Curriculum design in the age of AI

Benjamin Davies*

Draft version: July 20, 2026

Click here for latest version

Abstract

I develop a model of learning-by-doing and curriculum design, and use it to study the impact of artificial intelligence (AI). A myopic student faces a sequence of tasks that he can work on or delegate to AI. Work requires costly effort but builds skill; delegation requires no effort but builds no skill. A teacher designs a task sequence ("curriculum") that maximizes the student's skill development given his choices to work or delegate. Without AI, the teacher makes earlier tasks more effort-intensive and later tasks more skill-intensive. With AI, the teacher must distort the curriculum to incentivize effort, leading to less skill development. If AI complements effort, then improvements in AI quality make high-skill students learn faster but low-skill students learn slower.

JEL classification: C61, D24, I21, J24

Keywords: AI, curriculum design, delegation, education, human capital, skill

*Department of Economics, Stanford University; bldavies@stanford.edu. I thank seminar participants at Stanford for helpful comments.

1 Introduction

Using artificial intelligence (AI) can boost productivity (Noy and Zhang, 2023; Brynjolfsson et al., 2025; Dillon et al., 2026), but can also inhibit learning and cognition (Bastani et al., 2025; Kosmyna et al., 2025; Liu et al., 2026; Shen and Tamkin, 2026). In particular, using AI can "short-circuit the productive struggle essential for learning" (Poulidis et al., 2025, p. 2).

At the same time, AI use in schools is widespread (Doss et al., 2025; Campos and Singleton, 2026). Students use AI to augment their learning and automate their coursework, and view AI as beneficial (Contractor and Reyes, 2026). However, AI allows students to delegate coursework that was intended as an opportunity to develop skill. Chirikov (2026) offers empirical evidence of such delegation, while Bastani et al. (2025) and Strömberg et al. (2026) offer evidence that students who delegate develop less skill.[1]

These empirical results raise an important question: how should curricula be designed to ensure students develop skill when they can delegate to AI?

To answer this question, I develop a model of learning-by-doing and curriculum design. A myopic student (he) faces a sequence of tasks that he can work on or delegate to AI. Working requires costly effort but builds skill. Delegation does not require effort but does not build skill. A teacher (she) designs a task sequence (i.e., a "curriculum") that maximizes the student's overall skill development given his best-responses (i.e., his choices to work or delegate).

Working on a task involves the combination of effort and skill to produce an output. I interpret "effort" as the cognitive resources spent applying a production method and "skill" as the ability to identify suitable methods. Effort is a costly productive input but skill is costless. Moreover, effort and skill are complements because a student with high skill can allocate effort more efficiently.

I characterize tasks by their "effort intensity": the extent to which production relies on effort vis-à-vis skill. Effort-intensive tasks require the student to apply a given method (e.g., using the power rule to differentiate a polynomial function), while skill-intensive tasks require him to identify an appropriate method (e.g., building a tractable economic model).

First, I consider the "first-best" case in which the student prefers to work on every task. Then the teacher's optimal curriculum comprises tasks that become gradually less effort-intensive and more skill-intensive (Theorem 1). This is because the teacher maximizes the student's skill development by maximizing his effort. On earlier tasks, the student has less skill and so the teacher incentivizes effort by making it contribute more to output. On later tasks, the student has more skill, which makes effort more productive because effort and skill are complementary inputs. The teacher capitalizes on this complementarity by making tasks less effort-intensive and more skill-intensive.

[1] AI (mis)use also threatens to invalidate educational assessments (Chirikov et al., 2026): if students delegate tasks to AI, then performance on those tasks cannot be used to infer students' ability.

Second, I consider the "second-best" case in which the student may prefer to delegate a task. Then the first-best curriculum may not be incentive-compatible and the teacher must distort it to incentivize effort. She does so by making earlier tasks more skill-intensive than the first-best curriculum prescribes. As a result, the second-best curriculum comprises tasks with increasing-then-decreasing effort intensities (Theorem 2). This ensures the student develops some skill, but he develops less than under the first-best curriculum. The student gains more skill overall when he starts with more skill and earns a lower payoff from delegating (Theorem 3). Intuitively, skillful students can produce more on their own, are less tempted to delegate, and so can be incentivized to work with less curriculum distortion. In contrast, making delegation more attractive tightens the incentive constraint on the second-best curriculum, which the teacher must distort further to incentivize effort.

Finally, I extend my model to allow for the possibility that AI complements effort.[2] I capture this possibility by introducing an "AI quality" parameter that increases the marginal product of effort *and* the payoff from full delegation. Improvements in AI quality make high-skill students gain skill faster but low-skill students gain skill slower (Theorem 4). This is because high-skill students become more willing to work while low-skill students become more tempted to delegate.

After discussing related literature, the paper is organized as follows. Section 2 describes the model. Section 3 characterizes the student's best-responses. Section 4 characterizes the teacher's optimal curricula. Section 5 considers the possibility that AI complements effort. Section 6 concludes. Appendix A contains proofs of my mathematical claims.

1.1 Related literature

This paper connects to the literature on human capital formation. Schultz (1960) frames education as investment in human capital. Becker (1962) and Mincer (1962) present formal theories of such investment, while Ben-Porath (1967) studies human capital investment across the life cycle. Heckman (2000) and Cunha and Heckman (2007) emphasize the importance of early investment, and argue it promotes later investment due to dynamic complementarities ("skill begets skill"). Such complementarities arise in my model: exerting effort builds skill, which raises the future productivity of effort and leads to more skill development.

Similar complementarities arise in the economic literature on "learning-by-doing," surveyed by Thompson (2010). Its central idea is that knowledge and skill accumulate via experience, which can take the form of investment (Arrow, 1962), production (Rosen, 1972; Brueckner and Raymon, 1983), or technology use (Jovanovic and Nyarko, 1996). Likewise, in my model, skill accumulates via effortful production. However, I depart from the learning-by-doing literature by introducing

[2] For example, the student could use AI to perform effort-intensive subtasks he has used skill to identify, rather than delegate entire tasks and let the AI identify subtasks on its own.

"doing-without-learning" as an outside option.

The option to use AI appears in recent theoretical work on AI and skill formation. Fudenberg and Levine (2026) show that early access to AI can weaken the incentive to work on tasks that build and maintain expertise. Fudenberg and Levine's agent chooses between exogenous tasks; in contrast, I treat tasks as endogenous and study how a planner should design them. Garicano and Rayo (2026) study AI's impact on apprenticeships, which they model as scenarios in which "juniors pay for training by doing menial work" that AI performs increasingly well. In Garicano and Rayo's model, the planner is a master who controls the pace of knowledge transfer, and the agent's outside option is to quit and use AI independently. In contrast, in my model, the planner is a teacher who designs the task sequence and the agent's outside option is to delegate within a task. Huang and Vishnoi (2026) study the conditions under which repeated delegation makes skill collapse to a stable low equilibrium; I allow a designer who can prevent such collapses. Finally, Peterson (2025) studies an educational planner who misreads the labor-market return to easily taught skills. Peterson's friction is informational, whereas mine is incentive-based: the teacher in my model must prevent delegation.

A parallel strand studies incentive conflicts around AI in firms. Bastani and Cachon (2026) study a principal-agent model where the agent can shirk by not verifying AI output. Bastani and Cachon show that motivating human vigilance can become prohibitively costly, leading the principal to prohibit AI use despite its productivity benefits. Siderius et al. (2026) show that the optimal compensation for verifying AI output can be non-monotone in AI quality. Likewise, I show that the optimal curriculum designed in response to AI is non-monotone in AI quality (see Section 5.2). Caosun and Aral (2026) consider a manager who chooses how much a worker can use AI, trading off productivity gains and skill losses. Managerial incentives can lead to "augmentation traps" in which the worker uses AI but ends up worse off than if they had never used AI. In contrast, the "manager" (i.e., teacher) I study only wants the "worker" (i.e., student) to build skill and must design incentives around the worker's temptation to delegate.

My focus on incentives echoes the theoretical literature on educational standards. Becker and Rosen (1992), Costrell (1994), and Betts (1998) view education as a principal-agent problem: teachers set standards or grading policies to motivate effort from students. Adilov and Cline (2026) study how the possibility of AI-enabled cheating impacts optimal grading policies, showing that it precludes teachers from using high grade boundaries to motivate effort. Adilov and Cline also show that effort falls as AI quality rises. In contrast, I show effort can rise or fall, depending on whether the student uses AI as a complement or substitute for effort.

A separate literature treats how and what students learn as an optimization problem. Lazear (2001) studies optimal class sizes when students can disrupt their classmates' learning. Ellison and Pathak (2025) model curriculum design as the allocation of teaching time across skills. My model shares Ellison and Pathak's premise that curricula can be chosen optimally, but I focus on a different margin (effort intensity, rather than skill coverage), and introduce endogenous student effort and skill dynamics. Such dynamics are also part of Shaikh's (2025) model of curriculum design. Shaikh shows that if course content is sufficiently cumulative, then teachers should put more grading weight on earlier tasks so that students are motivated to build foundational skills. My analysis complements Shaikh's analysis by endogenizing the task sequence and allowing students to delegate to AI.

Recent empirical work motivates my focus on the delegation threat. Chirikov (2026) offers empirical evidence that students delegate to AI. Bastani et al. (2025), Shen and Tamkin (2026), and Strömberg et al. (2026) offer evidence that students and workers who delegate build less skill; Kosmyna et al. (2025), Liu et al. (2026), and Poulidis et al. (2025) document reduced cognitive engagement and weaker independent performance. Chirikov et al. (2026) argue that AI (mis)use calls for assessment reform, such as "requiring students to document their process [and] demonstrate understanding" (Chirikov et al., 2026, p. 820). Indeed, the teacher in my model prevents delegation by making tasks more skill-intensive (see Theorem 2), which makes understanding the task domain necessary for high performance.

2 Model

There is a myopic student (he) and a forward-looking teacher (she). The student faces a sequence of tasks $t \in \mathcal{T} \equiv \{0, \ldots, T-1\}$ with $T \geq 2$. He can work on a task or delegate it to an AI assistant. Work requires costly effort but builds skill. Delegation requires no effort but builds no skill. The teacher designs tasks that maximize the student's skill development given his best-responses.

Working. Working on a task involves the combination of effort and skill to produce an output. I interpret "effort" as the cognitive resources spent applying a production method, and "skill" as the ability to identify suitable methods and transfer them across tasks.[3] Effort is a costly productive input but skill is costless.[4] Moreover, effort and skill are complements because a student with high skill can allocate effort more efficiently.

I characterize tasks by their "effort intensity": the extent to which production relies on effort vis-à-vis skill. Effort-intensive tasks require the student to apply a given method, while skill-intensive tasks require him to identify an appropriate method.

[3]For example, a student could transfer methods across tasks by recognizing those tasks as instances of a common class. This would align with Anzai and Simon's (1979, p. 124) discussion of problem-solving strategies: "Since the most efficient strategies for a particular task may not be the most obvious ones to someone encountering the task for the first time ... Initially, the solver might hit on one of the 'obvious' strategies, and then gradually progress to more efficient ones with increasing familiarity with the problem domain."

[4]Intuitively, the student either knows how to approach a problem efficiently or does not, depending on his familiarity with the problem domain (see Footnote 3).

If the student works on task t, then he chooses the effort $x_t \geq 0$ that maximizes

$$u(x_t, s_t, \eta_t) \equiv \pi x_t^{\eta_t} s_t^{1-\eta_t} - \frac{x_t^2}{2} \tag{1}$$

given his overall productivity $\pi > 0$, his skill level $s_t > 0$, and the task's effort intensity $\eta_t \in (0,1]$. The first term on the right-hand side of (1) captures the output produced by combining effort x_t and skill s_t. The effort intensity η_t determines the marginal product of effort vis-à-vis skill.[5] The second term on the right-hand side of (1) captures the cost of exerting effort. The parameter π indexes the student's overall productivity relative to the cost of exerting effort.

Delegation. If the student delegates task t, then he exerts zero effort and receives payoff $\overline{u} \geq 0$. This payoff measures the (expected) output generated by an AI assistant given the entire task as an input (e.g., via the student prompting an AI chatbot with "here is my homework, give me the answers.").[6]

Best-responses. The student's best-response effort

$$x^*(s_t, \eta_t) \in \arg\max_{x \geq 0} u(x, s_t, \eta_t)$$

yields payoff

$$u^*(s_t, \eta_t) \equiv \max_{x \geq 0} u(x, s_t, \eta_t).$$

He works on task t if $u^*(s_t, \eta_t) \geq \overline{u}$ and delegates task t otherwise.

Skill. The student has initial skill $s_0 > 0$. On each task $t \in \mathcal{T}$, he gains skill $\Delta(s_t, \eta_t) \equiv s_{t+1} - s_t$ proportional to the effort he exerts:

$$\Delta(s_t, \eta_t) \equiv \begin{cases} r\, x^*(s_t, \eta_t) & \text{if student works} \\ 0 & \text{if student delegates} \end{cases} \tag{2}$$

with $r > 0$ a fixed skill accumulation rate.

[5] If $\eta_t = 0$, then output πs_t is independent of effort x_t and the student optimally chooses $x_t = 0$. But then he builds zero skill, which is never optimal for the teacher when she solves her problem (P). Thus, it is without loss to bound η_t away from zero.

[6] I say "expected" because (i) the student may know only the AI's typical performance over a range of tasks, rather than its exact performance on a specific task, and (ii) the AI output may contain random errors (e.g., due to "hallucination").

Teacher's problem. The teacher chooses effort intensities $(\eta_t)_{t\in\mathcal{T}}$ that maximize the student's overall skill gain

$$s_T - s_0 = \sum_{t\in\mathcal{T}} \Delta(s_t, \eta_t)$$

given his best-responses.[7] Formally, she solves

$$\max_{(\eta_t)_{t\in\mathcal{T}}} \; (s_T - s_0) \;\text{ subject to (2) and } 0 < \eta_t \leq 1 \text{ for each } t \in \mathcal{T}. \tag{P}$$

I characterize the solution to (P) in Section 4.[8]

3 Best-responses

Suppose the student has skill $s > 0$ and faces a task with effort intensity $\eta \in (0,1]$. If he works on the task, then he exerts best-response effort $x^*(s,\eta)$ and yields payoff $u^*(s,\eta)$. I characterize this effort and payoff in Lemmas 1 and 2.

Lemma 1 (Best-response effort)**.** *Given a skill level $s > 0$ and effort intensity $\eta \in (0,1]$, the student exerts best-response effort*

$$x^*(s,\eta) = s\left(\frac{\pi\eta}{s}\right)^{1/(2-\eta)}. \tag{3}$$

This effort is (i) positive, (ii) increasing in the productivity parameter π, (iii) increasing in s when $\eta < 1$ and constant in s when $\eta = 1$, and (iv) increasing in η if

$$\eta \exp\left(\frac{2}{\eta} - 1\right) > \frac{s}{\pi} \tag{4}$$

and non-increasing in η otherwise.

Lemma 2 (Best-response payoff)**.** *Given a skill level $s > 0$ and effort intensity $\eta \in (0,1]$, the student's best-response effort $x^*(s,\eta)$ yields payoff*

$$u^*(s,\eta) = \frac{1}{\eta}\left(1 - \frac{\eta}{2}\right)(x^*(s,\eta))^2. \tag{5}$$

This payoff is (i) positive, (ii) increasing in the productivity parameter π, (iii) increasing in s when $\eta < 1$ and constant in s when $\eta = 1$, and (iv) decreasing in η when $\eta < s/\pi$ and increasing in η otherwise.

[7] Thus, the teacher maximizes the curriculum's "value-added." This is consistent with the empirical literature that uses value-added to evaluate teachers (see, e.g., Bacher-Hicks and Koedel, 2023).

[8] A solution to (P) exists for all parameter values because choosing $\eta_t = 1$ for each $t \in \mathcal{T}$ is always feasible. Theorem 2 implies the solution is unique when (SB) and (ND) hold. If they do not hold, then (P) can have many solutions because, for example, the student could have final skill $s_T = s_0$ regardless of the chosen sequence $(\eta_t)_{t\in\mathcal{T}}$.

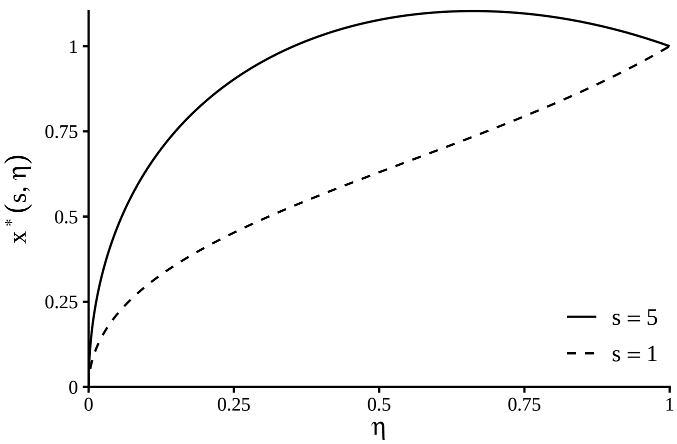

Figure 1: Best-response efforts $x^*(s, \eta)$ when $\pi = 1$ and $s \in \{1, 5\}$.

I prove Lemmas 1 and 2, and all subsequent results, in Appendix A.

The best-response effort $x^*(s, \eta)$ and $u^*(s, \eta)$ are both increasing in the productivity parameter π. Raising π makes the student more productive relative to the cost of exerting effort, making him more willing to exert effort because doing so yields a higher payoff.

