1982) the agent has unlimited ability to borrow and save, and their earnings equal their productivity. In our model, earnings can be both above and below that level, depending on the stage of the apprenticeship and the agent's bargaining power. Moreover, except in our first-best benchmark model, the agent can save but not borrow.

[9]In our setting, one way to punish the agent would be if the principal could forbid the agent from using the production technology. If this were possible, the principal would never slow the agent's training.

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 instead because the agent is liquidity constrained.

2 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 $\overline{X} > \underline{X}$. The agent cannot learn on their own, so unless they walk away from the relationship and find an alternative master, the only way their knowledge level can grow is by means of knowledge transfers from the principal. 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 $a_t \in [0, \overline{a}]$ at cost $d(a_t)$ and produce flow output $f(X_t) + a_t$. If they instead leave the principal, they can obtain an outside option worth $h(X_t)$ in flow terms. 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 bellow 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 $\mathbf{a}$ and $\mathbf{c}$) will denote time paths.

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. Because $u'' < 0$, $a^*(X)$ is decreasing in X. We normalize $u'\left(f\left(\overline{X}\right)+a^*(\overline{X})\right)=1$ and assume $a^*(\overline{X}) < \overline{a}$.

Assumption 2. *The agent's outside option h satisfies:*

1. $h(\overline{X}) = u(f(\overline{X})+a^*(\overline{X})) - d(a^*(\overline{X}))$.

2. *h is strictly increasing and has bounded first and second derivatives.*

Part 1 of this assumption says that, once fully trained, the agent is equally productive with or without the principal.[10] Before the agent is fully trained, other than being strictly increasing and having well-behaved derivatives, the outside option can take any value. This allows the outside option to capture a variety of different scenarios, such as the agent using the same technology elsewhere but gaining no further knowledge, switching to another master, or using what they have already learned in a related industry where knowledge has some value.

The speed $z_t := \dot{X}_t$ at which the principal can train the agent at time t is bounded by the *learning constraint*

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

as well as the constraints $0 \leq z_t$ and $X_t \leq \overline{X}$.[11] Other than these constraints, the principal can select any training rate they desire. 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.[12]

Assumption 3. *$L(X,a)$ is additively separable, strictly positive, weakly increasing and weakly concave in each argument, and twice differentiable with bounded first and second derivatives.*

[10] In Section 5 (footnote 30) we remark on the case where the fully-trained agent is more productive elsewhere.

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

[12] As 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.[13]

Assumption 4. *For all* $X \in [\underline{X}, \overline{X}]$ *and all* $a \in [0, \overline{a}]$,

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

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 4 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 $T \leq \overline{T}$, where $\overline{T}$ is an exogenous upper bound of say 200 years.[14] At the start of the relationship, the principal commits to a contract $\mathcal{C} := \left\langle T, (z_t, w_t, a_t)_{t=0}^{T} \right\rangle$, 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. 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 $h(X_T)$. If the agent happens to be fully trained by the time of graduation ($X_T = \overline{X}$) they continue using the same technology as with the principal, but now they keep all of their output.

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

[13] Alternate sufficient conditions involve restrictions on the third partials of L.

[14] Provided $\overline{T}$ is sufficiently large the constraint $T \leq \overline{T}$ will hold with strict inequality.

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

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

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 $h(X_t)/r$, and can also reject the principal's contract altogether and obtain utility $\underline{v}/r$ from an alternative occupation (one can interpret $\underline{v}/r$ as a measure of the agent's bargaining power). Consequently, the principal is bound by the *participation constraints*

$$v_t \geq h\left(X_t\right) \text{ for all } t \leq T, \tag{2}$$
$$v_0 \geq \underline{v}. \tag{3}$$

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

3 Benchmark: The agent 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. We call this benchmark the *(agent) first-best*.[15]

[15]The resulting first-best contract is also profit maximizing if the agent has commitment power and their initial outside option is so large that the best the principal could do is earn zero profits.

Formally, we solve

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

$$\text{s.t.} \quad (1)$$

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

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

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

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.[16]

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 non-negative function $\widetilde{D}_t$ such that their effort path satisfies*

$$d'\left(a_t\right) = \min\left\{d'\left(\overline{a}\right),\ u'\left(w^{**}\right)\left(1+\widetilde{D}_t\right)\right\}. \tag{5}$$

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

$$d'\left(a^{**}\right) = \min\left\{d'\left(\overline{a}\right),\ u'\left(w^{**}\right)\right\}. \tag{6}$$

Proof. See the Online Appendix. □

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

[16] We 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.

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 $\widetilde{D}_t$ in equation (5). This term is given by

$$\widetilde{D}_t = L_a\left(X_t, a_t\right) \int_t^T e^{-r(s-t)} f'\left(X_s\right) e^{\int_t^s L_X(X_\tau, a_\tau) d\tau} ds. \tag{7}$$

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).[17]

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

4 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 or distortions relative to the first best. These distortions, moreover, will be magnified by the principal's desire to extract rents.

The principal's problem is:

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

Note that varying the agent's outside payoff $\underline{v}$ traces the Pareto-efficient implementable payoffs via the initial participation constraint (3) (i.e. $v_0 \geq \underline{v}$).[18]

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

[18] A 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

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 $\overline{X}$ and exerting effort $a^*(\overline{X})$.

Assumption 5. *For every time path of effort* $\mathbf{a}$,

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

where $\widehat{X}_t(\mathbf{a})$ *is the knowledge path when the agent learns at rate* $L\left(X_t, a_t\right)$ *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.[19]

Our next assumption imposes some parametric restrictions.

Assumption 6. *The parameters of the model are such that:*

1. $\underline{v} > h\left(\underline{X}\right)$.
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.[20] Part 2 holds whenever $\underline{v}$ and $\underline{c}$ are sufficiently small.

We now introduce some notation. Given a fixed time path of knowledge, let

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

The ratio m_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 Pareto

solution and verify that it is optimal.

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

[20]This assumption guarantees that the learning constraint binds for some interval of time.

frontier in a world with no learning constraints, where the agent's ongoing participation constraint always binds and the agent is fully trained by time T. We call this the *unconstrained Pareto frontier*.

