Taking the Easy Way Out: AI, Self-Control, and Human Capital Formation*

Drew Fudenberg† David K. Levine‡

August 19, 2026

Abstract

Some researchers worry that AI use by junior researchers inhibits acquisition of human capital. We study a dynamic model with a short-term project with a direct payoff now and a long-term project with a larger direct payoff next period. AI increases success with low human capital, but prevents learning by doing. Increasing available options ordinarily cannot decrease welfare, but if direct payoffs are tempting and self-control is costly, AI reduces welfare even if it is not adopted, and can induce the researcher to switch to the short-term project. More surprisingly, AI may increase the acquisition of human capital.

*We thank Daron Acemoglu, Joshua Gans, Andrew Koh, and Stephanie Wang for helpful comments and suggestions, and NSF grant SES-2417162 and the Leverhulme Trust for financial support.

†Department of Economics, MIT

‡Department of Economics, RHUL and Washington University in St, Louis

1 Introduction

An AI agent may give a junior researcher a better chance of producing a useful output. But as Hogg [2026] and Karamanis [2026] emphasize, the value of long-term research work lies not only in the output it generates, but also in the skills and judgment acquired by doing it. We share their concern that by replacing hard research work, AI can weaken the process through which researchers acquire judgment and skill: A researcher who uses AI to generate results and text can produce a passable paper while giving up the learning by doing that would make future work better.

This paper studies the decision of whether to delegate to an AI agent in a simple dynamic model where the researcher can either delegate the task to an AI agent with minimal human input or not use it at all. Each period, the researcher chooses between a short-term project and a long-term one. The short-term project provides a direct payoff now, while the long-term project provides a direct payoff only later. In low-human-capital states, working without AI can generate learning by doing and lead to the high-human-capital state, whereas using AI reduces human-capital acquisition. In a standard model without a self-control problem, the availability of AI cannot lower the agent's average present value: They will not use it if doing so decreases their average present value, and if they do not use AI they are no worse off than if it were not available.

The welfare effect of AI is no longer straightforward when the agent is tempted to forgo future direct payoffs for current direct payoffs and so has a cost of self control. To model this, we assume that the cost of resisting temptation depends on the best direct payoff available in the current period, and that this cost may be incurred even if the AI is not used. By making a high-direct-payoff option available, AI raises the temptation benchmark. As a result, the availability of AI can worsen welfare and distort project choice.

The issue is not specific to academic research. Many valuable skills are acquired while performing tasks that are initially slow, frustrating, or only imperfectly productive. AI changes the private return to those tasks by making it possible to obtain acceptable output without going through the costly process of skill formation. Thus the central question is not only whether AI raises current productivity, but also whether it changes the incentives to acquire the human capital on which future productivity depends. This makes AI different from a pure productivity shock: it changes both what can be done today and what the worker has reason to learn for tomorrow.

The interaction of AI and self-control gives rise to four distinct effects.

- Blowing off: the researcher uses AI to carry out the long-term project in the low-human-capital state and forgoes the accumulation of human capital.
- Discouragement: when AI becomes available, the researcher uses AI to carry out the short-term project in the low-human-capital state even though the long-term project would have been chosen if AI were unavailable.
- Reversal: the arrival of AI causes the researcher to switch from the long-term project to the short-term project in the low-human-capital state even though AI is not used.
- Encouragement: when AI becomes available, the researcher acquires human capital even though, without AI, the researcher would have remained in the low-human-capital state.

The first three effects concern how AI changes behavior within the low human capital state. In blowing off, the researcher continues to choose the long-term project, but does so with AI and therefore gives up the learning by doing that would accompany working without AI. In discouragement, the arrival of AI changes project choice: long-term work would have been chosen without AI, but with AI available the researcher switches to the short-term project. In reversal, AI changes the ranking of the two non-AI tasks. The researcher switches from the long-term project to the short-term project even though AI is not used, because the availability of AI changes the self-control cost of doing long-term projects without it.

Encouragement concerns a different margin. Without AI, the researcher may choose a long-term project and still remain in the low-human-capital state. With AI, the average present value of remaining in that state changes, and the researcher may instead choose a path that leads to high human capital. Thus the analysis distinguishes effects of AI on current project choice from effects of AI on the average present value of remaining low skilled. Note that even without AI, self-control costs alone can prevent human-capital acquisition. Because moving from low to high human capital raises the tempting direct payoff as much as it raises productivity, remaining low-skilled can be optimal even when AI is unavailable. This is why the encouragement effect can occur: AI can promote human-capital acquisition because the relevant comparison is not between AI and a frictionless world, but between AI and a status quo in which the researcher may already be constrained by costly self-control.

While our baseline model is highly stylized, the results follow from three key assumptions. The first is that using AI is better than working without it when human capital is low and worse when human capital is high. The second is that human capital accumulation is enhanced by doing it yourself. A third assumption, that there is an increasing marginal cost of self-control, is needed for the reversal result because the independence of irrelevant alternatives is satisfied when the marginal cost of self-control is constant.