If a task is fully effort-intensive (i.e., if $\eta = 1$), then the student's skill when he faces the task is irrelevant, and he simply trades off the marginal product and cost of effort. Indeed, if $\eta = 1$, then his best-response effort $x^*(s, 1) = \pi$ and payoff $u^*(s, 1) = \pi^2/2$ do not depend on his skill level s. However, if a task is *not* fully effort-intensive (i.e., if $\eta < 1$), then having more skill makes the marginal product of effort higher and makes the student more willing to exert effort. Indeed, if $\eta < 1$, then $x^*(s, \eta)$ and $u^*(s, \eta)$ are both increasing in s.

Figure 1 shows how $x^*(s, \eta)$ depends on η when $\pi = 1$ and $s \in \{1, 5\}$. If $s = 1$, then the condition (4) holds for all $\eta \in (0, 1]$ and so Lemma 1 implies $x^*(1, \eta)$ is increasing in η. In contrast, if $s = 5$, then (4) holds precisely when η is small and so $x^*(5, \eta)$ has an inverted-U-shaped relationship with η. In general, for all $s > 0$, there is a unique effort intensity that maximizes $x^*(s, \eta)$. I characterize this maximizer in Section 4.1.

4 Optimal curricula

This section characterizes the effort intensities $(\eta_t)_{t \in \mathcal{T}}$ that solve the teacher's problem (P). She wants to maximize the sum of skill increments $\Delta(s_t, \eta_t) \equiv s_{t+1} - s_t$ defined by (2). Each increment depends on whether the student works or delegates, and, if he works, how much effort he exerts as a best-response to his skill level and the task's effort intensity.

The student works on task $t \in \mathcal{T}$ if and only if his best-response payoff $u^*(s_t, \eta_t)$ exceeds the delegation payoff $\overline{u}$:

$$u^*(s_t, \eta_t) \geq \overline{u}. \tag{IC}$$

By Lemmas 1 and 2, the best-response effort $x^*(s_t, \eta_t)$ and payoff $u^*(s_t, \eta_t)$ are non-decreasing in the student's skill level s_t when he faces task t. It follows that the skill increment

$$\Delta(s_t, \eta_t) = \begin{cases} r\, x^*(s_t, \eta_t) & \text{if (IC) holds} \\ 0 & \text{otherwise} \end{cases}$$

is also non-decreasing in s_t. As a result, the teacher cannot benefit from allowing the student to gain less skill from a task than is feasible. Thus, for each task $t \in \mathcal{T}$, she chooses an effort intensity $\eta_t \in (0, 1]$ that maximizes the student's best-response effort $x^*(s_t, \eta_t)$ subject to the incentive constraint (IC):

$$\eta_t \in \underset{\eta \in (0,1]}{\arg\max}\ x^*(s_t, \eta) \ \text{ subject to } \ u^*(s_t, \eta) \geq \overline{u}. \tag{6}$$

I analyze the relationship between η_t and s_t below. Section 4.1 considers the "first-best" case in which $\overline{u} = 0$ and so (IC) holds for all $\eta_t \in (0, 1]$. Section 4.2 considers the "second-best" case in which $\overline{u} > 0$ and so (IC) may not hold. Finally, Section 4.3 analyzes the student's overall skill gain $(s_T - s_0)$ when the teacher designs a second-best curriculum.

4.1 First-best curricula

Consider the first-best case in which the delegation payoff $\overline{u}$ equals zero. In this case, the student prefers to work on every task regardless of its effort intensity.[9] As a result, the teacher faces no incentive constraints when she chooses tasks' effort intensities; she can simply choose them to maximize the student's best-response efforts.

4.1.1 Optimal effort intensities

Consider the condition (4) that determines whether the best-response effort $x^*(s, \eta)$ is increasing in the effort intensity η. The left-hand side of (4) is continuously decreasing in η over $(0, 1]$ and attains its minimum of $e \equiv \exp(1)$ when $\eta = 1$. So if $s < \pi e$, then (4) holds for all $\eta \in (0, 1]$ and the teacher maximizes $x^*(s, \eta)$ by choosing $\eta = 1$. In contrast, if $s > \pi e$, then the intermediate value theorem implies there is a unique $\eta^{\text{FB}}(s) \in (0, 1)$ such that

$$\eta^{\text{FB}}(s) \exp\left(\frac{2}{\eta^{\text{FB}}(s)} - 1\right) = \frac{s}{\pi} \tag{7}$$

and the teacher maximizes $x^*(s, \eta)$ by choosing $\eta = \eta^{\text{FB}}(s)$. Moreover, since the left-hand side of (4) is decreasing in η but the right-hand side is increasing in s, the "first-best effort intensity" $\eta^{\text{FB}}(s)$ must decline as s rises. I illustrate this decline in Figure 2.

[9] If $\overline{u} = 0$, then $u^*(s, \eta) \geq u(0, s, \eta) = 0 = \overline{u}$ for all $s > 0$ and $\eta \in (0, 1]$, and so the student prefers to work on every task regardless of his skill level or the task's effort intensity.

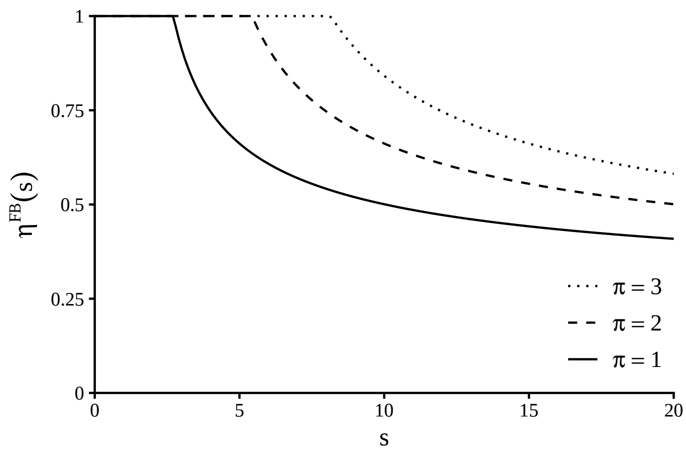

Figure 2: First-best effort intensities $\eta^{\text{FB}}(s)$ when $\pi \in \{1, 2, 3\}$.

Lemma 3 (First-best effort intensities)**.** *There is a unique function $\eta^{FB} : (0, \infty) \to (0, 1]$ such that (i) if $s \leq \pi e$, then $\eta^{FB}(s) = 1$, and (ii) if $s > \pi e$, then $\eta^{FB}(s) < 1$ satisfies* (7)*. This function is continuous, constant on $(0, \pi e]$, and decreasing on $(\pi e, \infty)$. Moreover, for all $s > 0$, the best-response effort $x^*(s, \eta)$ attains its maximum over $\eta \in (0, 1]$ when $\eta = \eta^{FB}(s)$.*

Theorem 1 uses the function $\eta^{\text{FB}} : (0, \infty) \to (0, 1]$ to characterize first-best curricula.

Theorem 1 (First-best curricula)**.** *Let $\overline{u} = 0$ and suppose $(\eta_t)_{t \in \mathcal{T}}$ solves the teacher's problem* (P)*. Then $\eta_t = \eta^{FB}(s_t)$ for each $t \in \mathcal{T}$. Moreover, there exists $t^{FB} \geq 0$ such that (i) if $t < t^{FB}$, then $\eta_t = 1$, and (ii) if $t \geq t^{FB}$, then $\eta_{t+1} < \eta_t$.*

If the student has skill s_t when he faces task t, then the first-best curriculum sets $\eta_t = \eta^{\text{FB}}(s_t)$. The student works on the task, exerts best-response effort $x^{\text{FB}}(s_t) \equiv x^*(s_t, \eta^{\text{FB}}(s_t))$, and increases his skill to $s_{t+1} = s_t + rx^{\text{FB}}(s_t)$. The next task has effort intensity $\eta^{\text{FB}}(s_{t+1})$, which is at most as large as $\eta^{\text{FB}}(s_t)$ because $s_{t+1} > s_t$ and η^{FB} is non-increasing on its domain.

Thus, under the first-best curriculum, earlier tasks are more effort-intensive and later tasks are more skill-intensive. The intuition is as follows. On all tasks, the teacher wants to incentivize as much effort as possible because this maximizes the student's skill development. On earlier tasks, the student has less skill and so the teacher incentivizes effort by making it contribute more to output. On later tasks, the student has more skill, which makes effort more productive because effort and skill are complementary inputs. The teacher capitalizes on this complementarity by making tasks less effort-intensive and more skill-intensive.

4.1.2 Best-responses

For all skill levels $s > 0$, let

$$x^{\text{FB}}(s) \equiv x^*(s, \eta^{\text{FB}}(s))$$

and

$$u^{\mathrm{FB}}(s) \equiv u^*(s, \eta^{\mathrm{FB}}(s))$$

denote the student's best-response effort and payoff under a first-best curriculum. Both are larger when s is larger:

Proposition 1 (First-best outcomes)**.** *The functions* $x^{FB} : (0,\infty) \to (0,\infty)$ *and* $u^{FB} : (0,\infty) \to (0,\infty)$ *are continuous, constant on* $(0, \pi e]$*, and increasing on* $(\pi e, \infty)$*.*

I use Proposition 1 to derive suitable parameter restrictions in my analysis of "second-best" curricula below.

4.2 Second-best curricula

Now I allow the delegation payoff $\overline{u}$ to exceed zero. This reintroduces the incentive constraint (IC) on the effort intensities $(\eta_t)_{t\in\mathcal{T}}$.

4.2.1 Parameter restrictions

Suppose the teacher chooses $\eta_t = \eta^{\mathrm{FB}}(s_t)$ for each task $t \in \mathcal{T}$. Then, by Proposition 1, the student's best-response payoff $u^*(s_t, \eta_t) = u^{\mathrm{FB}}(s_t)$ on task t is non-decreasing in his skill level s_t when he faces that task. But s_t is non-decreasing in t and so $u^{\mathrm{FB}}(s_t)$ is also non-decreasing in t. So if the constraint (IC) binds the choice of η_t for *any* task $t \in \mathcal{T}$, then it binds the choice of η_t for the initial task $t = 0$. For this reason, I assume

$$u^{\mathrm{FB}}(s_0) < \overline{u}; \tag{SB}$$

that is, the student prefers to delegate the initial task under the first-best curriculum. This is the minimal assumption needed to ensure the first- and second-best curricula differ: if (SB) does not hold, then the first-best curriculum induces work on every task and so the constraint (IC) has no bite.

Now consider the student's best-response payoff $u^*(s_0, \eta_0)$ from working on the initial task. If $u^*(s_0, \eta) < \overline{u}$ for all effort intensities $\eta \in (0, 1]$, then it is not possible for the teacher to induce work on the initial task. So $x^*(s_0, \eta_0) = 0$ and hence $s_1 = s_0$ regardless of the teacher's chosen η_0. But then $u^*(s_1, \eta) = u^*(s_0, \eta) < \overline{u}$ for all $\eta \in (0, 1]$, implying the teacher cannot induce work on task $t = 1$. Iterating this argument yields $s_T = s_{T-1} = \cdots = s_0$ regardless of $(\eta_t)_{t\in\mathcal{T}}$. Thus, if $u^*(s_0, \eta) < \overline{u}$ for all $\eta \in (0, 1]$, then the teacher's problem (P) is degenerate because she cannot induce work on *any* tasks. To avoid this degeneracy, I assume

$$\sup_{\eta\in(0,1]} u^*(s_0, \eta) > \overline{u}; \tag{ND}$$

that is, it is possible for the teacher to induce work on the initial task (and, thus, subsequent tasks).

Assumptions (SB) and (ND) impose restrictions on the model parameters:

Lemma 4 (Parameter restrictions)**.**

(i) If (SB) *holds, then* $\pi^2/2 < \overline{u}$*.*

(ii) If (SB) *and* (ND) *hold, then* $\overline{u} < \pi s_0$*.*

Together, assumptions (SB) and (ND) imply the delegation payoff $\overline{u}$ lies within the open interval $(\pi^2/2, \pi s_0)$ determined by the productivity parameter π and initial skill level s_0. Moreover, since s_t is non-decreasing in t, the assumptions imply $\overline{u} < \pi s_t$ for each $t \in \mathcal{T}$. Accordingly, I focus on skill levels $s > \overline{u}/\pi$ for the remainder of this section.

4.2.2 Incentive compatibility

Suppose (SB) and (ND) hold, and let $s > \overline{u}/\pi$. By Lemma 2, the best-response payoff $u^*(s, \eta)$ attains its minimum over $\eta \in (0, 1]$ at

$$\eta_{\min}(s) \equiv \min\left\{\frac{s}{\pi}, 1\right\}.$$

The payoff falls continuously from

$$\lim_{\eta \to 0} u^*(s, \eta) = \pi s$$

to

$$u^*(s, \eta_{\min}(s)) = \frac{\pi^2}{2}\begin{cases}\left(2 - \frac{s}{\pi}\right)\frac{s}{\pi} & \text{if } \frac{s}{\pi} \leq 1 \\ 1 & \text{if } \frac{s}{\pi} > 1\end{cases}$$

as η rises from zero to $\eta_{\min}(s)$, then rises continuously to

$$u^*(s, 1) = \frac{\pi^2}{2}$$

as η rises to one. So, by Lemma 4 and the intermediate value theorem, there is a unique effort intensity $\eta^{\mathrm{IC}}(s) \in (0, \eta_{\min}(s))$ such that $u^*(s, \eta^{\mathrm{IC}}(s)) = \overline{u}$. Moreover, since $u^*(s, \eta)$ is decreasing in η over $(0, \eta_{\min}(s))$ and bounded above by $\pi^2/2$ over $[\eta_{\min}(s), 1]$, we have $u^*(s, \eta) \geq \overline{u}$ if and only if $\eta \leq \eta^{\mathrm{IC}}(s)$. I characterize the function $\eta^{\mathrm{IC}} : (\overline{u}/\pi, \infty) \to (0, 1)$ in Lemma 5.

Lemma 5 (Incentive-compatible effort intensities)**.** *Suppose* (SB) *and* (ND) *hold. For all* $s > \overline{u}/\pi$*, there is a unique*

$$\eta^{IC}(s) \in \left(0, \min\left\{\frac{s}{\pi}, 1\right\}\right)$$

such that $u^*(s, \eta^{IC}(s)) = \overline{u}$*. The function* $\eta^{IC} : (\overline{u}/\pi, \infty) \to (0, 1)$ *is continuous and increasing, with*

$$\lim_{s \to \overline{u}/\pi} \eta^{IC}(s) = 0 \text{ and } \lim_{s \to \infty} \eta^{IC}(s) = 1.$$

Moreover, the incentive constraint (IC) *holds if and only if* $\eta_t \leq \eta^{IC}(s_t)$*.*

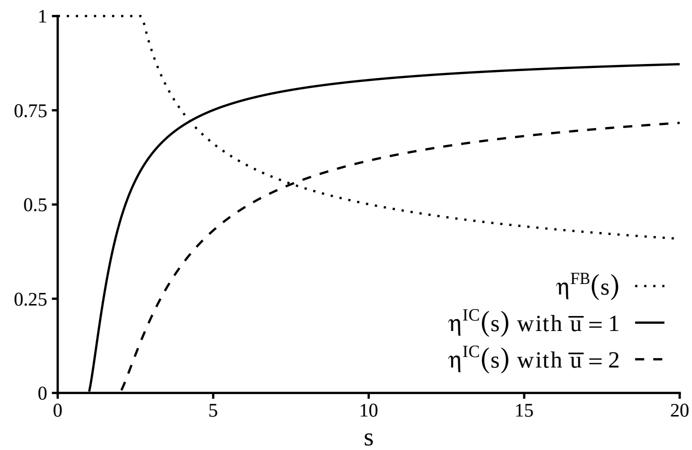

Figure 3: First-best and incentive-compatible effort intensities when $\pi = 1$ and $\overline{u} \in \{1, 2\}$.

Figure 3 shows that the incentive-compatible effort intensity $\eta^{\mathrm{IC}}(s)$ is increasing in the skill level s and decreasing in the delegation payoff $\overline{u}$. Intuitively, if the student has more skill, then he is more willing to work—by Lemma 2, his best-response payoff $u^*(s, \eta)$ is increasing in s—and so the incentive constraint (IC) has less bite. In contrast, if the delegation payoff $\overline{u}$ is larger, then the student is more tempted to delegate and so the incentive constraint has more bite.

4.2.3 Optimal effort intensities

The student works on task $t \in \mathcal{T}$ if and only if (IC) holds. By Lemma 5, this happens precisely when the task's effort intensity η_t is equal to at most the incentive-compatible intensity $\eta^{\mathrm{IC}}(s_t)$. So if $\eta^{\mathrm{IC}}(s_t)$ is larger than the first-best intensity $\eta^{\mathrm{FB}}(s_t)$ defined in Section 4.1, then the first-best curriculum is incentive-compatible because setting $\eta_t = \eta^{\mathrm{FB}}(s_t)$ makes the student prefer to work. However, if $\eta^{\mathrm{IC}}(s_t)$ is *smaller* than $\eta^{\mathrm{FB}}(s_t)$, then the first-best curriculum is *not* incentive-compatible because setting $\eta_t = \eta^{\mathrm{FB}}(s_t)$ makes the student prefer to delegate. Then, to satisfy (IC), the teacher must set η_t strictly below $\eta^{\mathrm{FB}}(s_t)$. But Lemma 3 implies the student's best-response effort $x^*(s_t, \eta)$ is increasing in η over $(0, \eta^{\mathrm{FB}}(s_t))$. So if $\eta^{\mathrm{IC}}(s_t) < \eta^{\mathrm{FB}}(s_t)$, then the teacher maximizes $x^*(s_t, \eta_t)$ subject to (IC) by setting $\eta_t = \eta^{\mathrm{IC}}(s_t)$. Thus, the teacher optimally sets η_t equal to the minimum of $\eta^{\mathrm{IC}}(s_t)$ and $\eta^{\mathrm{FB}}(s_t)$. I characterize this minimum in the following lemma.