To understand why the unconstrained Pareto frontier has slope S_t, suppose that at time t the principal commits to giving the agent $e^{r(T-t)}$ utils at time T when the agent's training is complete (which is worth 1 util from the standpoint of t). Because at T the agent's marginal utility is 1, this costs the principal $e^{r(T-t)}$ (or 1 from the standpoint of t), but also increases the agent's continuation payoff by $e^{r(s-t)}$ utils at all times $s < T$. These higher continuation payoffs allow the principal to transfer additional knowledge, and in particular raise the agent's productivity by $e^{r(s-t)}m_s$ at each $s < T$, so from the standpoint of t the principal recoups $\int_t^T m_s ds$.[21] Notice that as we move backward in time, S_t falls in absolute terms (the frontier becomes flatter). This is because the greater the remaining time in the apprenticeship, the larger the output loss from the agent's low knowledge level and lack of liquidity.[22]

Theorem 2 shows that the optimal contract consists of two phases, regardless of the agent's initial outside option $\underline{v}$. The first phase 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, the agent is only trained 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 themself. 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*

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

[21] The interest rate does not impact S_t because agent and principal discount payoffs at the same rate.

[22] In the absence of learning constraints, the profit-maximizing contract begins where $S_t = 0$. In the special case where the agent's outside option consists of working with the same output technology but gaining no further knowledge and utility is linear, we have $m_t = 1/r$ and $S_t = -1 + r(T-t)$, so the profit-maximizing contract absent learning constraints lasts $1/r$ years, as noted in FR.

and is assigned the effort path

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

where $D_\theta = 0$ *and* $D_t \geq 0$ *for all* $t < \theta$.[23]

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*

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

and is assigned the non-increasing effort path

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

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

As the agent's initial outside option grows, phase 1 becomes weakly longer and phase 2 weakly shorter, each of them strictly whenever the agent's initial participation constraint binds. Phase 2, however, never disappears.

Proof. See Appendix A. □

Figure 1 helps explain this result. It depicts two different Pareto frontiers. The higher one is the unconstrained frontier defined above, where the agent's ongoing participation constraint binds. The lower frontier, which we call the *constrained Pareto frontier*, is relevant when the learning constraint is present. It lies below the unconstrained frontier because at time 0 knowledge cannot be instantly raised to the point where the agent's ongoing participation constraint binds.[24]

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 $\underline{v}/r$ otherwise. (In the figure, points **a**, **b**, and **c** indicate three possible

[23] Specifically, $D_t = L_a(X_t, a_t)\int_t^\theta f'(X_s)\, e^{\int_t^s [L_X(X_\tau, a_\tau) - r]d\tau} ds$. This expression differs from (7) only in that the first integral is taken from t to θ instead of from t to T, since the learning constraint (1) becomes slack after θ.

[24] Recall that $\underline{v} > h(\underline{X})$ and the initial participation constraint requires that $v_0 \geq \underline{v}$. The fact that knowledge cannot be instantly raised therefore implies that $X_0 = \underline{X}$ and $v_0 > h(X_0)$.

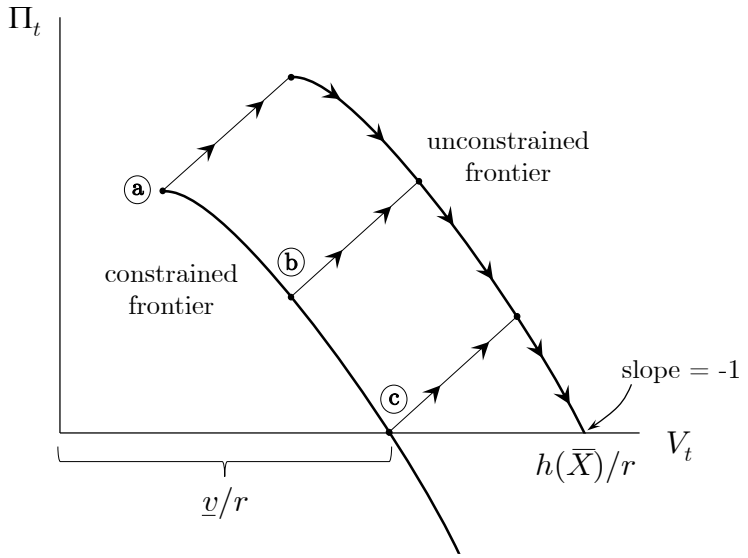

Figure 1: Pareto frontiers and evolution of payoffs

starting points.) Because the unconstrained frontier lies above the constrained one, the principal wishes to reach this higher frontier with as little delay as possible, 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 unconstrained frontier until the agent's training is complete. Further properties of the contract are as follows:

Wages. Wages are constant during phase 1 because the slack participation constraint allows for full consumption smoothing. For reasons that will soon be clear, this wage is higher the steeper the slope of the relaxed frontier at the moment of transition between the two phases. During phase 2, the principal offers an increasing wage path (i.e. backloads wages) in order to relax the binding participation constraint. In this second phase, $|S_t|$ represents the time-t shadow cost of paying the agent with knowledge (i.e. the loss in profits from raising the agent's outside option); modulo the

exogenous lower bound on wages, the optimal contract equates the marginal utility of wages to the inverse of this shadow cost, so that paying the agent with wages is equally costly in the margin as paying them with knowledge.

Effort. At all times before T, the contract prescribes effort strictly above the myopic optimum $a^*(X_t)$, which decreases over time. In phase 1 the effort distortion serves both to relax the learning constraint (with a greater D_t leading to a larger distortion) and to extract rents from the agent (with a smaller $|S_\theta|$ leading to a greater distortion). During this phase, effort need not be monotone.[25] During phase 2, in contrast, rent extraction is the single motive for the distortion. Because the principal pays for effort with knowledge, and the shadow cost of knowledge $|S_t|$ grows as regime 2 evolves, the distortion becomes smaller over time, until finally disappearing at time T; effort therefore decreases over time, strictly so whenever the effort upper bound is slack.

Transition between regimes. The path traced by the players' continuation payoffs must connect the two frontiers at points of equal slope. This condition, which determines the time θ that separates the two phases, is a form of smooth pasting that guarantees that wages and effort do not jump when the apprenticeship switches from one regime to the other. To see why the two slopes must be equal, notice that the participation constraint is slack during the first phase, so the principal can give the agent an extra dV utils in continuation value at the start of the next phase (e.g. by means of a slightly faster knowledge transfer shortly after θ) without altering any of phase 1's characteristics. As a result, the trade-off between profits and agent utility is the same at the start of phase 1 (or indeed at any time during that phase) as it is at the start of phase 2.[26] Throughout phase 1, wages and effort are distorted in inverse proportion to that tradeoff, and hence these distortions fall as $|S_\theta|$ grows.

Impact of agent's outside option. As the agent's initial outside option increases (perhaps because ex-ante competition between different masters becomes more intense) 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

[25]Effort is guaranteed to be weakly decreasing during this phase if $r \leq L_X(X, a)$ for all X and a, as this guarantees that the exponential in D_t (defined in footnote 23) falls over time, i.e., the compounding impact of knowledge on future learning via L_X (which maters less as time goes by and there is less of phase 1 left) is significant enough to overcome the principal's impatience. See footnote 40 in the Appendix.