Lemma 6 (Second-best effort intensity). *Suppose* (SB) *and* (ND) *hold, and define*

$$\eta^{SB}(s) \equiv \min\Big\{\eta^{IC}(s),\, \eta^{FB}(s)\Big\} \tag{8}$$

for all $s > \overline{u}/\pi$. There is a unique skill level

$$\overline{s} \in \left(\max\left\{\frac{\overline{u}}{\pi},\, \pi e\right\},\, \infty\right)$$

such that $\eta^{IC}(\overline{s}) = \eta^{FB}(\overline{s})$, and

$$\eta^{SB}(s) = \begin{cases} \eta^{IC}(s) & \text{if } s < \overline{s} \\ \eta^{FB}(s) & \text{if } s \geq \overline{s} \end{cases}$$

is increasing in s when $s < \overline{s}$ and decreasing in s when $s > \overline{s}$.

Figure 3 shows that $\eta^{\mathrm{SB}}(s) = \min\{\eta^{\mathrm{IC}}(s), \eta^{\mathrm{FB}}(s)\}$ is increasing at low skill levels s and decreasing at high levels.

Theorem 2 uses the function $\eta^{\mathrm{SB}} : (\overline{u}/\pi, \infty) \to (0,1)$ to characterize second-best curricula.

Theorem 2 (Second-best curricula). *Suppose* (SB) *and* (ND) *hold, and $(\eta_t)_{t\in\mathcal{T}}$ solves the teacher's problem* (P). *Then $\eta_t = \eta^{SB}(s_t)$ for each $t \in \mathcal{T}$. Moreover, there exists $t^{SB} \in \mathcal{T}$ such that (i) if $t^{SB} > 0$ and $0 < t \leq t^{SB}$, then $\eta_t > \eta_{t-1}$, and (ii) if $t > t^{SB}$, then $\eta_t < \eta_{t-1}$.*

First-best curricula comprise tasks with non-increasing effort intensities (see Theorem 1). In contrast, second-best curricula comprise tasks with increasing-then-decreasing effort intensities. This non-monotonicity stems from the student's option to delegate. On early tasks, his skill level s_t is low, so his payoff $u^{\mathrm{FB}}(s_t)$ under a first-best curriculum is also low (by Proposition 1) and delegation is relatively attractive. To prevent delegation, the teacher must raise the student's payoff from working, which she achieves by distorting effort intensities downward. As the student works on tasks and gains skill, his payoff $u^{\mathrm{FB}}(s_t)$ under a first-best curriculum rises, and so the teacher does not need to distort the curriculum as much to make the student prefer working over delegating. Thus, on early tasks, the second-best curriculum becomes gradually more effort-intensive. Eventually, the student has built enough skill that his payoff $u^{\mathrm{FB}}(s_t)$ under a first-best curriculum exceeds his payoff $\overline{u}$ from delegating (i.e., the curriculum is incentive-compatible). Then the teacher proceeds with a first-best curriculum, which becomes gradually less effort-intensive (by Theorem 1).

Theorem 2 formalizes a suggestion made by Chirikov et al. (2026) in their analysis of AI misuse and assessment reform. The authors suggest "requiring students to document their process, justify their choices, [and] demonstrate understanding [of] underlying reasoning" (Chirikov et al., 2026, p. 820). These redesign choices correspond to making tasks more skill-intensive: the student must be able to identify and explain appropriate methods for completing tasks.

4.2.4 Best-responses

Suppose (SB) and (ND) hold, let $s > \overline{u}/\pi$, and let

$$x^{\mathrm{FB}}(s) \equiv x^*(s, \eta^{\mathrm{FB}}(s)) \text{ and } x^{\mathrm{SB}}(s) \equiv x^*(s, \eta^{\mathrm{SB}}(s))$$

denote the student's best-response efforts under first- and second-best curricula. Since $x^{\mathrm{FB}}(s)$ maximizes $x^*(s,\eta)$ over $\eta \in (0,1]$ (by Lemma 3), we have

$$x^{\mathrm{SB}}(s) \leq x^{\mathrm{FB}}(s)$$

with equality if and only if $\eta^{\mathrm{SB}}(s) = \eta^{\mathrm{FB}}(s)$. Figure 4 illustrates this bound, and shows $x^{\mathrm{SB}}(s)$ is increasing in the skill level s when the student has productivity $\pi = 1$ and delegation payoff $\overline{u} = 2$. Proposition 2 establishes that $x^{\mathrm{SB}}(s)$ is increasing in s for *all* values of π and $\overline{u}$, despite the second-best effort intensity $\eta^{\mathrm{SB}}(s)$ being non-monotone in s.

Proposition 2 (Second-best efforts)**.** *Suppose* (SB) *and* (ND) *hold, and let $s > \overline{u}/\pi$. Then the second-best effort $x^{SB}(s)$ is increasing in s.*

Now let

$$u^{\mathrm{FB}}(s) \equiv u^*(s, \eta^{\mathrm{FB}}(s)) \text{ and } u^{\mathrm{SB}}(s) \equiv u^*(s, \eta^{\mathrm{SB}}(s))$$

denote the student's best-response payoffs under first- and second-best curricula. Lemmas 5 and 6 imply

$$u^{\mathrm{SB}}(s) = \max\{\overline{u}, u^{\mathrm{FB}}(s)\},$$

and so

$$u^{\mathrm{SB}}(s) \geq u^{\mathrm{FB}}(s)$$

with equality if and only if $\eta^{\mathrm{SB}}(s) = \eta^{\mathrm{FB}}(s)$.[10] Thus, at all skill levels, the student exerts weakly less effort but obtains weakly higher payoffs under a second-best curriculum than under a first-best curriculum. That he exerts less effort implies he develops less skill, making the second-best curriculum worse from the teacher's perspective even though it may be better from the student's perspective.[11]

[10] Since $u^{\mathrm{FB}}(s)$ is increasing in s (by Proposition 1) and exceeds $\overline{u}$ precisely when s exceeds the threshold $\bar{s}$ defined in Lemma 6, the second-best payoff $u^{\mathrm{SB}}(s)$ is constant in s when $s < \bar{s}$ and increasing in s when $s \geq \bar{s}$.

[11] Although the student has higher payoffs for a given skill level under a second-best curriculum than under a first-best curriculum, his skill levels $(s_t)_{t\in\mathcal{T}}$ under a second-best curriculum are lower because he does not exert as much effort on early tasks. Consequently, his payoff $u^{\mathrm{SB}}(s_t)$ on later tasks t under a second-best curriculum may be smaller than his payoff $u^{\mathrm{FB}}(s_t)$ on such tasks under a first-best curriculum. The student's overall preference between curricula depends on whether the increase in his payoffs on early tasks (from $u^{\mathrm{FB}}(s_t)$ to $\overline{u}$) dominates the decrease in his payoffs on later tasks.

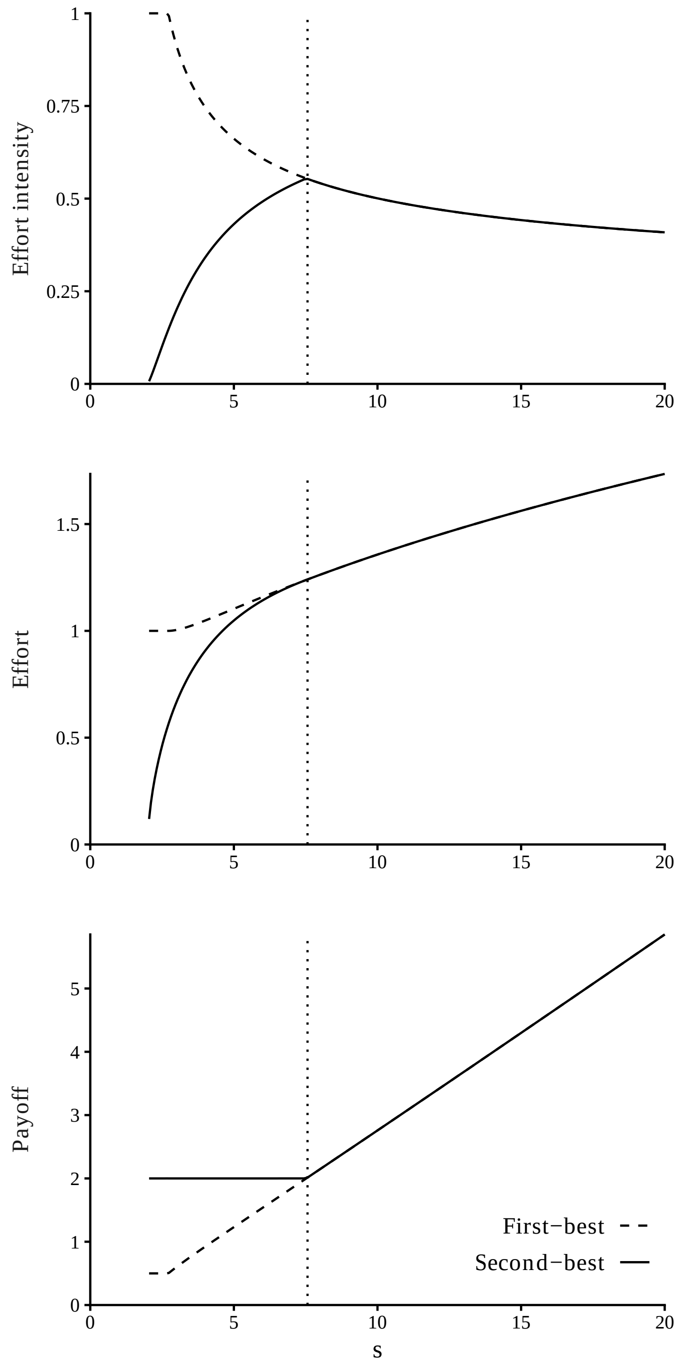

Figure 4: Optimal effort intensities, best-response efforts, and best-response payoffs when $\pi = 1$ and $\overline{u} = 2$. Dotted line indicates threshold $\bar{s}$ defined in Lemma 6.

4.3 Second-best skill gains

Suppose conditions (SB) and (ND) hold, and the teacher designs a curriculum that maximizes the student's overall skill gain

$$s_T - s_0 = \sum_{t \in \mathcal{T}} \Delta(s_t, \eta_t).$$

By Theorem 2, the teacher chooses $\eta_t = \eta^{\text{SB}}(s_t)$ for each task $t \in \mathcal{T}$. This choice satisfies the incentive constraint (IC), so the student exerts effort $x^{\text{SB}}(s_t) \equiv x^*(s_t, \eta^{\text{SB}}(s_t))$ and gains skill

$$\Delta(s_t, \eta_t) = r x^{\text{SB}}(s_t)$$

by working on each task $t \in \mathcal{T}$.

Lemma 7 (Skill increments)**.** *Suppose* (SB) *and* (ND) *hold, and* $(\eta_t)_{t \in \mathcal{T}}$ *solves the teacher's problem* (P)*. For each task* $t \in \mathcal{T}$*, the skill increment* $\Delta(s_t, \eta_t)$ *is increasing in the student's initial skill* s_0*, productivity parameter* π*, and skill accumulation rate* r*, but decreasing in the delegation payoff* $\overline{u}$*.*

The following result follows immediately from Lemma 7.

Theorem 3 (Overall skill gain)**.** *Suppose* (SB) *and* (ND) *hold, and* $(\eta_t)_{t \in \mathcal{T}}$ *solves the teacher's problem* (P)*. Then the student's overall skill gain* $(s_T - s_0)$ *is (i) increasing in his initial skill* s_0*, (ii) increasing in the productivity parameter* π*, (iii) increasing in the skill accumulation rate* r*, and (iv) decreasing in the delegation payoff* $\overline{u}$*.*

On each task, the student gains more skill when (i) he has more skill at the outset (i.e., s_0 is higher), since skill complements effort and grows across tasks; (ii) he is more productive (i.e., π is higher), since he is more willing to exert effort; (iii) he gains more skill per unit of effort exerted (i.e., r is higher). He gains *less* skill when delegation is more tempting (i.e., $\overline{u}$ is higher), since the teacher needs to distort the curriculum further from the first-best.

The effects of s_0, π, r, and $\overline{u}$ on the skill gained from each task also compound across tasks. This is because $s_{t+1} = s_t + r x^{\text{SB}}(s_t)$ is increasing in $x^{\text{SB}}(s_t)$, so any parameter change that raises $x^{\text{SB}}(s_t)$ also raises s_{t+1}, which, by Proposition 2, raises $x^{\text{SB}}(s_{t+1})$ too—a positive feedback loop. For example, a one-off boost to the student's initial skill raises the effort he exerts on the initial task *and* the effort he exerts on subsequent tasks, since he enters each task more skilled than he would have without the boost. Thus, as emphasized by Heckman (2000) and Cunha and Heckman (2007), "skill begets skill" due to dynamic complementarities.

5 Extension: Complementary AI

So far I have assumed AI substitutes for effort. Now I consider the possibility that AI complements effort. For example, rather than prompt AI to "do my homework," the student could use skill to break his homework into subtasks and use AI to accelerate the effortful completion of those subtasks.[12] So the student can use AI to avoid or enhance his effort, consistent with survey evidence on students' reported uses of AI (Contractor and Reyes, 2026).

To capture the possibility that AI can complement or substitute for effort, I introduce an "AI quality" parameter $q \geq 1$ that enters the working payoff

$$u(x, s, \eta; q) \equiv \pi(qx)^{\eta} s^{1-\eta} - \frac{x^2}{2}$$

and delegation payoff $q\overline{u}$. Letting $q = 1$ yields the payoffs defined in Section 2. Increasing q makes effort more productive *and* raises the delegation payoff. If the student works on a task, then his best-response effort

$$x^*(s, \eta; q) \in \arg\max_{x \geq 0} u(x, s, \eta; q)$$

and payoff

$$u^*(s, \eta; q) \equiv \max_{x \geq 0} u(x, s, \eta; q)$$

both depend on q, and so the first- and second-best curricula will also depend on q.[13] I study this dependence in Sections 5.1 and 5.2, following the steps in my analysis of the case with $q = 1$ (see Sections 4.1 and 4.2).

5.1 First-best curricula

Consider the first-best case in which $\overline{u} = 0$ and the student always prefers to work. Letting $y \equiv qx$ denote his AI-enhanced effort gives

$$u(x, s, \eta; q) = \frac{1}{q^2}\left(\pi q^2 y^{\eta} s^{1-\eta} - \frac{y^2}{2}\right) \tag{9}$$

for all skill levels $s > 0$ and effort intensities $\eta \in (0, 1]$. The bracketed term on the right-hand side of (9) equals the payoff from the case with $q = 1$, with x replaced by y and π replaced by πq^2. So, by Lemma 1, the student maximizes the bracketed term by choosing

$$y = s\left(\frac{(\pi q^2)\eta}{s}\right)^{1/(2-\eta)},$$

[12] For example, suppose the task is to prove a mathematical theorem. The student could pass the entire theorem to a Large Language Model (LLM) before submitting a simple user prompt: "prove the theorem." Alternatively, he could devise a proof strategy that involves proving a series of lemmas, then using the LLM to prove each of those lemmas.

[13] I write q (a primitive) after a semicolon to distinguish it from s (a state variable) and η (a choice variable). I include q as an argument but not π (also a primitive) because my focus is on how outcomes depend on q. Letting $q = 1$ yields the best-response effort $x^*(s, \eta)$ and payoff $u^*(s, \eta)$ defined in Lemmas 1 and 2. The qualitative properties of $x^*(s, \eta)$ and $u^*(s, \eta)$ established in those lemmas extend to $x^*(s, \eta; q)$ and $u^*(s, \eta; q)$.

which corresponds to exerting best-response effort

$$x^*(s,\eta;q) = s\left(\frac{\pi\eta q^\eta}{s}\right)^{1/(2-\eta)}. \tag{10}$$

The effort $x^*(s,\eta;q)$ is increasing in the AI quality parameter q—intuitively, raising q makes effort more productive, so the student is willing to exert more.

Replacing π by πq^2 also yields a more general version of the first-best effort intensity

$$\eta^{\mathrm{FB}}(s;q) \in \operatorname*{arg\,max}_{\eta\in(0,1]} x^*(s,\eta;q)$$

defined in Lemma 3:

(i) If $s \leq \pi e q^2$, then $\eta^{\mathrm{FB}}(s;q) = 1$.

(ii) If $s > \pi e q^2$, then $\eta^{\mathrm{FB}}(s;q) < 1$ is the unique solution to

$$\eta^{\mathrm{FB}}(s;q)\exp\left(\frac{2}{\eta^{\mathrm{FB}}(s;q)} - 1\right) = \frac{s}{\pi q^2}, \tag{11}$$

and is decreasing in s and increasing in q.[14]

Since the teacher wants to maximize $x^*(s_t,\eta_t;q)$ for each task $t \in \mathcal{T}$ (see Section 4 and Appendix Section A.4), she optimally chooses $\eta_t = \eta^{\mathrm{FB}}(s_t;q)$, which is non-decreasing in q for all $s_t > 0$ and increasing in q when $s_t > \pi e q^2$. Thus, she optimally assigns more effort-intensive tasks when AI has higher quality. She does so to capitalize on the student's higher willingness to exert effort.

Now let $s > 0$, and define the first-best effort and payoff

$$x^{\mathrm{FB}}(s;q) \equiv x^*(s,\eta^{\mathrm{FB}}(s;q);q) \quad\text{and}\quad u^{\mathrm{FB}}(s;q) \equiv u^*(s,\eta^{\mathrm{FB}}(s;q);q)$$

analogously to Section 4.1. Improvements in AI quality raise $x^{\mathrm{FB}}(s;q)$ but may lower $u^{\mathrm{FB}}(s;q)$:

Proposition 3 (First-best outcomes and AI quality)**.** *Let $s > 0$. The first-best effort $x^{FB}(s;q)$ is increasing in the AI quality parameter q. The first-best payoff $u^{FB}(s;q)$ is decreasing in q when $q < \sqrt{s/\pi e}$ and increasing in q when $q \geq \sqrt{s/\pi e}$.*

The intuition for Proposition 3 is as follows. The teacher chooses $\eta^{\mathrm{FB}}(s_t;q)$ to maximize the student's effort, rather than his payoff, and the effort-maximizing intensity is larger than the payoff-maximizing intensity because the former does not account for the cost of exerting effort. As AI quality improves, the teacher optimally makes tasks more effort-intensive to keep extracting maximal effort from the now-more-productive student. This pulls $\eta^{\mathrm{FB}}(s_t;q)$ further from the payoff-maximizing intensity, and so the student's payoff *falls* even as his effort and skill gain rise. Only when $\eta^{\mathrm{FB}}(s_t;q)$ reaches its ceiling of one—which, by the definition of $\eta^{\mathrm{FB}}(s;q)$, happens precisely when $q \geq \sqrt{s_t/\pi e}$—do further quality improvements lead to higher payoffs with no change in effort intensities.

[14] See Appendix Section A.3 for the derivation of $\eta^{\mathrm{FB}}(s;q)$ and its properties.

5.2 Second-best curricula

Now consider the second-best case in which $\overline{u} > 0$ and the student may prefer to delegate. Given a skill level $s > 0$ and effort intensity $\eta \in (0, 1]$, the student's best-response effort $x^*(s, \eta; q)$ yields payoff[15]

$$u^*(s, \eta; q) = \frac{1}{\eta}\left(1 - \frac{\eta}{2}\right)(x^*(s, \eta; q))^2. \tag{12}$$

He works on task $t \in \mathcal{T}$ if and only if his best-response payoff from working exceeds his payoff from delegating:

$$u^*(s_t, \eta_t; q) \geq q\overline{u}. \tag{IC;q}$$

5.2.1 Parameter restrictions

As in Section 4.2, I assume the first-best curriculum is not incentive-compatible:

$$u^{\mathrm{FB}}(s_0; q) < q\overline{u}. \tag{SB;q}$$

I also assume it is possible to design an incentive-compatible curriculum:

$$\sup_{\eta \in (0,1]} u^*(s_0, \eta; q) > q\overline{u}. \tag{ND;q}$$

Conditions (SB; q) and (ND; q) generalize conditions (SB) and (ND) from Section 4.2, and imply the restrictions

$$\frac{\pi^2 q^2}{2} < q\overline{u} < \pi s_0 \tag{13}$$

on the model parameters.[16]

5.2.2 Incentive compatibility

Suppose (SB; q) and (ND; q) hold, and let $s > q\overline{u}/\pi$. As in Section 4.2, there is a unique "incentive-compatible effort intensity"

$$\eta^{\mathrm{IC}}(s; q) \in \left(0, \min\left\{\frac{s}{\pi q^2}, 1\right\}\right)$$

such that

$$u^*(s, \eta^{\mathrm{IC}}(s; q); q) = q\overline{u}. \tag{14}$$

Moreover, the incentive constraint (IC; q) holds if and only if $\eta_t \leq \eta^{\mathrm{IC}}(s_t; q)$.[17]

Figure 5 shows that whether $\eta^{\mathrm{IC}}(s; q)$ is increasing in the AI quality parameter q depends on whether $\eta^{\mathrm{IC}}(s; q) > 2/3$. Lemma 8 characterizes this dependence in terms of the skill level s.

[15]See Appendix Section A.2 for the derivation of (12).
[16]See Appendix Section A.6 for the derivation of (13).
[17]See Appendix Section A.7 for the proof that $\eta^{\mathrm{IC}}(s; q)$ exists.

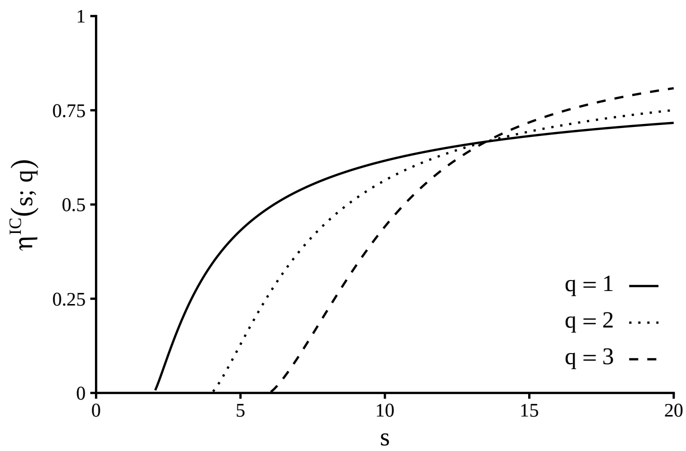

Figure 5: Incentive-compatible effort intensities $\eta^{\mathrm{IC}}(s;q)$ when $\pi = 1$ and $\overline{u} = 2$.

Lemma 8 (Incentive compatibility and AI quality)**.** *Suppose* (SB; q) *and* (ND; q) *hold, let* $s > q\overline{u}/\pi$, *and define*

$$s^{\dagger} \equiv \frac{27\,\overline{u}^2}{8\,\pi^3}. \tag{15}$$

Then $\eta^{IC}(s;q)$ *is decreasing in* q *when* $s < s^{\dagger}$, *constant in* q *when* $s = s^{\dagger}$, *and increasing in* q *when* $s > s^{\dagger}$. *Moreover, we have* $s < s^{\dagger}$ *if and only if* $\eta^{IC}(s;q) < 2/3$.

To see why the effect of raising q on $\eta^{\mathrm{IC}}(s;q)$ depends on whether $\eta^{\mathrm{IC}}(s;q) > 2/3$, fix $\eta \in (0,1]$, and decompose the student's best-response payoff (12) into the product

$$u^*(s,\eta;q) = q^{2\eta/(2-\eta)}\, u^*(s,\eta;1)$$

of terms that do and do not depend on q. Thus, the AI quality parameter q enters the best-response payoff $u^*(s,\eta;q)$ with exponent $2\eta/(2-\eta)$ and the delegation payoff $q\overline{u}$ with exponent 1. So if q rises, then $u^*(s,\eta;q)$ rises faster than $q\overline{u}$ precisely when $2\eta/(2-\eta) > 1$, which holds if and only if $\eta > 2/3$. But $u^*(s,\eta;q)$ is decreasing in η in a neighborhood of $\eta^{\mathrm{IC}}(s;q)$.[18] So if $u^*(s,\eta;q)$ rises faster than $q\overline{u}$ for η near $\eta^{\mathrm{IC}}(s;q)$, then $\eta^{\mathrm{IC}}(s;q)$ must fall to preserve the identity (14). Thus, the effect of increasing q on $\eta^{\mathrm{IC}}(s;q)$ depends on whether $\eta^{\mathrm{IC}}(s;q) > 2/3$. The threshold $s^{\dagger}$ is precisely the skill level at which $\eta^{\mathrm{IC}}(s^{\dagger};q) = 2/3$.

5.2.3 Optimal effort intensities

Suppose (SB; q) and (ND; q) hold so that $s_t > q\overline{u}/\pi$ for each $t \in \mathcal{T}$. As in Section 4.2, the teacher optimally sets each effort intensity $\eta_t \in (0,1]$ equal to the minimum of the incentive-compatible intensity $\eta^{\mathrm{IC}}(s_t;q)$ and first-best intensity $\eta^{\mathrm{FB}}(s_t;q)$. I characterize this minimum in Lemma 9, which generalizes Lemma 6 to cases with $q > 1$.

[18] Lemmas A.2 and A.5 imply $\eta^{\mathrm{IC}}(s;q)$ belongs to an interval over which $u^*(s,\eta;q)$ is decreasing in η.

Lemma 9 (Second-best effort intensities with complementary AI)**.** *Suppose* (SB; q) *and* (ND; q) *hold, and define*

$$\eta^{SB}(s;q) \equiv \min\left\{\eta^{IC}(s;q),\, \eta^{FB}(s;q)\right\} \tag{16}$$

for all $s > q\bar{u}/\pi$. There is a unique skill level

$$\bar{s}(q) \in \left(\max\left\{\frac{q\bar{u}}{\pi},\, \pi e q^2\right\},\, \infty\right)$$

such that $\eta^{IC}(\bar{s}(q);q) = \eta^{FB}(\bar{s}(q);q)$, and

$$\eta^{SB}(s;q) = \begin{cases} \eta^{IC}(s;q) & \text{if } s < \bar{s}(q) \\ \eta^{FB}(s;q) & \text{if } s \geq \bar{s}(q) \end{cases} \tag{17}$$

is increasing in s when $s < \bar{s}(q)$ and decreasing in s when $s > \bar{s}(q)$. Moreover, the threshold $\bar{s}(q)$ is increasing in the AI quality parameter q.

Thus, as in the case with $q = 1$ (see Section 4.2), the second-best effort intensity $\eta^{\mathrm{SB}}(s;q) < 1$ is increasing-then-decreasing in the skill level s:

$$\frac{\partial \eta^{\mathrm{SB}}(s;q)}{\partial s} > 0 \text{ if and only if } s < \bar{s}(q).$$

The threshold $\bar{s}(q)$ at which $\partial\eta^{\mathrm{SB}}(s;q)/\partial s$ changes sign is increasing in the AI quality parameter q. This is because $\bar{s}(q)$ belongs to the set $(\pi e q^2, \infty)$ of skill levels s at which increasing q lowers the first-best payoff $u^{\mathrm{FB}}(s;q)$ (see Proposition 3). As q rises, working on a first-best task becomes less attractive to the student while delegating becomes more attractive (since $q\bar{u}$ rises). As a result, it takes more accumulated skill for the payoff from working on a first-best task to be more attractive than delegating.

The skill level s also determines whether $\eta^{\mathrm{SB}}(s;q)$ rises or falls when AI quality improves. To see why, recall the thresholds $s^\dagger$ and $\bar{s}(q)$ defined in Lemmas 8 and 9. If $s < \bar{s}(q)$, then $\eta^{\mathrm{SB}}(s;q) = \eta^{\mathrm{IC}}(s;q)$ is decreasing in q when $s < s^\dagger$ and increasing in q when $s > s^\dagger$. In contrast, if $s \geq \bar{s}(q)$, then $\eta^{\mathrm{SB}}(s;q) = \eta^{\mathrm{FB}}(s;q)$ is increasing in q always. Thus, the effect of increasing q on $\eta^{\mathrm{SB}}(s;q)$ changes from negative to positive as s passes through $\min\{s^\dagger, \bar{s}(q)\}$ from below. Intuitively, if the student has low skill, then AI quality improvements raise his delegation payoff faster than they make effort more productive, and so the teacher must distort effort intensities downward to incentivize effort. In contrast, if the student has high skill, then AI quality improvements make effort more productive faster than they raise his delegation payoff, and so the teacher makes tasks more effort-intensive to capitalize on the student's higher willingness to exert effort.

5.2.4 Best-response efforts

Suppose (SB; q) and (ND; q) hold, and let $s > q\bar{u}/\pi$. Figure 6 shows how the second-best effort

$$x^{\mathrm{SB}}(s;q) \equiv x^*(s, \eta^{\mathrm{SB}}(s;q); q)$$

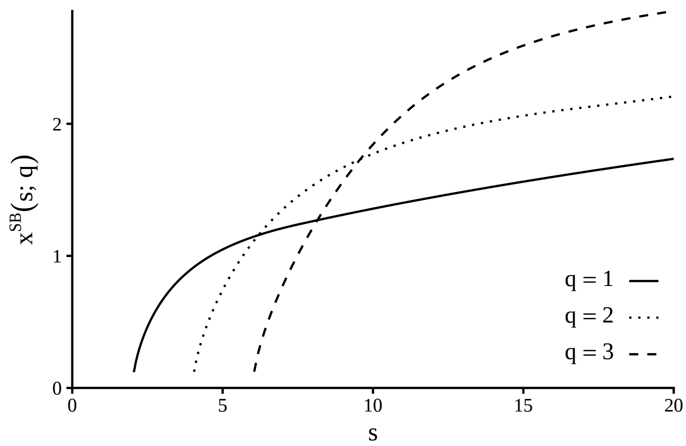

Figure 6: Second-best efforts $x^{\mathrm{SB}}(s;q)$ when $\pi = 1$ and $\overline{u} = 2$.

grows with the skill level $s > q\overline{u}/\pi$ when $(\pi, \overline{u}) = (1,2)$ and $q \in \{1,2,3\}$. If $s = 5$, then $x^{\mathrm{SB}}(s;1) > x^{\mathrm{SB}}(s;2)$. In contrast, if $s = 15$, then $x^{\mathrm{SB}}(s;1) < x^{\mathrm{SB}}(s;2)$. So $x^{\mathrm{SB}}(s;q)$ can be increasing or decreasing in the AI quality parameter q, depending on the skill level s.[19] I characterize this dependence in the main result of this section:

Theorem 4 (Second-best efforts and AI quality)**.** *Suppose* (SB; q) *and* (ND; q) *hold, let $s > q\overline{u}/\pi$, and define $s^{\dagger}$ and $\overline{s}(q)$ as in Lemmas 8 and 9. There is a unique skill level*

$$s^{\ddagger}(q) \in \left(\frac{q\overline{u}}{\pi}, \min\left\{s^{\dagger}, \overline{s}(q)\right\}\right)$$

such that the second-best effort $x^{\mathrm{SB}}(s;q)$ is decreasing in q when $s < s^{\ddagger}$ and increasing in q when $s > s^{\ddagger}$. Moreover, the threshold $s^{\ddagger}(q)$ is increasing in the AI quality parameter q.

Under a second-best curriculum, the student gains skill $s_{t+1} - s_t = rx^{\mathrm{SB}}(s_t;q)$ proportional to the effort $x^{\mathrm{SB}}(s_t;q)$ he exerts on each task $t \in \mathcal{T}$. Thus, Theorem 4 says that under a second-best curriculum, improvements in AI quality make high-skill students gain skill faster but low-skill students gain skill slower. This is because high-skill students become more willing to work (since their working payoff rises by more than their delegation payoff) while low-skill students become more tempted to delegate (since their delegation payoff rises faster than their working payoff).

5.3 Second-best skill gains

Suppose (SB; q) and (ND; q) hold, and the teacher designs a second-best curriculum: $\eta_t = \eta^{\mathrm{SB}}(s_t;q)$ for each task $t \in \mathcal{T}$. Then the student's overall skill gain

$$s_T - s_0 = r \sum_{t \in \mathcal{T}} x^{\mathrm{SB}}(s_t;q)$$

[19]The second-best effort $x^{\mathrm{SB}}(s;q)$ can also be non-monotone in q. For example, Figure 6 shows $x^{\mathrm{SB}}(s;3) < x^{\mathrm{SB}}(s;1) < x^{\mathrm{SB}}(s;2)$ when $(\pi, \overline{u}) = (1,2)$ and $s = 7$.

Table 1: Overall skill gains $(s_T - s_0)$ under second-best curriculum when $(s_0, \pi, \overline{u}) = (5, 1, 2)$.

	$r = 0.25$		$r = 0.5$		$r = 1$	
T	$q = 1$	$q = 2$	$q = 1$	$q = 2$	$q = 1$	$q = 2$
5	1.38	1.13	2.86	2.61	6.04	6.34
10	2.88	2.73	6.15	6.79	13.64	16.60
20	6.20	7.01	13.85	17.24	32.96	41.89

depends on his best-response efforts $x^{\text{SB}}(s_t; q)$, which depend on his skill levels s_t and the AI quality parameter q. By Theorem 4, increasing q raises $x^{\text{SB}}(s_t; q)$ when s_t is large but lowers $x^{\text{SB}}(s_t; q)$ when s_t is small. The overall effect of increasing q on $(s_T - s_0)$ depends on whether the positive effect at high skill levels dominate the negative effects at low skill levels.

For example, suppose the student has initial skill $s_0 = 5$, productivity parameter $\pi = 1$, and delegation payoff $q\overline{u}$ with $\overline{u} = 2$. He builds skill at rate $r = 0.25$. Figure 7 shows that he exerts more effort $x^{\text{SB}}(s_0; q)$ under a second-best curriculum when $q = 1$ than when $q = 2$. As a result, his skill $s_1 = s_0 + rx^{\text{SB}}(s_0; q)$ after completing the initial task is higher when $q = 1$ than when $q = 2$. However, his effort on the next few tasks grows faster when $q = 2$. For each task $t \geq 6$, the student exerts more effort—and so gains skill faster—when $q = 2$. Eventually his skill level s_t under $q = 2$ surpasses the level under $q = 1$. Indeed, for tasks $t \geq 13$ onwards, the student's skill level s_t is higher when $q = 2$ than when $q = 1$.