[26]Specifically, gifting the agent dV at time θ 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 $d\Pi_0/dV_0 = S_\theta$.

means higher wages and lower effort distortions (since $|S_\theta|$ grows), and thus a longer phase 1 and 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 **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 better smooth their consumption, without needing to pay it all back in phase 1. The principal collects on the remaining 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]

Training rate. Throughout phase 1 the agent is trained at rate $L(X_t, a_t)$ to reach the unconstrained frontier as quickly as possible given the desired effort path. In phase 2, the agent is instead trained at a rate that keeps them indifferent between staying with the principal and walking away, given the prescribed wages and effort.[28] We call this the *zero-rent training rate*; it generalizes the zero-rent training rates in GR and FR to allow for the effect of both effort and the agent's desire for income smoothing. It is given by

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

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 principal extracts all gains from further training. As the apprenticeship ends, $|S_t|$ converges to 1, effort converges to its steady-state level $a^*(\overline{X})$, wages converge to steady-state output $f(\overline{X}) + a^*(\overline{X})$, and

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

[28] The training rate z_t is not guaranteed to be monotone in either phase. In phase 1, this is because z_t is positively impacted by both effort and knowledge, and the former may decrease over time. In phase 2, it is because knowledge and the agent's flow payoff (both of which increase over time) have opposing effects on z_t, as shown below (and h' affects the training rate was well).

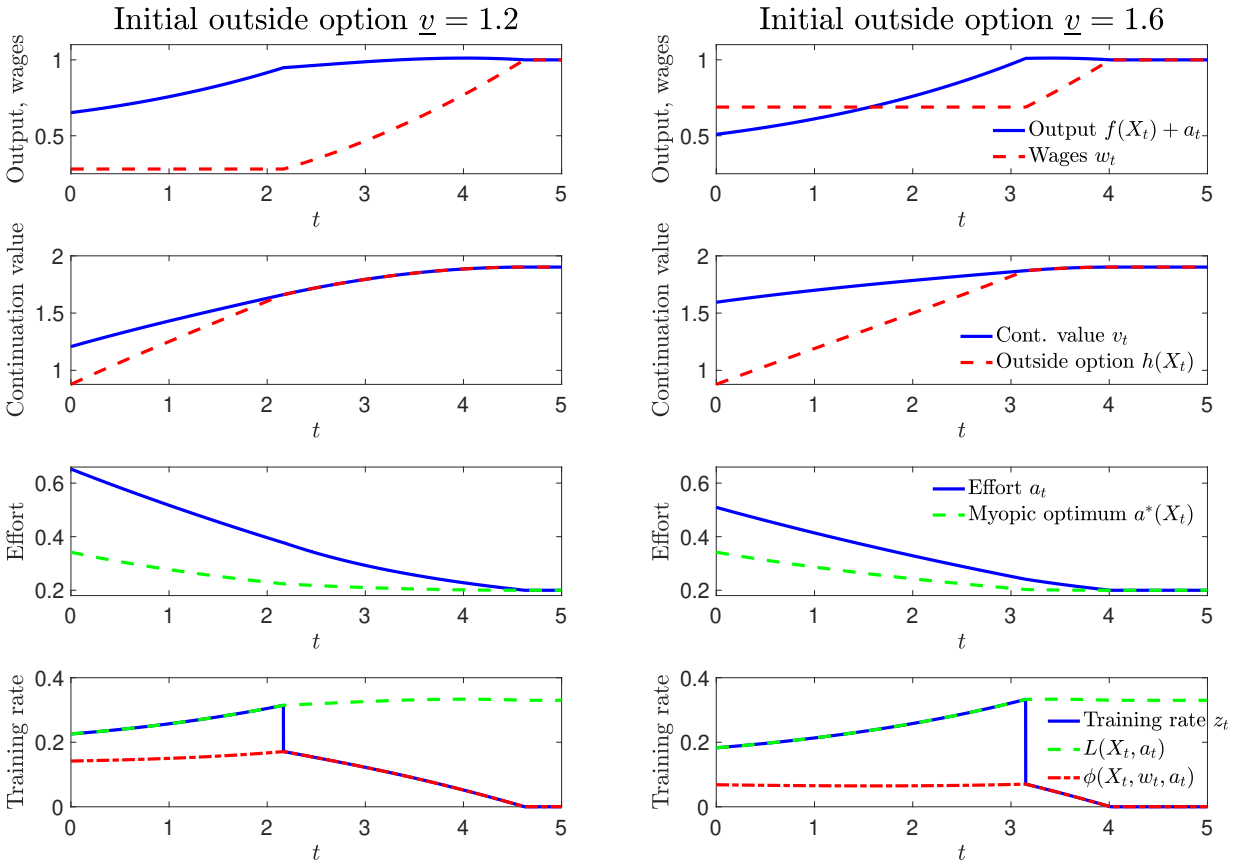

Figure 2: Illustration of optimal contract for two different initial outside options for the agent.

the training rate converges to zero. As a result, neither effort, wages, nor training have jumps at time T.

Figure 2 illustrates the trajectories of key variables for two optimal contracts, corresponding to two different levels of $\underline{v}$; each column in the figure depicts one such contract.[29] Under the lower outside option, the principal earns positive flow profits throughout the apprenticeship; under the higher one, the principal first implicitly

[29]The contract on the left is in the interior of the constrained Pareto frontier, corresponding to a point like **b** in Figure 1; the contract on the right is the most preferred by the agent, corresponding to point **c**. The contract most preferred by the principal (point **a**) has an even shorter phase 1, throughout which effort is at its upper bound. To construct this figure we assume $\underline{X} = 0$, $\overline{X} = 0.8$, $f(X) = X^{0.99}$, $L(X, a) = 0.3(0.1 + X + a)$, $u(w) = 2\sqrt{w}$, $d(a) = 2.5a^2$, $h(X) = u(f(X) + a^*(X)) - d(a^*(X))$, $\overline{a} = 1$, $\underline{c} = 0.1$, and $r = 0.2$.

lends money to the agent. The higher outside option leads to higher wages, less distorted effort, and a longer amount of time before the agent's outside option $h(X_t)$ reaches their continuation value v_t. In both contracts, the training rate drops discontinuously at the moment of transition.

In practice, apprentices might experience a discrete jump in wages (and perhaps also effort) 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.[30]

The apprenticeships our model predicts have features that are roughly in line with the real-life practices discussed in the introduction. Phase 1 in our model, for instance, represents a period of intense 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.[31] The weakly increasing wage path, with wages potentially higher than output at first, also seems to mirror the examples noted there.