This example shows that the effect of increasing the AI quality q on the student's overall skill gain $(s_T - s_0)$ depends on the number of tasks T. In the example, increasing q from one to two raises $(s_T - s_0)$ when $T \geq 13$ but lowers $(s_T - s_0)$ when $T \leq 12$.

The effect of increasing q on $(s_T - s_0)$ also depends on the skill accumulation rate r. Intuitively, the larger is r, the faster the student's skill level s_t passes the threshold $s^{\ddagger}(q)$, and so the higher is the proportion of tasks on which increasing q has positive effects (by raising the marginal product of effort) rather than negative effects (by making delegation more attractive).

I quantify this intuition in Table 1. It shows how $(s_T - s_0)$ depends on the number of tasks T and skill accumulation rate r when $(s_0, \pi, \overline{u}) = (5, 1, 2)$ and $q \in \{1, 2\}$ as in the example above. If there are few tasks ($T = 5$), then increasing q from one to two lowers $(s_T - s_0)$ when $r \in \{0.25, 0.5\}$ but raises $(s_T - s_0)$ when $r = 1$. In contrast, if there are many tasks $(T = 20)$, then increasing q from one to two raises $(s_T - s_0)$ when $r \in \{0.25, 0.5\}$. Intuitively, if the curriculum is short, then the student must gain skill quickly for the positive effects of AI quality improvements on later tasks to outweigh the negative effects on early tasks. If the curriculum is longer, then the student does not need to gain skill as quickly for AI quality improvements to have positive effects overall.

The patterns above suggest AI quality improvements increase the return on "learning how to learn" (Novak and Gowin, 1984). If the student gains more skill per unit of effort (i.e., r is higher),

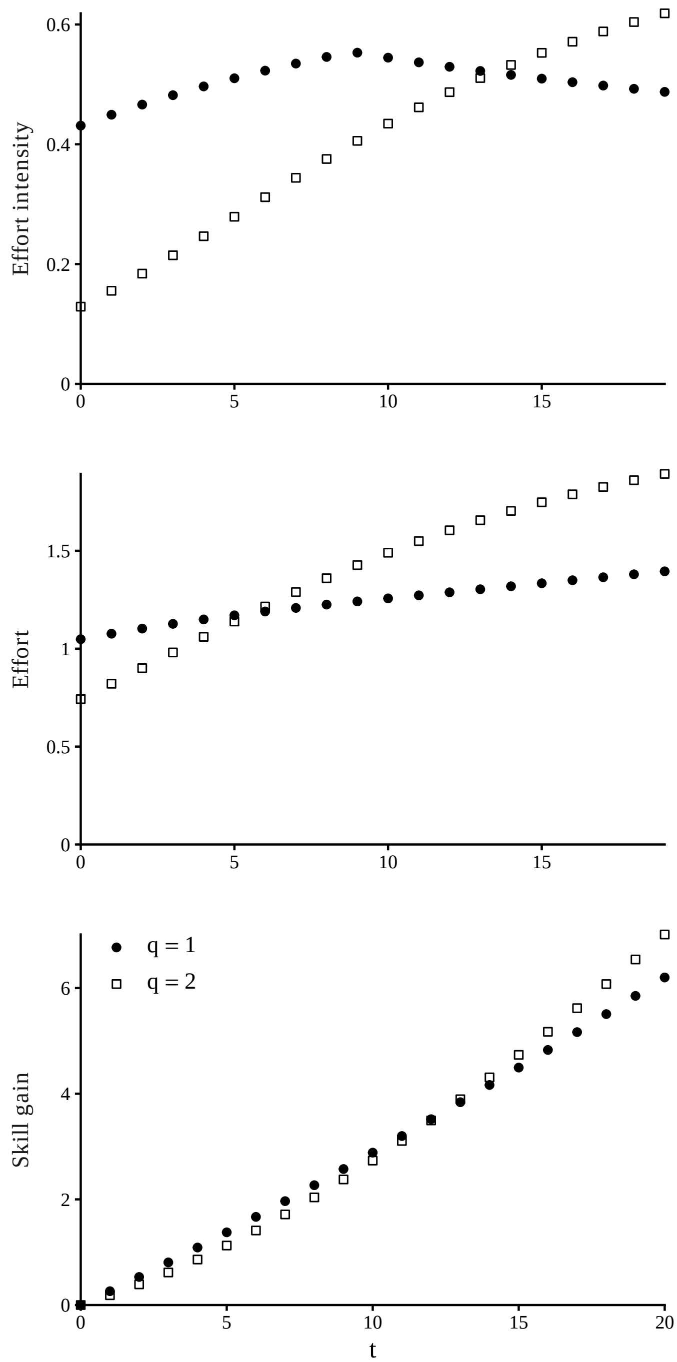

Figure 7: Effort intensities η_t, best-response efforts $x^*(s_t, \eta_t; q)$, and consequent skill gains $(s_t - s_0)$ under second-best curriculum when $(s_0, \pi, \overline{u}, r) = (5, 1, 2, 0.25)$.

then he overcomes his temptation to delegate earlier in the curriculum and so is faster to realize the productivity-enhancing benefits of AI.

6 Conclusion

This paper presents a model of learning-by-doing and curriculum design, and uses it to analyze the impact of AI on a student's skill development. The model delivers four main results. First, if the student is never tempted to delegate to AI, then the teacher optimally designs curricula that become gradually less effort-intensive and more skill-intensive (Theorem 1). Second, if the student can be tempted to delegate, then the teacher must distort the curriculum to be more skill-intensive (Theorem 2). Third, the student's overall skill gain rises when he starts with more skill (i.e., "skill begets skill") and falls when delegation is more tempting (Theorem 3). Fourth, if AI complements effort, then improvements in AI quality make high-skill students build skill faster but low-skill students build skill slower (Theorem 4).

To make the model tractable, I make several assumptions that could be relaxed in future research. One is that the student is myopic. My model maps to a setting where a student cares only about current performance (e.g., grade) and not future skill. However, many real-world students pursue education precisely because they want to gain skill. A skill-seeking student would internalize the future benefit of building skill, making them more willing to exert effort. The teacher would capitalize on this willingness by making tasks more effort-intensive.

Another assumption is that the teacher optimizes her curriculum for one student with a known initial skill level. Yet real-world teachers design curricula for many students with unknown initial skill levels. In my model, students with more initial skill gain more skill overall (see Theorem A.3). So if the teacher designs tasks for students with a distribution of initial skill levels, she may prefer to forgo low-skill students' development for high-skill students'. In particular, she may "leave low-skill students behind" by allowing them to delegate. Doing so would loosen the incentive constraints on high-skill students and allow the teacher to better optimize her curriculum for such students. Thus, the model speaks to a debate among real-world teachers: should they teach to the best, median, or worst student?

The model also speaks to issues beyond AI. For example, a student may "delegate" to other (non-AI) students by copying their answers on homework and exams. The student would delegate if the payoff from doing the work himself exceeded the payoff from copying—a choice similar to the one I study. Beyond education, the model embeds a choice similar to one faced by industrial policymakers: should production be in- or off-shored? Governments may want to incentivize domestic production so that their citizens build skill. Doing so would involve taxes and subsidies. In contrast, the teacher in my model designs incentives by changing the production technology, rather than providing transfers.

References

Adilov, N. and Cline, J. W. (2026). The Carrot or the Stick? An Economic Model of Optimal Course Difficulty, Student Effort and Cheating Using AI. *Studies in Microeconomics*, 14(1):95–117.

Anzai, Y. and Simon, H. A. (1979). The Theory of Learning by Doing. *Psychological Review*, 86(2):124–140.

Arrow, K. J. (1962). The Economic Implications of Learning by Doing. *Review of Economic Studies*, 29(3):155.

Bacher-Hicks, A. and Koedel, C. (2023). Estimation and interpretation of teacher value added in research applications. In *Handbook of the Economics of Education*, volume 6, pages 93–134. Elsevier.

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., and Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences*, 122(26):e2422633122.

Bastani, H. and Cachon, G. P. (2026). The Human-AI Contracting Paradox.

Becker, G. S. (1962). Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy*, 70(5, Part 2):9–49.

Becker, W. E. and Rosen, S. (1992). The learning effect of assessment and evaluation in high school. *Economics of Education Review*, 11(2):107–118.

Ben-Porath, Y. (1967). The Production of Human Capital and the Life Cycle of Earnings. *Journal of Political Economy*, 75(4, Part 1):352–365.

Betts, J. R. (1998). The Impact of Educational Standards on the Level and Distribution of Earnings. *American Economic Review*, 88(1):266–275.

Brueckner, J. K. and Raymon, N. (1983). Optimal production with learning by doing. *Journal of Economic Dynamics and Control*, 6:127–135.

Brynjolfsson, E., Li, D., and Raymond, L. (2025). Generative AI at Work. *Quarterly Journal of Economics*, 140(2):889–942.

Campos, C. and Singleton, J. D. (2026). AI Diffusion Gaps: Unequal Integration of AI Across K-12 Schools. NBER Working Paper w35347, National Bureau of Economic Research, Cambridge, MA.

Caosun, M. and Aral, S. (2026). The Augmentation Trap: AI Productivity and the Cost of Cognitive Offloading. arXiv preprint 2604.03501.

Chirikov, I. (2026). Artificial Intelligence and Grade Inflation. Higher Education Working Paper, CSHE.

Chirikov, I., Smirnov, I., and Kizilcec, R. F. (2026). Generative AI use and misuse call for assessment reform in higher education. *Science*, 392(6800):818–820.

Contractor, Z. and Reyes, G. (2026). Generative AI in Higher Education: Evidence from an Elite College. arXiv preprint 2508.00717.

Costrell, R. M. (1994). A Simple Model of Educational Standards. *American Economic Review*, 84(4):956–971.

Cunha, F. and Heckman, J. (2007). The Technology of Skill Formation. *American Economic Review*, 97(2):31–47.

Dillon, E. W., Jaffe, S., Immorlica, N., and Stanton, C. T. (2026). Shifting Work Patterns with Generative AI. *American Economic Review: Insights*, forthcoming.

Doss, C. J., Bozick, R., Schwartz, H. L., Chu, L., Rainey, L. R., Woo, A., Reich, J., and Dukes, J. (2025). AI Use in Schools Is Quickly Increasing but Guidance Lags Behind: Findings from the RAND Survey Panels. Technical report, RAND Corporation, Santa Monica, CA.

Ellison, G. and Pathak, P. (2025). Optimal School System and Curriculum Design: Theory and Evidence. NBER Working Paper w34091, National Bureau of Economic Research, Cambridge, MA.

Fudenberg, D. and Levine, D. K. (2026). Taking the Easy Way Out: AI, Self-Control, and Human Capital Formation.

Garicano, L. and Rayo, L. (2026). Training in the Age of AI: A Theory of Career Viability. CEPR Discussion Paper No. 20634, CEPR Press, Paris and London.

Heckman, J. J. (2000). Policies to foster human capital. *Research in Economics*, 54(1):3–56.

Huang, L. and Vishnoi, N. K. (2026). Path Dependence under Adaptive AI Delegation. arXiv preprint 2603.02950.

Jovanovic, B. and Nyarko, Y. (1996). Learning by Doing and the Choice of Technology. *Econometrica*, 64(6):1299.

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., and Maes, P. (2025). Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task.

Lazear, E. P. (2001). Educational Production. *Quarterly Journal of Economics*, 116(3):777–803.

Liu, G., Christian, B., Dumbalska, T., Bakker, M. A., and Dubey, R. (2026). AI Assistance Reduces Persistence and Hurts Independent Performance. arXiv preprint 2604.04721.

Mincer, J. (1962). On-the-Job Training: Costs, Returns, and Some Implications. *Journal of Political Economy*, 70(5, Part 2):50–79.

Novak, J. D. and Gowin, D. B. (1984). *Learning how to learn*. Cambridge University Press, New York.

Noy, S. and Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654):187–192.

Peterson, A. J. (2025). Training for Obsolescence? The AI-Driven Education Trap. arXiv preprint 2508.19625v2.

Poulidis, S., Bastani, H., and Bastani, O. (2025). Self-Regulated AI Use Hinders Long-Term Learning.

Rosen, S. (1972). Learning by Experience as Joint Production. *Quarterly Journal of Economics*, 86(3):366.

Schultz, T. W. (1960). Capital Formation by Education. *Journal of Political Economy*, 68(6):571–583.

Shaikh, H. (2025). Identifying a Cumulative Learning Technology: Evidence from Online Learning.

Shen, J. H. and Tamkin, A. (2026). How AI Impacts Skill Formation. arXiv preprint 2601.20245.

Siderius, J., Shumsky, R. A., and Taylor, A. (2026). Use it or Slowly Lose it: Expertise Atrophy with Organizational AI Usage.

Strömberg, D., Lei, V., and Wu, Y. (2026). The Generative AI Learning Penalty: Evidence from Chinese Secondary Education. CEPR Discussion Paper No. 21577, CEPR Press, Paris and London.

Thompson, P. (2010). Learning by Doing. In *Handbook of the Economics of Innovation*, volume 1, pages 429–476. Elsevier.

A Proofs

For each result presented in Sections 2–4, I present general versions that allow for the AI quality parameter $q \geq 1$ (introduced in Section 5) to exceed one.

A.1 Proof of Lemma 1

Lemma 1 is the special case of Lemma A.1 with $q = 1$.

Lemma A.1. *Given a skill level $s > 0$ and effort intensity $\eta \in (0, 1]$, the student exerts best-response effort $x^*(s, \eta; q)$ defined as in* (10). *This effort is (i) positive, (ii) increasing in the productivity parameter π, (iii) increasing in s when $\eta < 1$ and constant in s when $\eta = 1$, and (iv) increasing in η if*

$$\eta \exp\left(\frac{2}{\eta} - 1\right) > \frac{s}{\pi q^2} \tag{A.1}$$

and non-increasing in η otherwise.

Proof. The student's payoff $u(x, s, \eta; q)$ is concave in x, so his best-response effort $x^* \equiv x^*(s, \eta; q)$ satisfies the first-order condition

$$0 = \frac{\partial u(x^*, s, \eta; q)}{\partial x} = \pi \eta q^\eta (x^*)^{\eta - 1} s^{1-\eta} - x^*, \tag{A.2}$$

which can be rearranged for

$$x^* = s\left(\frac{\pi \eta q^\eta}{s}\right)^{1/(2-\eta)}.$$

Clearly x^* is positive since s, η, π, and q are positive.

Now $x \mapsto \log(x)$ is an increasing transformation on $(0, \infty)$, so x^* is increasing in a parameter if and only if

$$\log(x^*) = \frac{1}{2-\eta}\log(\pi\eta) + \frac{\eta}{2-\eta}\log(q) + \frac{1-\eta}{2-\eta}\log(s)$$

is increasing in that parameter. Differentiating with respect to π gives

$$\frac{\partial \log(x^*)}{\partial \pi} = \frac{1}{\pi(2-\eta)} > 0.$$

Differentiating with respect to s gives

$$\frac{\partial \log(x^*)}{\partial s} = \frac{1-\eta}{s(2-\eta)} \geq 0$$

with equality if and only if $\eta = 1$. Differentiating with respect to η gives

$$\begin{aligned}
\frac{\partial \log(x^*)}{\partial \eta} &= \frac{1}{(2-\eta)^2}\log(\pi\eta) + \frac{1}{\eta(2-\eta)} + \frac{2}{(2-\eta)^2}\log(q) - \frac{1}{(2-\eta)^2}\log(s) \\
&= \frac{1}{(2-\eta)^2}\left(\frac{2-\eta}{\eta} - \log\left(\frac{s}{\pi\eta q^2}\right)\right),
\end{aligned}$$

which is positive if and only if (A.1) holds. □

A.2 Proof of Lemma 2

Lemma 2 is the special case of Lemma A.2 with $q = 1$.

Lemma A.2. *Given a skill level $s > 0$ and effort intensity $\eta \in (0, 1]$, the student's best-response effort $x^*(s, \eta; q)$ yields payoff $u^*(s, \eta; q)$ defined as in* (12). *This payoff is (i) positive, (ii) increasing in the productivity parameter π, (iii) increasing in s when $\eta < 1$ and constant in s when $\eta = 1$, and (iv) decreasing in η when $\eta < s/\pi q^2$ and non-decreasing in η otherwise.*

Proof. By Lemma A.1, the best-response effort $x^* \equiv x^*(s, \eta; q)$ yields payoff

$$\begin{aligned} u^*(s, \eta; q) &\equiv u(x^*, s, \eta; q) \\ &= \pi(qx^*)^\eta s^{1-\eta} - \frac{(x^*)^2}{2} \\ &\overset{\star}{=} \frac{1}{\eta}\left(1 - \frac{\eta}{2}\right)(x^*)^2 \end{aligned}$$

where $\star$ uses the substitution $\pi\eta(qx^*)^\eta s^{1-\eta} = (x^*)^2$ implied by the first-order condition (A.2). Parts (i)–(iii) follow immediately from Lemma A.1. For part (iv), we have

$$\begin{aligned} \frac{\partial x^*}{\partial \eta} &= x^* \frac{\partial \log(x^*)}{\partial \eta} \\ &= \frac{x^*}{(2-\eta)^2}\left(\frac{2-\eta}{\eta} - \log\left(\frac{s}{\pi\eta q^2}\right)\right) \end{aligned}$$

from the proof of Lemma A.1 and therefore

$$\begin{aligned} \frac{\partial u^*(s, \eta; q)}{\partial \eta} &= -\frac{1}{\eta^2}(x^*)^2 + \frac{2}{\eta}\left(1 - \frac{\eta}{2}\right)x^* \frac{\partial x^*}{\partial \eta} \\ &= -\frac{(x^*)^2}{\eta(2-\eta)} \log\left(\frac{s}{\pi\eta q^2}\right), \end{aligned}$$

which is positive if and only if $\eta > s/\pi q^2$. □

A.3 Proof of Lemma 3

Lemma 3 is the special case of Lemma A.3 with $q = 1$.

Lemma A.3. *For all skill levels $s > 0$, there is a unique $\eta^{FB}(s;q) \in (0,1]$ such that:*

(i) if $s \leq \pi e q^2$, then $\eta^{FB}(s;q) = 1$;

(ii) if $s > \pi e q^2$, then $\eta^{FB}(s;q) < 1$ satisfies (11)*, and is decreasing in s and increasing in q;*

(iii) the best-response effort (10) *attains its maximum over $\eta \in (0,1]$ when $\eta = \eta^{FB}(s;q)$.*

Moreover, $\eta^{FB}(s;q)$ is continuous in both s and q, and tends to zero as $s \to \infty$.

Proof. By Lemma A.1, the best-response effort $x^*(s,\eta;q)$ is increasing in η when (A.1) holds and non-increasing otherwise. The left-hand side of (A.1) falls continuously from ∞ to e as η rises from zero to one. So there are three cases to consider:

1. Suppose $s < \pi e q^2$. Then (A.1) holds for all $\eta \in (0,1)$ and so $x^*(s,\eta;q)$ attains its maximum only when $\eta = \eta^{\mathrm{FB}}(s;q) \equiv 1$.

2. Now suppose $s = \pi e q^2$. Then the proof of Lemma A.1 implies

$$\frac{\partial x^*(s,\eta;q)}{\partial \eta} > 0 \text{ if and only if } \frac{2-\eta}{\eta} > 0,$$

   which holds for all $\eta \in (0,1]$. So $x^*(s,\eta;q)$ is increasing in η and attains its maximum only when $\eta = \eta^{\mathrm{FB}}(s;q) \equiv 1$.

3. Finally, suppose $s > \pi e q^2$. Then, by the intermediate value theorem, there is a unique effort intensity $\eta^{\mathrm{FB}}(s;q) \in (0,1)$ satisfying (11) that maximizes $x^*(s,\eta;q)$ over $\eta \in (0,1]$. The left-hand side of (11) is decreasing in $\eta^{\mathrm{FB}}(s;q)$, while the right-hand side of (11) is increasing in s and decreasing in q. It follows that $\eta^{\mathrm{FB}}(s;q)$ is decreasing in s and increasing q. Moreover, the partial derivatives of the left- and right-hand sides of (11) are bounded away from zero, and so the implicit function implies $\eta^{\mathrm{FB}}(s;q)$ is differentiable (and, thus, continuous) in s and q when $s > \pi e q^2$. We also have $\eta^{\mathrm{FB}}(s;q) \to 1$ as $s \to \pi e q^2$ from above, from which it follows that $\eta^{\mathrm{FB}}(s;q)$ is globally continuous in both s and q. □

Finally, the right-hand side of (11) grows without bound as $s \to \infty$, but the left-hand side diverges to ∞ only as $\eta^{\mathrm{FB}}(s;q) \to 0$. It follows that $\eta^{\mathrm{FB}}(s;q) \to 0$ as $s \to \infty$.

A.4 Proof of Theorem 1

In the general setting described in Section 5, the student gains skill

$$s_{t+1} - s_t = \begin{cases} r\, x^*(s_t, \eta_t; q) & \text{if (IC; } q\text{) holds} \\ 0 & \text{if otherwise} \end{cases} \tag{A.3}$$

on each task $t \in \mathcal{T}$. The teacher chooses $(\eta_t)_{t\in\mathcal{T}}$ to solve

$$\max_{(\eta_t)_{t\in\mathcal{T}}} (s_T - s_t) \text{ subject to (A.3) and } 0 < \eta_t \leq 1 \text{ for each } t \in \mathcal{T}. \tag{P; q}$$

Following the logic at the beginning of Section 4, for each task $t \in \mathcal{T}$ we have

$$\eta_t \in \arg\max_{\eta\in(0,1]} x^*(s_t, \eta; q) \text{ subject to } u^*(s_t, \eta; q) \geq q\bar{u}. \tag{A.4}$$

Letting $q = 1$ recovers (6). So Theorem 1 is the special case of Theorem A.1 with $q = 1$.

Theorem A.1. *Let $\bar{u} = 0$ and suppose $(\eta_t)_{t\in\mathcal{T}}$ solves* (P; q). *Then $\eta_t = \eta^{FB}(s_t; q)$ for each $t \in \mathcal{T}$. Moreover, there exists $t^{FB} \geq 0$ such that if $t < t^{FB}$, then $\eta_t = 1$, and if $t \geq t^{FB}$, then $\eta_{t+1} < \eta_t$.*

Proof. Lemma A.3 implies

$$\arg\max_{\eta\in(0,1]} x^*(s_t, \eta; q) = \left\{\eta^{\text{FB}}(s_t; q)\right\}$$

for each $t \in \mathcal{T}$. Moreover, since $\bar{u} = 0$, Lemma A.2 implies $u^*(s_t, \eta; q) > q\bar{u}$ for all $\eta \in (0, 1]$. It follows from (A.4) that $\eta_t = \eta^{\text{FB}}(s_t; q)$ for each $t \in \mathcal{T}$.

Now $s_t \leq s_{t+1}$ for each $t \in \mathcal{T}$. So if $s_T \leq \pi e q^2$, then Lemma A.3 implies $\eta_t = 1$ for each $t \in \mathcal{T}$ and letting $t^{\text{FB}} = T$ yields the result. In contrast, if $s_T > \pi e q^2$, then there is a least $t^{\text{FB}} \in \mathcal{T}$ such that $s_t > \pi e q^2$ for each $t \geq t^{\text{FB}}$, and Lemma A.3 then implies $\eta_t = 1$ when $t < t^{\text{FB}}$ and $\eta_{t+1} < \eta_t$ when $t \geq t^{\text{FB}}$. □

A.5 Proof of Proposition 1

Proposition 1 is the special case of Proposition A.1 with $q = 1$.

Proposition A.1. *The first-best effort $x^{FB}(s; q)$ and payoff $u^{FB}(s; q)$ defined in Section 5.1 are continuous in s, constant in s when $s \leq \pi e q^2$, and increasing in s when $s > \pi e q^2$.*

Proof. Suppose $s \leq \pi e q^2$. Then $\eta^{\text{FB}}(s; q) = 1$ by Lemma A.3, and so $x^{\text{FB}}(s; q) = x^*(s, 1; q) = \pi q$ and $u^{\text{FB}}(s; q) = u^*(s, 1; q) = \pi^2 q^2/2$ are constant in s.

Now suppose $s > \pi e q^2$. Then $\eta^{\text{FB}}(s; q) < 1$ is the interior solution to

$$\max_{\eta} x^*(s, \eta; q) \text{ subject to } 0 < \eta \leq 1.$$

So the envelope theorem implies

$$\begin{aligned}
\frac{\partial x^{\text{FB}}(s; q)}{\partial s} &= \left.\frac{\partial x^*(s, \eta; q)}{\partial s}\right|_{\eta=\eta^{\text{FB}}(s;q)} \\
&= \left.\frac{1-\eta}{2-\eta}\left(\frac{\pi\eta q^\eta}{s}\right)^{1/(2-\eta)}\right|_{\eta=\eta^{\text{FB}}(s;q)} \\
&= \frac{1-\eta^{\text{FB}}(s; q)}{2-\eta^{\text{FB}}(s; q)} \cdot \frac{x^{\text{FB}}(s; q)}{s} \\
&> 0.
\end{aligned}$$

Thus $x^{\text{FB}}(s;q)$ is increasing in s. Then, by Lemmas A.2 and A.3, the best-response payoff

$$u^{\text{FB}}(s;q) = \frac{1}{\eta^{\text{FB}}(s;q)} \cdot \left(1 - \frac{\eta^{\text{FB}}(s;q)}{2}\right) \cdot (x^{\text{FB}}(s;q))^2$$

equals the product of three factors that are increasing in s. Thus $u^{\text{FB}}(s;q)$ is also increasing in s.

Finally, since $x^{\text{FB}}(s;q)$ and $u^{\text{FB}}(s;q)$ are continuous functions of s and $\eta^{\text{FB}}(s;q)$, and $\eta^{\text{FB}}(s;q)$ is continuous in s, both $x^{\text{FB}}(s;q)$ and $u^{\text{FB}}(s;q)$ are continuous in s. □

A.6 Proof of Lemma 4

Lemma 4 is the special case of Lemma A.4 with $q = 1$.

Lemma A.4.

(i) If (SB; q) *holds, then* $\pi^2 q^2/2 < q\bar{u}$.

(ii) If (SB; q) *and* (ND; q) *hold, then* $q\bar{u} < \pi s_0$.

Proof. If $s_0 \leq \pi e q^2$, then $u^{\text{FB}}(s_0;q) = u^*(s_0,1;q) = \pi^2 q^2/2$. If $s_0 > \pi e q^2$, then $\eta^{\text{FB}}(s_0;q) < 1 < e < s_0/\pi q^2$ and so $u^{\text{FB}}(s_0;q) = u^*(s_0,\eta^{\text{FB}}(s_0;q);q) > u^*(s_0,1;q) = \pi^2 q^2/2$ by Lemma A.2. Thus $u^{\text{FB}}(s_0;q) \geq \pi^2 q^2/2$ independently of s_0. So if (SB; q) holds, then $\pi^2 q^2/2 < q\bar{u}$.

Now suppose (SB; q) and (ND; q) hold. Then, by Lemma A.2, the best-response payoff $u^*(s_0,\eta;q)$ is decreasing in η when $\eta < s_0/\pi q^2$ and increasing in η when $\eta > s_0/\pi q^2$. It follows that

$$\begin{aligned}\sup_{\eta\in(0,1]} u^*(s_0,\eta;q) &= \max\left\{\lim_{\eta\to 0} u^*(s_0,\eta;q),\ u^*(s_0,1;q)\right\} \\ &= \max\left\{\pi s_0,\ \frac{\pi^2 q^2}{2}\right\}. \end{aligned} \tag{A.5}$$

But $\pi^2 q^2/2 < q\bar{u}$ from above. So if $\pi s_0 \leq \pi^2 q^2/2$, then (A.5) contradicts (ND; q). Thus $\pi s_0 > \pi^2 q^2/2$ and therefore

$$q\bar{u} < \sup_{\eta\in(0,1]} u^*(s_0,\eta;q) = \pi s_0.$$ □

A.7 Proof of Lemma 5

Lemma 5 is the special case of Lemma A.5 with $q = 1$.

Lemma A.5. *Suppose* (SB; q) *and* (ND; q) *hold. For all $s > q\overline{u}/\pi$, there is a unique*

$$\eta^{IC}(s;q) \in \left(0, \min\left\{\frac{s}{\pi q^2}, 1\right\}\right)$$

such that $u^*(s, \eta^{IC}(s;q);q) = q\overline{u}$. The mapping $s \mapsto \eta^{IC}(s;q)$ is continuous and increasing on its domain, and

$$\lim_{s \to \overline{u}/\pi q^2} \eta^{IC}(s;q) = 0 \text{ and } \lim_{s \to \infty} \eta^{IC}(s;q) = 1.$$

Moreover, the incentive constraint (IC; q) *holds if and only if $\eta_t \leq \eta^{IC}(s_t;q)$.*

Proof. Suppose (SB; q) and (ND; q) hold. By Lemma A.2, the best-response payoff $u^*(s, \eta; q)$ falls continuously from πs to

$$\min_{\eta \in (0,1]} u^*(s, \eta; q) = \frac{\pi^2 q^2}{2} \begin{cases} \left(2 - \frac{s}{\pi q^2}\right) \frac{s}{\pi q^2} & \text{if } \frac{s}{\pi q^2} < 1 \\ 1 & \text{if } \frac{s}{\pi q^2} \geq 1 \end{cases}$$

as η rises from zero to

$$\eta_{\min}(s;q) \equiv \min\left\{\frac{s}{\pi q^2}, 1\right\},$$

then rises continuously to $u^*(s, 1; q) = \pi^2 q^2/2$ as η rises to one. So, by Lemma A.4 and the intermediate value theorem, there is a unique effort intensity

$$\eta^{\mathrm{IC}}(s;q) \in (0, \eta_{\min}(s;q))$$

such that $u^*(s, \eta^{\mathrm{IC}}(s;q); q) = q\overline{u}$. Moreover, since $u^*(s, \eta; q)$ is decreasing in η over $(0, \eta_{\min}(s;q))$ and bounded above by $\pi^2 q^2/2$ over $[\eta_{\min}(s;q), 1]$, the incentive constraint $u^*(s, \eta; q) \geq q\overline{u}$ holds if and only if $\eta \leq \eta^{\mathrm{IC}}(s;q)$.

Next, I show $\eta^{\mathrm{IC}}(s;q)$ is continuous and increasing in $s > q\overline{u}/\pi$. Now $u^*(s, \eta; q)$ is continuously differentiable in s and η, and Lemma A.2 implies

$$\left.\frac{\partial u^*(s, \eta; q)}{\partial \eta}\right|_{\eta = \eta^{\mathrm{IC}}(s;q)} < 0$$

because $0 < \eta^{\mathrm{IC}}(s;q) < \eta_{\min}(s;q)$. So, by the implicit function theorem, the mapping $s \mapsto \eta^{\mathrm{IC}}(s;q)$ is differentiable and, thus, continuous. Its derivative $\partial \eta^{\mathrm{IC}}(s;q)/\partial s$ satisfies

$$\left.\frac{\partial u^*(s, \eta; q)}{\partial s}\right|_{\eta = \eta^{\mathrm{IC}}(s;q)} + \left.\frac{\partial u^*(s, \eta; q)}{\partial \eta}\right|_{\eta = \eta^{\mathrm{IC}}(s;q)} \frac{\partial \eta^{\mathrm{IC}}(s;q)}{\partial s} = 0,$$

obtained by applying the chain rule to the identity $u^*(s, \eta^{\mathrm{IC}}(s;q); q) = q\overline{u}$. But Lemma A.2 implies

$$\left.\frac{\partial u^*(s, \eta; q)}{\partial s}\right|_{\eta = \eta^{\mathrm{IC}}(s;q)} > 0 \text{ and } \left.\frac{\partial u^*(s, \eta; q)}{\partial \eta}\right|_{\eta = \eta^{\mathrm{IC}}(s;q)} < 0,$$

from which it follows that $\partial\eta^{\mathrm{IC}}(s;q)/\partial s > 0$.

Finally, I use the identity $u^*(s,\eta^{\mathrm{IC}}(s;q);q) = q\overline{u}$ to derive the limits of $\eta^{\mathrm{IC}}(s;q)$ as $s \to q\overline{u}/\pi$ and $s \to \infty$. We must have $\eta^{\mathrm{IC}}(s;q) \to 0$ as $s \to q\overline{u}/\pi$ from above, since $u^*(s,\eta;q) \to \pi s$ only as $\eta \to 0$. We must also have $\eta^{\mathrm{IC}}(s;q) \to 1$ as $s \to \infty$, since $u^*(s,\eta;q)$ grows for all $\eta < 1$ (by Lemma A.2) but $q\overline{u}$ does not change. $\square$

A.8 Proof of Lemma 6

Lemma 6 is the special case of Lemma 9 with $q = 1$. I prove Lemma 9 in Section A.15.

A.9 Proof of Theorem 2

Theorem 2 is the special case of Theorem A.2 with $q = 1$.

Theorem A.2. *Suppose* (SB;q) *and* (ND;q) *hold, and* $(\eta_t)_{t\in\mathcal{T}}$ *solves* (P;q). *Then* $\eta_t = \eta^{SB}(s_t;q)$ *for each* $t \in \mathcal{T}$. *Moreover, there exists* $t^{SB} \in \mathcal{T}$ *such that if* $t^{SB} > 0$ *and* $0 < t \leq t^{SB}$, *then* $\eta_t > \eta_{t-1}$, *and if* $t > t^{SB}$, *then* $\eta_t < \eta_{t-1}$.

Proof. By Lemma A.4, we have $s_t > q\overline{u}/\pi$ for each $t \in \mathcal{T}$.

Fix $t \in \mathcal{T}$. Lemma A.1 implies $x^*(s_t,\eta;q)$ is increasing in η when $\eta < \eta^{\mathrm{FB}}(s_t;q)$ and decreasing in η when $\eta > \eta^{\mathrm{FB}}(s_t;q)$. But, by Lemma A.5, we have $u^*(s_t,\eta;q) \geq q\overline{u}$ if and only if $\eta \leq \eta^{\mathrm{IC}}(s_t;q)$. It follows from (A.4) that $\eta_t = \eta^{\mathrm{SB}}(s_t;q)$. Then $u^{\mathrm{SB}}(s_t;q) \equiv u^*(s_t,\eta^{\mathrm{SB}}(s_t;q);q) \geq q\overline{u}$ because $\eta_t = \eta^{\mathrm{SB}}(s_t;q)$ satisfies (IC;q).