While some of these practices may seem abusive to the agent, our analysis shows that they might actually be beneficial. With this in mind, we turn to the problem of optimal regulation.

[30]The training rate would also remain strictly positive at time T if, contrary to Assumption 2, the agent was more productive elsewhere once fully trained, as this puts pressure on the principal to complete their training. In this case, wages and effort would also be distorted at time T, but we can no longer say whether or not they will jump upon graduation, and in which direction, as at that time the agent will potentially switch to a different technology.

[31]In 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).

5 Regulation

Apprenticeships are a frequent target of regulation. The G20, for instance, has expressed interest in making apprenticeship programs "attractive to both employers and employees" and protecting apprentices from being underpaid and undertrained. In some cases, regulators intervene in nearly all aspects of the apprenticeship (e.g. wages, duration, curriculum, and even location), as occurs for instance in the German and Swiss dual-education models.[32] In other instances regulators seem especially concerned with specific aspects of the relationship, such as the ACGME restricting the hours of medical residents in the U.S.[33]

Caps on effort (i.e. hours worked) and floors on wages should be much easier to enforce than restrictions on the rate of knowledge transfer, as the agent's knowledge can be difficult to monitor by an external regulator. Proposition 1 shows that a planner who is able to set upper bounds on effort and lower bounds on wages is able to implement any contract on the constrained Pareto frontier (i.e. any among the family of contracts characterized in Theorem 2) subject to the players ex-ante outside options being met, without needing to control the knowledge path.

Proposition 1. *Select any contract $\mathcal{C}^*$ that is Pareto efficient (i.e. lies on the constrained Pareto frontier) and that satisfies the agent's initial participation constraint and gives the principal non-negative profits; let $(\boldsymbol{w}^*, \boldsymbol{a}^*, \boldsymbol{X}^*)$ denote the agent's lifetime wages, effort, and knowledge paths when trained under this contract. If the planner restricts the effort path to be pointwise weakly below $\boldsymbol{a}^*$ and the wage path to be pointwise weakly above $\boldsymbol{w}^*$ for the duration of the apprenticeship, then $\mathcal{C}^*$ is the unique profit-maximizing contract.*

Proof. See the Online Appendix. □

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 lower profits for the principal. Moreover, because

[32] See, e.g., "OECD Note on 'Quality Apprenticeships' for the G20 Task Force on Employment" (www.oecd.org/els/emp/OECD%20Apprenticeship%20Note%2026%20Sept.pdf, accessed 7/19/21), the German Vocational Training Act (www.gesetze-im-internet.de/bbig_2005/, accessed 7/12/20), and Wyman (2017).

[33] These restrictions include maximum weakly hours (80 on average) and limits on hours worked straight (24 in some cases). See www.acgme.org/Portals/0/PDFs/dh-ComparisonTable2003v2011.pdf (accessed 7/12/20).

the knowledge path $\boldsymbol{X}^*$ maximizes the agent's output given $\boldsymbol{a}^*$ and $\boldsymbol{w}^*$, while also meeting the learning and ongoing participation constraints, the principal will opt for that path.

Observe that the planner need not worry about capping the overall length of the contract, even though the principal might in principle be temped 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.[34]

There is empirical precedent for time-varying wage floors. For example, in formal U.S. apprenticeships, the paychecks of novices are "guaranteed to increase as their training progresses," and employers in Germany "must grant apprentices reasonable remuneration [and] remuneration increases with progressive vocational training, at least annually."[35] Time-varying caps on hours should also not be difficult to implement.

A regulator who has the power to solve the agent's commitment problem and to impose time-varying effort caps and wage floors is also able to implement the agent-optimal (first-best) contract in Theorem 1, which generates constant wages throughout the agent's lifetime and trains the agent as quickly as feasible until fully trained.

Proposition 2. *Let $\mathcal{C}^{**}$ denote the agent-optimal (first-best) contract in Theorem 1, and let $(\boldsymbol{w}^{**}, \boldsymbol{a}^{**}, \boldsymbol{X}^{**})$ denote the agent's lifetime wages, effort, and knowledge paths when trained under this contract. If 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 ongoing participation constraint, and restricts the effort path to be pointwise weakly below $\boldsymbol{a}^{**}$ and the wage path to be pointwise weakly above $\boldsymbol{w}^{**}$ for the duration of the apprenticeship, then $\mathcal{C}^{**}$is the unique profit-maximizing contract.*[36]

Proof. See the Online Appendix. □

[34] In 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.

[35] See www.dol.gov/apprenticeship/toolkit/toolkitfaq.htm (accessed 7/7/20) and www.gesetze-im-internet.de/bbig_2005/ (accessed 7/12/20).

[36] Unlike the contracts in Proposition 1, this contract lasts forever, so the agent can pay 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 $\int_T^\infty e^{-r(t-T)}[f(\overline{X}) + a_t^{**} - w_t^{**}]dt$.

In this intervention, certification requirements and non-compete clauses, which are frequently observed, should be accompanied by restrictions on both the wage and effort paths. However, if these restrictions are optimally set, there is no need to regulate the knowledge path or contract length. The intuition is similar to that of Proposition 1, save that the principal no longer needs to worry about the ongoing participation constraint.

The regulator can use the same type of intervention to implement any desired Pareto efficient contract (in the world with full commitment by both parties) so long as players earn no less than their ex-ante outside options. Such contracts are very similar to that characterized in Theorem 1, with a binding learning constraint throughout the training period and flat wages throughout the agent's lifetime.[37]

6 Conclusion

We have studied the problem of training a novice who must work as an apprentice in order to learn. To do so, we introduced the idea of learning-by-doing as an inequality constraint instead of as an equality, as this allows the master to strategically slow training down even when the agent works hard. 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." A 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 to a shorter apprenticeship without the slow-training phase.

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 simultaneously restricting 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

[37] Any such Pareto efficient contract satisfies all properties in Theorem 1 upon substituting a new (lower) wage level in the place of w^{**} (see Online Appendix, footnote 46), and can be implemented by the planner in the same manner as the agent first-best contract (see Online Appendix, footnote 47).

intrinsic ability during the apprenticeship, which would likely 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.

Landrigan, C.P., Rothschild, J.M., Cronin, J.W., Kaushal, R., Burdick, E., Katz, J.T., Lilly, C.M., Stone, P.H., Lockley, S.W., Bates, D.W. and Czeisler, C.A., 2004. Effect of reducing interns' work hours on serious medical errors in intensive care units. *New England Journal of Medicine*, 351(18), pp.1838-1848.