Now define $\overline{s}(q)$ as in Lemma 9. If $s_0 \geq \overline{s}(q)$, then $\eta^{\mathrm{SB}}(s_0;q) = \eta^{\mathrm{FB}}(s_0;q)$ and therefore $q\overline{u} \leq u^{\mathrm{SB}}(s_0;q) = u^{\mathrm{FB}}(s_0;q) < q\overline{u}$, contradicting (SB;$q$). Hence $s_0 < \overline{s}(q)$. Moreover, since each η_t satisfies the incentive constraint (IC;q), the student works on each task and Lemma A.1 implies $s_{t+1} > s_t$ for each $t \in \mathcal{T}$. So there is a maximal task $t^{\mathrm{SB}} \in \mathcal{T}$ such that $s_t \leq \overline{s}(q)$ if $t \leq t^{\mathrm{SB}}$ and $s_t > \overline{s}(q)$ if $t > t^{\mathrm{SB}}$. Then Lemma 9 implies $\eta_t = \eta^{\mathrm{SB}}(s_t;q) < \eta^{\mathrm{SB}}(s_{t-1};q) = \eta_{t-1}$ for each $t > t^{\mathrm{SB}}$. Moreover, if $t^{\mathrm{SB}} > 0$ and $0 < t \leq t^{\mathrm{SB}}$, then Lemma 9 implies $\eta_t = \eta^{\mathrm{SB}}(s_t;q) > \eta^{\mathrm{SB}}(s_{t-1};q) = \eta_{t-1}$. $\square$

A.10 Proof of Proposition 2

Proposition 2 follows immediately from letting $q = 1$ in the statement of Proposition A.2.

Proposition A.2. *Suppose* (SB; q) *and* (ND; q) *hold, let* $s > q\overline{u}/\pi$, *and define* $\overline{s}(q)$ *as in Lemma 9. Then the second-best effort* $x^{\text{SB}}(s;q)$ *is increasing in* s *and* π, *decreasing in* $\overline{u}$ *when* $s < \overline{s}(q)$, *and constant in* $\overline{u}$ *when* $s \geq \overline{s}(q)$.

Proof. We have $\eta^{\text{SB}}(s;q) = \eta^{\text{IC}}(s;q)$ when $s < \overline{s}(q)$ and $\eta^{\text{SB}}(s;q) = \eta^{\text{FB}}(s;q)$ when $s \geq \overline{s}(q)$. I consider these two cases separately:

1. Suppose $s < \overline{s}(q)$. Then $\eta^{\text{SB}}(s;q) = \eta^{\text{IC}}(s;q)$ and hence

$$\begin{aligned} q\overline{u} &= u^*(s, \eta^{\text{IC}}(s;q); q) \\ &= \frac{1}{\eta^{\text{IC}}(s;q)}\left(1 - \frac{\eta^{\text{IC}}(s;q)}{2}\right)\left(x^{\text{SB}}(s;q)\right)^2, \end{aligned}$$

which can be rearranged for

$$x^{\text{SB}}(s;q) = \sqrt{\frac{2\eta^{\text{IC}}(s;q)q\overline{u}}{2 - \eta^{\text{IC}}(s;q)}}.$$

Differentiating with respect to s then gives

$$\frac{\partial x^{\text{SB}}(s;q)}{\partial s} = h(s;q)\frac{\partial \eta^{\text{IC}}(s;q)}{\partial s},$$

where I define

$$h(s;q) \equiv \frac{2q\overline{u}}{x^{\text{SB}}(s;q)(2 - \eta^{\text{IC}}(s;q))^2} > 0$$

for convenience. But Lemma A.5 implies $\partial \eta^{\text{IC}}(s;q)/\partial s > 0$. Thus $x^{\text{SB}}(s;q)$ is increasing in s. Now Lemma A.2 implies

$$\left.\frac{\partial u^*(s,\eta;q)}{\partial \pi}\right|_{\eta=\eta^{\text{IC}}(s;q)} > 0 \text{ and } \left.\frac{\partial u^*(s,\eta;q)}{\partial \eta}\right|_{\eta=\eta^{\text{IC}}(s;q)} < 0$$

because $\eta^{\text{IC}}(s;q) < s/\pi q^2$. So differentiating $u^*(s, \eta^{\text{IC}}(s;q); q) = q\overline{u}$ with respect to π and rearranging the result gives

$$\frac{\partial \eta^{\text{IC}}(s;q)}{\partial \pi} = -\left.\frac{\partial u^*(s,\eta;q)/\partial \pi}{\partial u^*(s,\eta;q)/\partial \eta}\right|_{\eta=\eta^{\text{IC}}(s;q)} > 0.$$

Thus

$$\frac{\partial x^{\text{SB}}(s;q)}{\partial \pi} = h(s;q)\frac{\partial \eta^{\text{IC}}(s;q)}{\partial \pi} > 0$$

and therefore $x^{\text{SB}}(s;q)$ is increasing in π. Finally, differentiating $u^*(s, \eta^{\text{IC}}(s;q); q) = q\overline{u}$ with respect to $\overline{u}$ and rearranging the result gives

$$\frac{\partial \eta^{\text{IC}}(s;q)}{\partial \overline{u}} = \left.\frac{q}{\partial u^*(s,\eta;q)/\partial \eta}\right|_{\eta=\eta^{\text{IC}}(s;q)} < 0.$$

Thus

$$\frac{\partial x^{\text{SB}}(s;q)}{\partial \overline{u}} = h(s;q)\frac{\partial \eta^{\text{IC}}(s;q)}{\partial \overline{u}} < 0$$

and therefore $x^{\text{SB}}(s;q)$ is decreasing in $\overline{u}$.

2. Now suppose $s \geq \bar{s}(q)$. Then $\eta^{\text{SB}}(s;q) = \eta^{\text{FB}}(s;q)$ and therefore $x^{\text{SB}}(s;q) = x^{\text{FB}}(s;q)$. It follows from Proposition A.1 that $x^{\text{SB}}(s;q)$ is increasing in s. Now

$$x^{\text{FB}}(s;q) = \max_{\eta \in (0,1]} x^*(s,\eta;q)$$

by definition and $\eta^{\text{FB}}(s;q) < 1$ by Lemma 6. So the envelope theorem and Lemma A.1 imply

$$\frac{\partial x^{\text{SB}}(s;q)}{\partial \pi} = \left.\frac{\partial x^*(s,\eta;q)}{\partial \pi}\right|_{\eta=\eta^{\text{FB}}(s;q)} > 0.$$

Thus $x^{\text{SB}}(s;q)$ is increasing in π. Finally, Lemmas A.2 and A.3 imply $u^*(s,\eta;q)$ and $\eta^{\text{FB}}(s;q)$ are both constant in $\bar{u}$, which implies $x^{\text{SB}}(s;q) = u^*(s,\eta^{\text{FB}}(s;q);q)$ is also constant in $\bar{u}$. □

A.11 Proof of Lemma 7

Lemma 7 is the special case of Lemma A.6 with $q = 1$.

Lemma A.6. *Suppose* (SB; q) *and* (ND; q) *hold, and* $(\eta_t)_{t\in\mathcal{T}}$ *solves* (P; q). *For each task* $t \in \mathcal{T}$, *the skill increment* $\Delta(s_t,\eta_t;q) \equiv s_{t+1} - s_t$ *is increasing in* s_0, π, *and* r, *but decreasing in* $\bar{u}$.

Proof. Define $x^{\text{SB}}(s;q) \equiv x^*(s,\eta^{\text{SB}}(s;q);q)$ for all $s > q\bar{u}/\pi$. Theorem A.2 implies $\Delta(s_t,\eta_t;q) = rx^{\text{FB}}(s_t;q)$ for each $t \in \mathcal{T}$. Moreover, Lemma A.4 implies $s_0 > q\bar{u}/\pi$.

I consider each parameter s_0, π, $\bar{u}$, and r separately:

(s_0) Proposition A.2 says $x^{\text{SB}}(s;q)$ is increasing in s. So if s_0 rises, then $x^{\text{SB}}(s_0;q)$ rises, which makes $s_1 = s_0 + rx^{\text{SB}}(s_0;q)$ rise, which makes $x^{\text{SB}}(s_1;q)$ rise, which makes $s_2 = s_1 + rx^{\text{SB}}(s_1;q)$ rise, and so on. Thus, if s_0 rises, then $x^{\text{SB}}(s_t;q)$ and $\Delta(s_t,\eta_t;q)$ rise for each $t \in \mathcal{T}$.

(π) Proposition A.2 says $x^{\text{SB}}(s;q)$ is increasing in s and π. So if π rises, then $x^{\text{SB}}(s_0;q)$ rises, which makes $s_1 = s_0 + rx^{\text{SB}}(s_0;q)$ rise, which together with π rising makes $x^{\text{SB}}(s_1;q)$ rise, which makes $s_2 = s_1 + rx^{\text{SB}}(s_1;q)$ rise, and so on. Thus, if π rises, then $x^{\text{SB}}(s_t;q)$ and $\Delta(s_t,\eta_t;q)$ rise for each $t \in \mathcal{T}$.

(r) Proposition A.2 says $x^{\text{SB}}(s;q)$ is increasing in s and Lemma A.1 implies $x^{\text{SB}}(s_0;q) > 0$. So if r rises, then $s_1 = s_0 + rx^{\text{SB}}(s_0;q)$ rises, which makes $x^{\text{SB}}(s_1;q)$ rise, which together with r rising makes $s_2 = s_1 + rx^{\text{SB}}(s_1;q)$ rise, and so on. Thus, if r rises, then $\Delta(s_t,\eta_t) = s_{t+1} - s_t$ rises for each $t \in \mathcal{T}$.

($\bar{u}$) Define $\bar{s}(q)$ as in Lemma 9. The proof of Theorem A.2 shows $s_0 \leq \bar{s}(q)$. Then Proposition A.2 implies $x^{\text{FB}}(s_0;q)$ is decreasing in $\bar{u}$, and $x^{\text{SB}}(s;q)$ is increasing in s and non-increasing in $\bar{u}$. So if $\bar{u}$ rises, then $x^{\text{SB}}(s_0;q)$ falls, which makes $s_1 = s_0 + rx^{\text{SB}}(s_0;q)$ fall, which together with $\bar{u}$ rising makes $x^{\text{SB}}(s_1;q)$ fall, which makes $s_2 = s_0 + rx^{\text{SB}}(s_1;q)$ fall, and so on. Thus, if $\bar{u}$ rises, then $x^{\text{SB}}(s_t;q)$ and $\Delta(s_t,\eta_t)$ fall for each $t \in \mathcal{T}$. □

A.12 Proof of Theorem 3

Theorem 3 is the special case of Theorem A.3 with $q = 1$.

Theorem A.3. *Suppose* (SB; q) *and* (ND; q) *hold, and* $(\eta_t)_{t\in\mathcal{T}}$ *solves* (P; q). *Then* $(s_T - s_0)$ *is increasing in* s_0, π, *and* r, *but decreasing in* $\overline{u}$.

Proof. The result follows immediately from Lemma A.6 and the fact that

$$s_T - s_0 = \sum_{t\in\mathcal{T}}(s_{t+1} - s_t).$$

□

A.13 Proof of Proposition 3

Suppose $s \leq \pi e q^2$. Then $\eta^{\text{FB}}(s;q) = 1$ by Lemma A.3, and so $x^{\text{FB}}(s;q) = x^*(s,1;q) = \pi q$ and $u^{\text{FB}}(s;q) = u^*(s,1;q) = \pi^2 q^2/2$ are increasing in q.

Now suppose $s > \pi e q^2$. Then $\eta^{\text{FB}}(s;q)$ satisfies (11), which can be rewritten as

$$\frac{1}{q^2}\exp\left(-\left(\frac{2}{\eta^{\text{FB}}(s;q)} - 1\right)\right) = \frac{\pi\eta^{\text{FB}}(s;q)}{s}. \tag{A.6}$$

Then Lemma A.1 implies

$$\begin{aligned} x^{\text{FB}}(s;q) &= s\left(\frac{\pi\eta^{\text{FB}}(s;q)q^{\eta^{\text{FB}}(s;q)}}{s}\right)^{1/(2-\eta^{\text{FB}}(s;q))} \\ &= s\left(\frac{1}{q^2}\exp\left(-\left(\frac{2}{\eta^{\text{FB}}(s;q)} - 1\right)\right)q^{\eta^{\text{FB}}(s;q)}\right)^{1/(2-\eta^{\text{FB}}(s;q))} \\ &= \frac{s}{q}\exp\left(-\frac{1}{\eta^{\text{FB}}(s;q)}\right), \end{aligned}$$

which is increasing in $\eta^{\text{FB}}(s;q)$. But Lemma A.3 implies $\eta^{\text{FB}}(s;q)$ is increasing in q, and it follows that $x^{\text{FB}}(s;q)$ is increasing in q. Moreover, Lemma A.2 implies

$$\begin{aligned} u^{\text{FB}}(s;q) &= \frac{2 - \eta^{\text{FB}}(s;q)}{2\eta^{\text{FB}}(s;q)}\left(x^{\text{FB}}(s;q)\right)^2 \\ &= \frac{\left(2 - \eta^{\text{FB}}(s;q)\right)s^2}{2\eta^{\text{FB}}(s;q)eq^2}\exp\left(-\left(\frac{2}{\eta^{\text{FB}}(s;q)} - 1\right)\right) \\ &\overset{\star}{=} \frac{\pi s}{2e}\left(2 - \eta^{\text{FB}}(s;q)\right), \end{aligned}$$

where $\star$ uses the substitution (A.6). So $u^{\text{FB}}(s;q)$ is decreasing in $\eta^{\text{FB}}(s;q)$, which is increasing in q. Thus $u^{\text{FB}}(s;q)$ is decreasing in q.[20] □

[20]Recall from Lemma A.3 that $\eta^{\text{FB}}(s;q) \to 1$ as $s \to \pi e q^2$ from above. So

$$\lim_{s\to\pi e q^2}\frac{\pi s}{2e}\left(2 - \eta^{\text{FB}}(s;q)\right) = \frac{\pi(\pi e q^2)}{2e} = \frac{\pi^2 q^2}{2} = u^{\text{FB}}(s;q)\Big|_{s\leq\pi e q^2},$$

confirming the continuity of $u^{\text{FB}}(s;q)$ in s as claimed in Proposition A.1.

A.14 Proof of Lemma 8

Differentiating (14) implicitly with respect to q gives

$$\left.\frac{\partial u^*(s,\eta;q)}{\partial \eta}\right|_{\eta=\eta^{\mathrm{IC}}(s;q)} \frac{\partial \eta^{\mathrm{IC}}(s;q)}{\partial q} + \left.\frac{\partial u^*(s,\eta;q)}{\partial q}\right|_{\eta=\eta^{\mathrm{IC}}(s;q)} = \overline{u}.$$

Since $u^*(s,\eta;q)$ is decreasing in η at $\eta = \eta^{\mathrm{IC}}(s;q)$, we have $\partial\eta^{\mathrm{IC}}(s;q)/\partial q > 0$ if and only if

$$\left.\frac{\partial u^*(s,\eta;q)}{\partial q}\right|_{\eta=\eta^{\mathrm{IC}}(s;q)} > \overline{u}. \tag{A.7}$$

But

$$\frac{\partial u^*(s,\eta;q)}{\partial q} = \frac{2\eta}{2-\eta}\cdot\frac{u^*(s,\eta;q)}{q},$$

and therefore

$$\left.\frac{\partial u^*(s,\eta;q)}{\partial q}\right|_{\eta=\eta^{\mathrm{IC}}(s;q)} = \frac{2\eta^{\mathrm{IC}}(s;q)\overline{u}}{2-\eta^{\mathrm{IC}}(s;q)}$$

by (14). So (A.7) holds if and only if

$$\frac{2\eta^{\mathrm{IC}}(s;q)}{2-\eta^{\mathrm{IC}}(s;q)} > 1,$$

which, since $0 < \eta^{\mathrm{IC}}(s;q) < 1$, holds precisely when $\eta^{\mathrm{IC}}(s;q) > 2/3$.

Since $\eta^{\mathrm{IC}}(s;q)$ rises continuously from zero to one as s rises from $q\overline{u}/\pi$ to ∞ (see Appendix Section A.7), there is a unique skill level $s^\dagger > 0$ such that $\eta^{\mathrm{IC}}(s^\dagger;q) = 2/3$ and

$$\frac{\partial \eta^{\mathrm{IC}}(s;q)}{\partial q} \gtreqless 0 \text{ as } s \gtreqless s^\dagger.$$

This skill level satisfies

$$\begin{aligned} q\overline{u} &= u^*\left(s^\dagger, \frac{2}{3}; q\right) \\ &= \frac{1}{2/3}\left(1-\frac{2/3}{2}\right)\left(x^*\left(s^\dagger,\frac{2}{3};q\right)\right)^2 \\ &= \left(s^\dagger\left(\frac{\pi(2/3)q^{2/3}}{s^\dagger}\right)^{1/(2-2/3)}\right)^2 \\ &= \left(\frac{2\pi}{3}\right)^{3/2} q(s^\dagger)^{1/2}, \end{aligned}$$

which can be rearranged for (15). □

A.15 Proof of Lemma 9

By Lemma A.5, the incentive-compatible effort intensity $\eta^{\mathrm{IC}}(s;q) < 1$ rises continuously from zero to one as s rises from $q\overline{u}/\pi$ to ∞. In contrast, by Lemma A.3, the first-best effort intensity $\eta^{\mathrm{FB}}(s;q)$ equals one when $s \le \pi e q^2$, and falls continuously from one to zero as s rises from $\pi e q^2$ to ∞. So, by the intermediate value theorem, there is a unique

$$\bar{s} \in \left(\max\left\{\frac{q\overline{u}}{\pi}, \pi e q^2\right\}, \infty\right)$$

such that $\eta^{\mathrm{IC}}(\bar{s};q) = \eta^{\mathrm{FB}}(\bar{s};q)$ and (17) holds for all $s > q\overline{u}/\pi$. It follows that $\eta^{\mathrm{SB}}(s;q)$ is increasing in s when $s < \bar{s}$ and decreasing in s when $s > \bar{s}$.

It remains to show $\bar{s}$ is increasing in q.

Define $\overline{\eta} \equiv \eta^{\mathrm{IC}}(\bar{s};q) = \eta^{\mathrm{FB}}(\bar{s};q)$ for convenience. Then (11) implies

$$\bar{s} = \pi\overline{\eta}q^2 \exp\left(\frac{2}{\overline{\eta}} - 1\right)$$

and the definition of η^{IC} implies $u^*(\bar{s},\overline{\eta};q) = q\overline{u}$. Lemmas A.1 and A.2 imply

$$\begin{aligned} u^*(s,\eta;q) &= \frac{2-\eta}{2\eta}\left(s\left(\frac{\pi\eta q^{\eta}}{s}\right)^{1/(2-\eta)}\right)^2 \\ &= \left(1-\frac{\eta}{2}\right)\pi^{2/(2-\eta)}\eta^{\eta/(2-\eta)}q^{2\eta/(2-\eta)}s^{2(1-\eta)/(2-\eta)} \end{aligned} \tag{A.8}$$

for all $s > 0$ and $\eta \in (0,1]$, so

$$\begin{aligned} q\overline{u} &= u^*(\bar{s},\overline{\eta};q) \\ &= \left(1-\frac{\overline{\eta}}{2}\right)\pi^{2/(2-\overline{\eta})}\overline{\eta}^{\overline{\eta}/(2-\overline{\eta})}q^{2\overline{\eta}/(2-\overline{\eta})}\left(\pi\overline{\eta}q^2\exp\left(\frac{2}{\overline{\eta}}-1\right)\right)^{2(1-\overline{\eta})/(2-\overline{\eta})} \\ &= \frac{\pi^2\overline{\eta}(2-\overline{\eta})q^2}{2}\exp\left(\frac{2(1-\overline{\eta})}{\overline{\eta}}\right). \end{aligned}$$

Defining

$$Q(\eta) \equiv \frac{2\overline{u}}{\pi^2\eta(2-\eta)}\exp\left(-\frac{2(1-\eta)}{\eta}\right)$$

then yields an implicit equation $Q(\overline{\eta}) = q$ defining $\overline{\eta}$ in terms of q. But the function $Q : (0,1] \to \mathbb{R}$ is increasing on its domain because $Q(\eta) > 0$ and

$$\begin{aligned} \frac{Q'(\eta)}{Q(\eta)} &= \frac{\partial \log(Q(\eta))}{\partial \eta} \\ &= \frac{\partial}{\partial \eta}\left[\log\left(\frac{2\overline{u}}{\pi^2}\right) - \log(\eta(2-\eta)) - \frac{2(1-\eta)}{\eta}\right] \\ &= \frac{2\left((\eta-1)^2+1\right)}{\eta^2(2-\eta)} \\ &> 0 \end{aligned}$$

for all $\eta \in (0,1]$. So if q rises, then $\overline{\eta}$ must also rise to maintain $Q(\overline{\eta}) = q$. □

A.16 Proof of Theorem 4

The result contains three claims:

1. There exists $s^{\ddagger} \in (q\overline{u}/\pi, \overline{s}(q))$ such that $x^{\mathrm{SB}}(s;q)$ is decreasing in q when $s < s^{\ddagger}$ and increasing in q when $s > s^{\ddagger}$.
2. $s^{\ddagger} < s^{\dagger}$.
3. $s^{\ddagger}$ is increasing in q.

I prove these claims below.

A.16.1 Proof of Claim 1

My proof of Claim 1 invokes the following lemma.

Lemma A.7. *Suppose* (SB; q) *and* (ND; q) *hold, define* $\overline{s}(q)$ *as in Lemma 9, let* $s \in (q\overline{u}/\pi, \overline{s}(q))$*, and define*

$$h(s,\eta;q) \equiv \log\left(\frac{s}{\pi\eta q^2}\right)$$

for all $\eta \in (0,1]$*. Then*

$$\frac{\partial\eta^{SB}(s;q)}{\partial q} = \frac{(2-\eta^{SB}(s;q))(3\eta^{SB}(s;q)-2)}{2q\,h(s,\eta^{SB}(s;q);q)}. \tag{A.9}$$

Proof. Now $\eta^{\mathrm{SB}}(s;q) = \eta^{\mathrm{IC}}(s;q)$ satisfies the implicit equation

$$u^*(s,\eta^{\mathrm{SB}}(s;q);q) = q\overline{u}$$

Differentiating this equation with respect to q gives

$$\left.\frac{\partial u^*(s,\eta;q)}{\partial\eta}\right|_{\eta=\eta^{\mathrm{SB}}(s;q)} \frac{\partial\eta^{\mathrm{SB}}(s;q)}{\partial q} + \left.\frac{\partial u^*(s,\eta;q)}{\partial q}\right|_{\eta=\eta^{\mathrm{SB}}(s;q)} = \overline{u}. \tag{A.10}$$

To evaluate $\partial\eta^{\mathrm{SB}}(s;q)/\partial q$, we need to evaluate $\partial u^*(s,\eta;q)/\partial\eta$ and $\partial u^*(s,\eta;q)/\partial q$ at $\eta = \eta^{\mathrm{SB}}(s;q)$.

From Lemma A.2 and its proof, we have

$$(x^*(s,\eta;q))^2 = \frac{2\eta u^*(s,\eta;q)}{2-\eta}$$

and

$$\begin{aligned}\frac{\partial u^*(s,\eta;q)}{\partial\eta} &= -\frac{(x^*(s,\eta;q))^2}{\eta(2-\eta)}\log\left(\frac{s}{\pi\eta q^2}\right)\\ &= -\frac{2u^*(s,\eta;q)\,h(s,\eta;q)}{(2-\eta)^2}.\end{aligned}$$

We also have

$$\log(u^*(s,\eta;q)) = \log\left(\frac{1}{\eta}\left(1-\frac{\eta}{2}\right)\right) + 2\log(x^*(s,\eta;q))$$

and hence

$$\begin{aligned}
\frac{1}{u^*(s,\eta;q)}\frac{\partial u^*(s,\eta;q)}{\partial q} &= \frac{\partial \log(u^*(s,\eta;q))}{\partial q} \\
&= 2\frac{\partial \log(x^*(s,\eta;q))}{\partial q} \\
&= 2\frac{\partial}{\partial q}\left[\frac{1}{2-\eta}\log(\pi\eta) + \frac{\eta}{2-\eta}\log(q) + \frac{1-\eta}{2-\eta}\log(s)\right] \\
&= \frac{2\eta}{(2-\eta)q}.
\end{aligned}$$

But $\eta^{\text{SB}}(s;q) = \eta^{\text{IC}}(s;q)$ and therefore $u^*(s,\eta^{\text{SB}}(s;q);q) = q\bar{u}$. Thus

$$\left.\frac{\partial u^*(s,\eta;q)}{\partial \eta}\right|_{\eta=\eta^{\text{SB}}(s;q)} = -\frac{2q\bar{u}\, h(s,\eta^{\text{SB}}(s;q);q)}{(2-\eta^{\text{SB}}(s;q))^2}$$

and

$$\left.\frac{\partial u^*(s,\eta;q)}{\partial q}\right|_{\eta=\eta^{\text{SB}}(s;q)} = \frac{2\eta^{\text{SB}}(s;q)\bar{u}}{2-\eta^{\text{SB}}(s;q)}.$$

Substituting these expressions into (A.10) and rearranging the result yields (A.9). □

Proof of Claim 1. Suppose $s \geq \bar{s}(q)$. Then $x^{\text{SB}}(s;q) = x^{\text{FB}}(s;q)$, which, by Proposition 3, is increasing in q.

Now suppose $s < \bar{s}(q)$. Then $\eta^{\text{SB}}(s;q) = \eta^{\text{IC}}(s;q) < 1$ satisfies

$$u^*(s,\eta^{\text{SB}}(s;q);q) = q\bar{u},$$

which, by Lemma A.2, implies

$$x^{\text{SB}}(s;q) = \sqrt{\frac{2\eta^{\text{SB}}(s;q)\; q\bar{u}}{2-\eta^{\text{SB}}(s;q)}}.$$

Differentiating with respect to q gives

$$\frac{\partial x^{\text{SB}}(s;q)}{\partial q} = \frac{2q\bar{u}}{x^{\text{SB}}(s;q)(2-\eta^{\text{SB}}(s;q))^2}\left(\frac{\eta^{\text{SB}}(s;q)(2-\eta^{\text{SB}}(s;q))}{2q} + \frac{\partial \eta^{\text{SB}}(s;q)}{\partial q}\right),$$

which is positive if and only if

$$\frac{\partial \eta^{\text{SB}}(s;q)}{\partial q} > -\frac{\eta^{\text{SB}}(s;q)(2-\eta^{\text{SB}}(s;q))}{2q}. \tag{A.11}$$

Our goal is to express (A.11) as a condition on the skill level s. Define $h(s,\eta;q)$ as in Lemma A.7 and

$$g(s) \equiv \eta^{\text{SB}}(s;q)\Big(3 + h(s,\eta^{\text{SB}}(s;q);q)\Big) - 2$$

for all $s \in (q\bar{u}/\pi, \bar{s}(q))$. Substituting (A.9) into (A.11) and rearranging the result gives

$$g(s) > 0. \tag{A.12}$$

Conditions (A.11) and (A.12) are equivalent, and so $x^{\text{SB}}(s;q)$ is increasing in q if and only if (A.12) holds. Now $g(s) \to -2$ as $s \to q\bar{u}/\pi$ from above because $\eta^{\text{SB}}(s;q)\, h(s, \eta^{\text{SB}}(s;q);q) \to 0$. Moreover, since $h(s,\eta;q)$ is continuous in its arguments and $\eta^{\text{SB}}(s;q) \to \eta^{\text{FB}}(s;q)$ as $s \to \bar{s}(q)$ from below, we have $h(s,\eta^{\text{SB}}(s;q);q) \to h(s,\eta^{\text{FB}}(s;q);q) = 2/\eta^{\text{FB}}(s;q) - 1$ and hence $g(s) \to 2\eta^{\text{FB}}(s;q) > 0$ as $s \to \bar{s}(q)$ from below. But $g : (q\bar{u}/\pi, \bar{s}(q)) \to \mathbb{R}$ is differentiable (and thus continuous) on its domain, and its derivative

$$g'(s) = \frac{\partial \eta^{\text{SB}}(s;q)}{\partial s}\Big(2 + h(s, \eta^{\text{SB}}(s;q);q)\Big) + \frac{\eta^{\text{SB}}(s;q)}{s}$$

is positive because $\eta^{\text{SB}}(s;q) = \eta^{\text{IC}}(s;q)$ is increasing in s (see Lemma 5). Thus, the value $g(s)$ rises continuously from a negative number to a positive number as s rises from the $q\bar{u}/\pi$ to $\bar{s}(q)$. So, by the intermediate value theorem, there is a unique skill level

$$s^\ddagger \in \left(\frac{q\bar{u}}{\pi}, \bar{s}(q)\right)$$

such that $g(s^\ddagger) = 0$, and (A.12) holds if and only if $s > s^\ddagger$. It follows that $x^{\text{SB}}(s;q)$ is decreasing in q when $s < s^\ddagger$ and increasing in q when $s > s^\ddagger$. □

A.16.2 Proof of Claim 2

Suppose $s^\dagger < \bar{s}(q)$, and define $g(s)$ and $h(s,\eta;q)$ as in the proof of Claim 1. Then Lemma A.5 implies

$$\begin{aligned}
g(s^\dagger) &= \eta^{\text{SB}}(s^\dagger;q)\Big(3 + h(s^\dagger, \eta^{\text{SB}}(s^\dagger;q);q)\Big) - 2 \\
&= \frac{2}{3}\left(3 + h\left(s^\dagger, \frac{2}{3}; q\right)\right) - 2 \\
&= \frac{2}{3} h\left(s^\dagger, \frac{2}{3}; q\right) \\
&> 0,
\end{aligned}$$

and therefore $s^\ddagger < s^\dagger$ because $g(s)$ is increasing in s and $g(s^\ddagger) = 0 < g(s^\dagger)$.

Now suppose $s^\dagger \geq \bar{s}(q)$. Then $s^\ddagger < \bar{s}(q) \leq s^\dagger$ and thus $s^\ddagger < s^\dagger$. □

A.16.3 Proof of Claim 3

Define $g(s)$ as in the proof of Claim 1 and define $\eta^\ddagger \equiv \eta^{\text{SB}}(s^\ddagger;q) \in (0, 2/3)$ for convenience. Then

$$0 = g(s^\ddagger) = \eta^\ddagger\left(3 + \log\left(\frac{s^\ddagger}{\pi\eta^\ddagger q^2}\right)\right) - 2,$$

which can be rewritten as

$$s^\ddagger = \pi\eta^\ddagger q^2 \exp\left(\frac{2}{\eta^\ddagger} - 3\right). \tag{A.13}$$

Now $s \mapsto \log(s)$ is an increasing transformation on $(0, \infty)$, so $s^\ddagger$ is increasing in q if and only if

$$\begin{aligned}\frac{\partial \log(s^\ddagger)}{\partial q} &= \frac{\partial}{\partial q}\left[\log(\pi) + \log(\eta^\ddagger) + 2\log(q) + \frac{2}{\eta^\ddagger} - 3\right] \\ &= \frac{1}{\eta^\ddagger}\left(1 - \frac{2}{\eta^\ddagger}\right)\frac{\partial \eta^\ddagger}{\partial q} + \frac{2}{q} \end{aligned} \tag{A.14}$$

is positive.

To determine (A.14), we need to determine $\partial\eta^\ddagger/\partial q$. Substituting $(s, \eta) = (s^\ddagger, \eta^\ddagger)$ and (A.13) into (A.8) gives

$$u^*(s^\ddagger, \eta^\ddagger; q) = \left(1 - \frac{\eta^\ddagger}{2}\right)\pi^2\eta^\ddagger q^2 \exp\left(\frac{2(2 - 3\eta^\ddagger)(1 - \eta^\ddagger)}{\eta^\ddagger(2 - \eta^\ddagger)}\right).$$

But $u^*(s^\ddagger, \eta^\ddagger; q) = q\bar{u}$ by definition. Combining these two expressions yields an implicit equation

$$q = Q(\eta^\ddagger) \equiv \frac{2}{\pi^2\eta^\ddagger(2 - \eta^\ddagger)} \exp\left(-\frac{2(2 - 3\eta^\ddagger)(1 - \eta^\ddagger)}{\eta^\ddagger(2 - \eta^\ddagger)}\right) \tag{A.15}$$

defining $\eta^\ddagger$.

Now $Q : (0, 2/3) \to \mathbb{R}$ is differentiable on its domain. Denoting its derivative by Q' gives

$$\begin{aligned}\frac{Q'(\eta)}{Q(\eta)} &= \frac{\partial \log(Q(\eta))}{\partial \eta} \\ &= \frac{\partial}{\partial \eta}\left[\log\left(\frac{q}{\pi^2}\right) - \log(\eta(2 - \eta)) - \frac{2(2 - 3\eta)(1 - \eta)}{\eta(2 - \eta)}\right] \\ &= -\frac{2P_1(\eta)}{\eta^2(2 - \eta)^2},\end{aligned}$$

where $P_1(\eta) \equiv \eta^3 - 2\eta^2 + 6\eta - 4 < 0$ for all $\eta \in (0, 2/3)$. Differentiating (A.15) with respect to q and rearranging the result then gives

$$\begin{aligned}\frac{\partial \eta^\ddagger}{\partial q} &= \frac{1}{Q'(\eta^\ddagger)} \\ &= -\frac{(\eta^\ddagger)^2(2 - \eta^\ddagger)^2}{2qP_1(\eta^\ddagger)}\end{aligned} \tag{A.16}$$

because $Q(\eta^\ddagger) = q$. Moreover, substituting (A.16) into (A.14) gives

$$\begin{aligned}\frac{\partial \log(s^\ddagger)}{\partial q} &= \frac{1}{\eta^\ddagger}\left(1 - \frac{2}{\eta^\ddagger}\right)\left(-\frac{(\eta^\ddagger)^2(2 - \eta^\ddagger)^2}{2qP_1(\eta^\ddagger)}\right) + \frac{2}{q} \\ &= \frac{P_2(\eta^\ddagger)}{2qP_1(\eta^\ddagger)},\end{aligned}$$

where $P_2(\eta) \equiv 3\eta^3 - 2\eta^2 + 12\eta - 8 < 0$ for all $\eta \in (0, 2/3)$. Thus $\partial \log(s^\ddagger)/\partial q > 0$, and so $s^\ddagger$ is increasing in q. $\square$