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 introduces 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. It determines the agent's initial payoff v_0 and the thresholds θ, and characterizes the trajectories of X_t and v_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:

$$
\begin{aligned}
S(\omega) = \max \; & r \int_0^{\overline{T}} e^{-rt} \left[f(X_t) + a_t - w_t \right] dt + \omega v_0 && (9) \\
\text{s.t. } & \dot{X}_t = z_t && (10) \\
& \dot{v}_t = r \left[v_t - u(w_t) + d(a_t) \right] && (11) \\
& v_0 \text{ free and } v_T = h(X_T) \\
& z_t \leq L(X_t, a_t) && (12) \\
& v_t \geq h(X_t) && (13) \\
& X_0 = \underline{X}, \; X_t \leq \overline{X} && (14) \\
& a_t \in [0, \overline{a}] \; , w_t \geq \underline{c}.
\end{aligned}
$$

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. Note that we multiplied the integral in the objective by r. We also fixed the horizon to be equal to $\overline{T}$. This will turn out to be without loss of generality, because after the agent's knowledge reaches $\overline{X}$, 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

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

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 $S_t = -1 + \int_t^T m(X_s)ds$ and the agent's continuation payoff v_t satisfy the system of ordinary differential equations*

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

subject to the boundary conditions $S_T = -1$ and $v_T = h(\overline{X})$.
(iii) *$\theta = \min\{t \geq 0 : S_t = -\omega \text{ or } h^{-1}(v_t) = \underline{X}\}$, and $X_t = X_\theta + \int_\theta^T \phi(X_s, w_s, a_s)ds$ with $X_T = \overline{X}$.*
(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*

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

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

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

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

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

where q_t^L, q_t^h, and $q_t^{\overline{X}}$ 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 $\overline{X}$, respectively.[38]

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^h, and $q_t^{\overline{X}}$ 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 the adjoint equation

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

$$\dot{p}_t^v = -\frac{\partial \mathcal{L}}{\partial v_t} = -p_t^v r - q_t^h r, \tag{16}$$

respectively.

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

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

(C.4) The transversality condition

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

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

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^h, and $q_t^{\overline{X}}$ 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 $(p_t^X, p_t^v, q_t^L, q_t^h, q_t^{\overline{X}})$ satisfy conditions (C.1-5).

Let us begin with (C.5). Since $f(X)$ is strictly concave, $h(X)$ 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 yields:

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

We want to show that 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. Because $w_t \geq \underline{c}$ and $a_t \leq \overline{a}$, 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 $d'(a_t) = \min\left\{d'(\overline{a}), -\left[q_t^L L_a(X_t, a_t) + e^{-rt}\right]/p_t^v\right\}$. 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 $T \leq \overline{T}$ and a $\theta \in (0, T)$, and characterize the variables $(p_t^X, p_t^v, q_t^L, q_t^h, q_t^{\overline{X}})$ such that (C.1-4) are satisfied. In our solution, (12) is slack for $t > \theta$, (13) is slack for $t < \theta$, and (14) is slack for $t < T$. Thus, the complementary slackness conditions in (C.3) can be rewritten as

$$q_t^L \begin{cases} \geq 0 & \text{if } t \leq \theta \\ = 0 & \text{if } t > \theta \end{cases}, \; q_t^h \begin{cases} = 0 & \text{if } t < \theta \\ \geq 0 & \text{if } t \geq \theta \end{cases} \text{ and } q_t^{\overline{X}} \begin{cases} = 0 & \text{if } t < T \\ \geq 0 & \text{if } t \geq T. \end{cases} \tag{18}$$

Next, we characterize the trajectories of the co-state variables by solving the corresponding adjoint equations. Solving (16) yields

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

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

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

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

and this ODE admits the following solution:

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

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 $q_t^h = 0$ for all $t < \theta$. Because $p_t^X = 0$ for all $t \in (\theta, T)$, $e^{-rs} f'(X_s) - q_s^h h'(X_s) = 0$, or equivalently,

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

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

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

$$q_t^{\overline{X}} = re^{-rt} f'(\overline{X}) - r q_t^h h'(\overline{X}).$$

Let us guess that for all $t > T$, $q_t^h = 0$ and $p_t^v = -1e^{-rt}$.[39] Then we have the following

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

expressions for $(p_t^X, p_t^v, q_t^L, q_t^h, q_t^{\overline{X}})$:

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

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

$$q_t^h = \begin{cases} 0 & \text{if } t < \theta \\ e^{-rt} m(X_t)/r & \text{if } \theta < t < T \\ 0 & \text{if } t > T \end{cases}, \text{ and } q_t^{\overline{X}} = \begin{cases} 0 & \text{if } t < T \\ re^{-rt} f'(\overline{X}) & \text{if } t > T. \end{cases} \tag{20}$$

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

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

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

Because $q_t^L \geq 0$, and $d'(0) = 0$, $d'' > 0$ and $L_{aa} \leq 0$, for $t < \theta$ there is a unique $a_t \in [0, \overline{a}]$ 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

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

So far, we have fixed an arbitrary T and $\theta < T$, and characterized the functions

$(z_t, w_t, a_t, p_t^X, p_t^v, q_t^L, q_t^h, q_t^{\overline{X}})$ such that conditions (C.1-4) are satisfied, and we argued that (C.5) is satisfied by assumption. Moreover, the agent's continuation value v_t must satisfy $v_t > h(X_t)$ for all $t < \theta$ and $v_\theta = h(X_\theta)$. (By the construction of $\phi(X, w, a)$, $v_t = h(X_t)$ 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 $v_t = h(X_t)$) 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 $v_t = h(X_t)$ on this interval and $h(X)$ is strictly increasing, it will be convenient to define the function $\xi(y) := m\left(h^{-1}(y)\right)$ and recall that $m(X) = rf'(X)/h'(X)$. 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 $d'(a_t) = \min\{d'(\overline{a}), -1/S_t\}$, respectively. During Phase 2, the trajectories of S_t and v_t satisfy the following system of ODE:

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

subject to the initial conditions $S_T = -1$ and $v_T = h(\overline{X})$. Because u, d, f, and h 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 $X_t = h^{-1}(v_t)$ 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 from (19) can equivalently be rewritten as $S_\theta \leq -\omega$ ($' =' $ if $v_0 > h(\underline{X})$). This condition implies that either (I) $S_\theta \leq -\omega$ and $v_0 = h(\underline{X})$, or (II) $S_\theta = -\omega$ and $v_0 > h(\underline{X})$. In Case (I), Phase 1 is non-existent, $v_\theta = h(\underline{X})$, and hence θ is the first time that v_t hits $h(\underline{X})$. In Case (II), Phase 1 has strictly positive duration, $v_\theta > h(\underline{X})$, and hence θ is the first time that S_t hits $-\omega$. Thus, given a solution to the system of ODE (23), we define $\theta := \min\left\{t : S_t = -\omega \text{ or } h^{-1}(v_t) = \underline{X}\right\}$. That is,

starting at T and moving backward in time, θ is the first time that S_t hits $-\omega$ or $h^{-1}(v_t) = X_t$ 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 $\tilde{t} = T - \theta$ so that $X_0 = \underline{X}$. Then Phase 2 starts at $\theta = 0$, and the agent's level of knowledge reaches $\overline{X}$ 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 $X_\theta = h^{-1}(v_\theta)$. 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 $q_t^h = q_t^{\overline{X}} = 0$ during Phase 1, we obtain the following expression for the trajectory of μ_t:

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

The trajectory of X_t, v_t, and μ_t satisfies the following system of ODE:

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

subject to the initial conditions $X_\theta = h^{-1}(v_\theta^*)$, $v_\theta = v_\theta^*$, 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 $w_0 = \max\{\underline{c},\, u'^{-1}(1/\omega)\}$, and effort a_t is implicitly defined by the equation $d'(a_t) = \min\{d'(\overline{a}),\, [L_a(X_t, a_t)\mu_t + 1]/\omega\}$.[40] 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

[40] Differentiating this expression with respect to t shows that when effort is interior, $\dot{a}_t = L_a(X_t, a_t)\dot{\mu}_t / [\omega d''(a_t) - L_{aa}(X_t, a_t)\mu_t]$. Because $L_a, \mu_t \geq 0$, $d'' > 0$ and $L_{aa} \leq 0$, for any $\omega > 0$ effort is decreasing if and only if $\dot{\mu}_t \leq 0$, which is the case whenever $r \leq L_X(X, a)$ for all X and a. (If $\omega = 0$, then $a_t = \overline{a}$ for all t.)

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 $v_0 = v_{t_0}$. Finally, because the system of ODE is autonomous, we can replace t with $\tilde{t} = t - t_0$. Then the above solution continues to satisfy (24) with $X_0 = \underline{X}$, $X_{\theta - t_0} = h^{-1}(v^*_\theta)$, $v_{\theta - t_0} = v^*_\theta$, 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, $\overline{T}$ is sufficiently large such that $\widetilde{t} < \overline{T}$.

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 $d'(a_t) = \min\{d'(\overline{a}), (1 + D_t)/\omega\}$, where $D_t := L_a(X_t, a_t)\mu_t$, and their training rate is $z_t = L(X_t, a_t)$.[41] During Phase 2, which lasts from $t = \theta$ until T, the wage, effort and training rates are $u'(w_t) = \min\{u'(\underline{c}), 1/|S_t|\}$, $d'(a_t) = \min\{d'(\overline{a}), 1/|S_t|\}$, and $z_t = \phi(X_t, w_t, a_t)$. that $S_t < 0$. After T, the agent's knowledge stays constant at $\overline{X}$, their effort satisfies $d'(a_t) = \min\{d'(\overline{a}), 1\}$, and they earn $w_t = f(\overline{X}) + a_t$, while the principal earns zero (since $f(X_t) + a_t - w_t = 0$ for all $t \geq T$). □

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. 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.

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

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), allowing us to apply their Corollary 8.2. First, we reverse time in (9) so that time "starts" at $\overline{T}$ and "ends" at 0, and hence the term $-\omega v_0$ becomes a terminal value function. Second, in every optimal solution to (9), as well as the transformed problem, $X_{\overline{T}} = \overline{X}$.[42] Therefore, we fix the "initial" values of the state variables $X_{\overline{T}} = \overline{X}$ and $v_{\overline{T}} = h(\overline{X})$, 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 C.5) is replaced by the condition that the function

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

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

[42] To see why, towards a contradiction, suppose that there exists another optimal contract C' with $X'_{\overline{T}} < \overline{X}$. Consider a modified version of the relaxed problem given in (9) where $\overline{X}$ is replaced with $X'_{\overline{T}}$; i.e. the principal is endowed with knowledge $X'_{\overline{T}}$ instead of $\overline{X}$. 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 $T < \overline{T}$, $X_T = X'_{\overline{T}}$ 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 $\overline{a}$, and is trained at the zero-rent rate $\phi(X_t, \underline{c}, \overline{a})$. This modified contract is feasible for (9) and because $f(X_t) + \overline{a} > \underline{c}$ 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.

purpose, we write the Lagrangian for the static problem:

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

where $\kappa \geq 0$ is a dual multiplier. This problem is convex, and for any κ, the optimal controls satisfy[43]

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

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. The Hessian of $\mathcal{H}^0(X, v, p^X, p^v, t)$,

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

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, so the trajectory of 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, 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 $w_t = u^{-1}\left(k_t + d(a_t)\right)$ is also unique on $[0, \theta)$. Next, consider $t > \theta$. Any optimal solution must satisfy the first-order conditions $u'(w_t) =$

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

$\min\{u'(\underline{c}),\, -e^{-rt}/p_t^v\}$ and $d'(a_t) = \min\{d'(\overline{a}),\, -[q_t^L L_a(X_t, a_t) + e^{-rt}]/p_t^v\}$ 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 $k_t = u(\underline{c}) - d(\overline{a})$, 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 $k_t > u(\underline{c}) - d(\overline{a})$. Therefore, w_t and a_t are also unique for such t. □

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 $z_t = \phi(X_t, w_t, a_t) = \dot{v}_t / h'(X_t)$, where we have used the fact that for such t, the ongoing participation constraint binds so $v_t = h(X_t)$, and that $\dot{v}_t = r\,[v_t - u(w_t) + d(a_t)]$. Since $h'(X) > 0$ 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 $u'(w_t) = \min\{u'(\underline{c}),\, -1/S_t\}$, $d'(a_t) = \min\{d'(\overline{a}),\, -1/S_t\}$, 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 $X_{\overline{T}} = \overline{X}$ and $X_t \leq \overline{X}$ for all t. Therefore, we conclude that such t' cannot exist, and hence $z_t \propto \dot{v}_t \geq 0$ for all t. □

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,

$$\omega' \left[v_0^*(\omega') - v_0^*(\omega) \right] > \pi_0^*(\omega) - \pi_0^*(\omega') > \omega \left[v_0^*(\omega') - v_0^*(\omega) \right], \tag{25}$$

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 $v_0^*(\infty) := \lim_{\omega \to \infty} v_0^*(\omega) > \underline{v}$. Note first that as $\omega \to \infty$, 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 $v_0^*(\infty) := \lim_{\omega \to \infty} v_0^*(\omega) \leq \underline{v}$. 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 $\omega^* = \inf \{\omega \geq 0 \, : \, v_0^*(\omega) \geq \underline{v}\}$. 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.

Note also that $\omega^* < 1$, and hence Phase 2 has strictly positive length. Otherwise, the agent would earn no less than steady state wages throughout the apprenticeship, which per Assumption 5, would create losses for the principal. By Assumption 6.2, however, there is a contract that allows the principal to make a non-negative profit.

Finally, we show that as the agent's initial outside option $\underline{v}$ increases, the length of Phase 1 increases and the length of Phase 2 decreases, each strictly so if the initial participation constraint binds. Note that if this constraint is slack, the length of each phase is independent of $\underline{v}$, and so we shall restrict attention to the case in which it binds.

From (23) and the fact that $S_\theta = -\omega$ that the length of Phase 2, $T - \theta$, strictly decreases in ω. Since $v_0^*(\omega)$ strictly increases in ω and $v_0^*(\omega^*) = \underline{v}$, it follows that ω^* strictly increases in $\underline{v}$, and hence $T - \theta$ strictly decreases in $\underline{v}$.

To establish that phase 1 becomes strictly longer, fix two initial outside options $\underline{v}, \underline{v}'$ such that $\underline{v} > \underline{v}'$, and denote the associated contracts by C and C', respectively. Assume the initial participation binds under both contracts. From (23) and because $\omega^* > \omega^{*\prime}$, we have $X_\theta > X'_{\theta'}$ and $|S_\theta| > |S'_{\theta'}|$. Now suppose towards a contradiction that $\theta \leq \theta'$; i.e., the length of Phase 1 when the initial outside option is $\underline{v}'$ is at least as large as when it is $\underline{v}$.

We claim that there is a time $\tilde{t} < \theta$ such that $a_{\tilde{t}} > a'_{\tilde{t}}$ and $X_t \geq X'_t$ for all $t \geq \tilde{t}$. To see why this must be the case, define s to be the largest time before θ such that $X_s = X'_s$. Note $s < \theta$ and by construction $X_t > X'_t$ for all $s < t \leq \theta$. If the claim is not true, then it must be that $a_t \leq a'_t$ for all $s \leq t \leq \theta$. But then the learning constraint implies that $X_\theta \leq X'_\theta$, a contradiction.

Next, since f' is decreasing in X, L_a is weakly decreasing in a, and L_X is weakly decreasing in X, the expression for D_t (given in footnote 23) implies that $D_{\tilde{t}} < D'_{\tilde{t}}$. Because $|S_\theta| > |S'_{\theta'}|$, it follows from the expression for a_t that $a_{\tilde{t}} \leq a'_{\tilde{t}}$, contradicting the claim established above. This completes the proof.

Working to Learn: Online Appendix

Proof of 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:

$$S(\beta) = \max \int_0^\tau e^{-rt} \left[u(w_t) - d(a_t)\right] dt + \beta \int_0^\tau e^{-rt} \left[f(X_t) + a_t - w_t\right] dt \tag{26}$$

$$\begin{aligned} \text{s.t. } & \dot{X}_t = z_t \\ & z_t \leq L(X_t, a_t) \\ & X_0 = \underline{X},\ X_t \leq \overline{X},\ X_\tau \text{ free} \\ & a_t \in [0, \overline{a}]. \end{aligned} \tag{27}$$

for some finite but large τ.[44] 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 (26).

The following lemma characterizes the optimal solution for this problem.

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 $d'(a_t) = \min\{d'(\overline{a}),\ \beta[1 + \mu_t L_a(X_t, a_t)]\}$, and*

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

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

(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 $X_0 = \underline{X}$, $X_T = \overline{X}$, and $\mu_T = 0$.
(iii) *For $t > T$, $X_t = \overline{X}$ and $\mu_t = 0$.*

Proof of Lemma 4. Define the Hamiltonian

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

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

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

where q_t^L and $q_t^{\overline{X}}$ are the multipliers associated with the agent's learning constraint and the constraint that their knowledge level X_t does not exceed $\overline{X}$, 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 (26) if there exists a piecewise continuously differentiable function p_t^X, and piecewise continuous functions q_t^L and $q_t^{\overline{X}}$ 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

$$\dot{p}_t^X = -\frac{d\mathcal{L}}{dX} = -\beta e^{-rt} f'(X_t) - q_t^L L_X(X_t, a_t) + q_t^{\overline{X}}. \tag{28}$$

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

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

(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 $q_t^{\overline{X}}$ 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 $(p_t^X, q_t^L, q_t^{\overline{X}})$ 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:

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

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 $a_t \leq \overline{a}$, 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 $d'(a_t) = \min\left\{d'(\overline{a}),\, q_t^L e^{rt} L_a(X_t, a_t) + \beta\right\}$. 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 $(p_t^X, q_t^L, q_t^{\overline{X}})$ 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 $X_t \leq \overline{X}$ binds if and only if $t \geq T$. Therefore, by the complementary slackness conditions in (C.3), the following must be true:

$$q_t^L \begin{cases} \geq 0 & \text{if } t < T \\ = 0 & \text{if } t > T \end{cases} \quad \text{and} \quad q_t^{\overline{X}} \begin{cases} = 0 & \text{if } t < T \\ \geq 0 & \text{if } t > T. \end{cases} \tag{29}$$

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 $q_t^{\overline{X}} = 0$, (28) 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 (29) 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 (28) that

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

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

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

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

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

$$u'(w_t) = \beta, \text{ and} \tag{30}$$

$$d'(a_t) = \min\left\{d'(\overline{a}),\, \beta\left[1 + \mu_t L_a(X_t, a_t)\right]\right\}. \tag{31}$$

Finally, because the learning constraint binds for $t < T$, while the constraint that $X_t \leq \overline{X}$ 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 $(z_t, w_t, a_t, p_t^X, q_t^L, q_t^{\overline{X}})$ 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 $X_T = \overline{X}$. 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:

$$\begin{bmatrix} \dot{X}_t \\ \dot{\mu}_t \end{bmatrix} = F(X_t, \mu_t) := \begin{bmatrix} L(X_t, a_t) \\ -f'(X_t) + \left[r - L_X(X_t, a_t)\right]\mu_t \end{bmatrix}, \tag{32}$$

where a_t is the unique solution of (31), 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 $X_T = \overline{X}$ 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 $\widetilde{t} = t - t_0$ so that for $\widetilde{t} = 0$, $X_{\widetilde{t}} = \underline{X}$. Thus, the training phase begins at $\tilde{t} = 0$ and ends that $\tilde{t} = T - t_0$.[45] During this interval, the agent's learning constraint binds,

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

and the trajectories of X_t and μ_t satisfy (32). The agent's wage and effort satisfies (30) and (31), 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 (26) 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

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

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 conditions, (30) and (31), 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 (31) observe that a_t is uniquely determined by X_t and μ_t, and hence its trajectory is also unique. □

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 $d'(a_t) = \min\{d'(\overline{a}),\ \beta\}$. 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 (26) is identical. Therefore, this solution also uniquely solves (26) when $\tau = \infty$.

To complete the proof of Theorem 1, we now show that for an appropriately chosen $\beta \geq 0$, the solution of (26) solves the original problem, (I). For each β (and $\tau = \infty$), denote the solution of (26) 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

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

Therefore,

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

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 $\beta \to 0$. Because u is strictly increasing by assumption, $\lim_{\beta \to 0} w_t^\beta = \infty$, while the corresponding effort remains bounded (since $a_t^\beta \leq \overline{a}$). Therefore, the agent's credit balance $\lim_{\beta \to 0} Y(\beta) = -\infty$. On other other hand, for β sufficiently large, $w_t^\beta \simeq 0$ and $a_t^\beta = \overline{a}$ for all t. In this case, $Y(\beta) \simeq \int_0^\infty e^{-rt}\left[f(X_t) + \overline{a}\right] dt > 0$ where $\dot{X}_t = L(X_t, \overline{a})\mathbb{I}_{\{X_t < \overline{X}\}}$. 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 (26) with $\beta = \beta^*$ solves the original problem, (I). Towards a contradiction, suppose there exists another solution with four-tuple $(\widetilde{X}_t, \widetilde{a}_t, \widetilde{w}_t, \widetilde{z}_t)$ such that the agent's payoff $\widetilde{V} \geq V(\beta^*)$ and (4) is satisfied. But then this implies that

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

which contradicts the fact that $(X_t^{\beta^*}, a_t^{\beta^*}, w_t^{\beta^*}, z_t^{\beta^*})$ uniquely solves (26) when $\beta =$

β^*.[46] □

Proof of Proposition 1. We will show that $\mathcal{C}^*$ is the unique solution to the problem that is identical to (II) with the addition of the planner's constraints, that is,

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

Let v_0^* denote the agent's initial payoff under the contract $\mathcal{C}^*$. By Theorem 2, $\mathcal{C}^*$ 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 $\mathcal{C}'$ that solves (III), yet it does not coincide with $\mathcal{C}^*$ almost everywhere. If $\mathcal{C}'$ prescribes a wage or effort path that differs from $\boldsymbol{w}^*$ or $\boldsymbol{a}^*$, 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 $\boldsymbol{X}^*$. 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

[46] Because $Y(\beta)$ is strictly increasing and $V(\beta)$ is strictly decreasing in the multiplier β, by varying this multiplier from β^* to some $\overline{\beta}$, it is possible to trace the entire Pareto frontier (in the absence of commitment problems) subject to the agent's initial participation constraint being satisfied and the principal not making losses. It follows from Lemma 4 that for any such β, the optimal contract has the same structure as that in Theorem 1, except for the (constant) wage w_t that is different. (Note that $\widetilde{D}_t = \mu_t L_a(X_t, a_t)$ and the constant wage satisfies $u'(w_t) = \beta$.)

Theorem 2, $\boldsymbol{X}^*$ 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. □

Proof of Proposition 2. We will show that the contract $\mathcal{C}^{**}$ uniquely solves

$$\begin{aligned}
\max_{\mathcal{C}} \quad & \int_0^\infty e^{-rt} \left[f\left(X_t\right) + a_t - w_t \right] dt && \text{(IV)} \\
\text{s.t.} \quad & z_t \leq L(X_t, a_t) \\
& v_0 = \int_0^\infty e^{-rt} \left[u(w_t) - d(a_t) \right] dt \geq \underline{v} && (33) \\
& z_t \geq 0, \ X_0 = \underline{X}, \ X_t \leq \overline{X} \\
& a_t \in [0, \overline{a}], \ w_t \geq \underline{c} \\
& a_t \leq a_t^{**} \quad \text{and} \quad w_t \geq w_t^{**} \text{ for all } t. && (34)
\end{aligned}$$

By construction, $\mathcal{C}^{**}$ is feasible for (IV). Let v_0^{**} and Y_0^{**} denote the agent's initial payoff and the principal's profits under this contract. Because the planner's constraints (34) bind under $\mathcal{C}^{**}$ for all t, any effort and wage path that satisfies (34) 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 (34) 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:

$$\begin{aligned}
S'(\gamma) = \max \quad & \int_0^\infty e^{-rt} \left[f\left(X_t\right) + a_t - w_t \right] dt + \gamma \int_0^\infty e^{-rt} \left[u(w_t) - d(a_t) \right] dt && (35) \\
\text{s.t.} \quad & \dot{X}_t = z_t \\
& z_t \leq L(X_t, a_t) \\
& X_0 = \underline{X}, \ X_t \leq \overline{X} \\
& w_t \geq \underline{c}, \ a_t \in [0, \overline{a}].
\end{aligned}$$

This problem is identical to (26) 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 $\mathbf{X}^\gamma$, $\mathbf{a}^\gamma$ and $\mathbf{w}^\gamma$ denote the knowledge, effort, and wage path that solves this problem, respectively, and define the corresponding principal's payoff $Y(\gamma) := \int_0^\infty e^{-rt}\left[f\left(X_t^\gamma\right) + a_t^\gamma - w_t^\gamma\right] dt$ and agent's payoff $V(\gamma) := \int_0^\infty e^{-rt}\left[u(w_t^\gamma) - d(a_t^\gamma)\right] dt$. 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 $Y(\gamma^*) = Y_0^{**}$. 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 (34) and the agent's initial participation constraint $v_0 \geq \underline{v}$ with the constraint $v_0 \geq v_0^{**}$. Suppose instead there is another solution with four-tuple $(\widetilde{X}_t, \widetilde{a}_t, \widetilde{w}_t, \widetilde{z}_t)$ such that the principal's payoff $\widetilde{Y} \geq Y(\gamma^*)$ and the constraint $\widetilde{v}_0 \geq v_0^{**}$ is satisfied. But this implies that

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

contradicting the fact that $\mathcal{C}^{**}$ uniquely solves (35) when $\gamma = \gamma^*$.[47] □

[47] In the same way, the regulator can implement any Pareto efficient contract that gives the agent initial utility $\widehat{v}_0$ for any $\widehat{v}_0 \in [\underline{v}, v_0^{**}]$. This follows from footnote 46 and the fact that $Y(\gamma)$ is strictly decreasing in γ, whereas $V(\gamma)$ is strictly increasing in γ.