Section 5 develops a more general model that allows AI to be used even when human capital is high, allows high human capital to decay, and permits the probability of reaching high human capital to depend on both the current state and the action taken. In this richer environment, the main qualitative forces from the baseline model continue to hold: AI is still never used in the high-human-capital state, and in the low-human-capital state AI is used when the discount factor is sufficiently small because it raises current success. The section also shows that, under additional restrictions on the human-capital transition probabilities, the opposite comparative static holds for sufficiently patient researchers: when the continuation value of high human capital is large enough, AI is not used in the low state.

The more general model also yields a comparative static that has no counterpart in the baseline model. If the researcher is already choosing the long-term project in the low-human-capital state, then reducing the probability that a short-term action leads to high human capital does not overturn that choice. In other words, weaker learning from the short-term task cannot make the short-term task more attractive relative to a long-term task that was already optimal. This result isolates the role of human-capital accumulation from the direct-payoff and self-control margins emphasized in the baseline model.

Related work. Becker [1962] frames education and training as forward-looking investments that trade off current costs against future returns. Mincer [1962] extends that logic to on-the-job training and explicitly treats learning from experience as human-capital investment, and Ben-Porath [1967] embeds these choices in a dynamic life-cycle model in which individuals allocate effort and time between current production and skill accumulation. We introduce AI into the human capital literature as a technology that can raise current productivity while simultaneously introducing a self-control problem. Much of the economics of automation studies how new technologies substitute for or complement existing skills, see, for example, Autor, Levy, and Murnane [2003], Autor [2015], Acemoglu and Restrepo [2018], and Agrawal, Gans, and Goldfarb [2018]. Our focus is instead on how technology

changes the incentive to acquire those skills in the first place.

Acemoglu, Kong, and Ozdaglar [2026] studies how agentic AI can reduce human learning effort and erode collectively produced knowledge. Ide [2025] studies how automating entry-level tasks can weaken the intergenerational transmission of tacit knowledge. Both papers share our concern that AI can weaken learning, but their mechanisms work through social knowledge transmission and ours is instead an individual self-control problem. Davies [2026] studies a teacher's optimal curriculum design when students can delegate tasks to the AI instead of doing them. This leads the teacher to distort the curriculum away from the first-best, reducing overall skill development.

Garicano and Rayo [2017] and Fudenberg and Rayo [2019] study knowledge transfer in apprenticeships; Fudenberg, Georgiadis, and Rayo [2021] adds an explicit learning by doing constraint so that productive work is a vehicle for skill acquisition. These papers all feature contractual frictions between a trainer and an apprentice, while our model features an intrapersonal friction: AI changes the menu of direct payoffs, increasing temptation, so it can discourage skill-building even in the absence of an external supervisor or contract. Shen and Tamkin [2026] provides complementary evidence from a randomized experiment with software engineers: AI assistance reduced understanding and skills without delivering significant efficiency gains, with the sharpest skill losses among participants who fully delegated tasks to AI rather than remaining cognitively engaged.

Our analysis is based on past theoretical studies of temptation and self-control. Gul and Pesendorfer [2001] provides an axiomatic characterization of a model where self-control costs are proportional to the difference between the maximum currently attainable direct payoff and that of the chosen action. This paper instead uses the dual-self approach of Fudenberg and Levine [2006], which emphasizes that self-control is often a strictly convex function of foregone direct payoff. With such convex costs, enlarging the opportunity set can change the agent's preferences over previously available options. By contrast, with constant marginal self-control costs, a new tempting alternative shifts the period payoffs of the non-AI actions uniformly, and the reversal result disappears.[1]

The temptation interpretation is consistent with empirical evidence on overuse of digital technologies. Allcott, Gentzkow, and Song [2022] documents self-control problems in social-

[1] O'Donoghue and Rabin [1999] show that present bias can cause agents to procrastinate on tasks that require bearing an immediate cost for a delayed return. Our model shares that basic tension, but the introduction of AI changes the menu of available actions and thereby changes the cost of self-control, generating effects that do not arise in a present-bias framework.

media use, and the experiments of Shaw and Nave [2026] demonstrate the risk of "cognitive surrender" by users who delegate reasoning to AI even when doing so harms understanding or performance. These findings support modeling AI not only as a productivity-enhancing technology but also as a source of tempting direct payoffs.

We also add to the literature on skill erosion. Snower [1996] argues that sectors can become stuck in "low-skill, bad job" traps where low productivity and weak training incentives reinforce one another. Similarly, our reversal result shows how AI can shift behavior away from skill-building work. Our blowing-off result provides an individual-level analog to this phenomenon: once AI makes it possible to obtain output without learning, the private incentive to build expertise is weakened, potentially sustaining a low-skill equilibrium. This logic extends to the "shadow learning" observed by Beane [2019] in robotic surgery, where trainees must navigate organizational constraints to get the hands-on experience needed to learn. More generally, our model shows how a technology that raises short-run performance can also weaken the process through which expertise is formed.

Recent experimental evidence validates our concern that technologies that boost short-term productivity can weaken the skill-building process on which long-term expertise depends, creating the risks of skill erosion and "blowing off" that we identify. Dell'Acqua et al. [2023] identifies a "jagged technological frontier" where over-reliance on AI leads to systematically flawed output when tasks fall outside the tool's capabilities. This over-reliance mirrors our discouragement effect, where the direct payoff from AI outweighs the incentive to build the human judgment necessary to navigate the frontier. Similarly, Bastani et al. [2025] shows that unrestricted AI access can undermine performance on subsequent exams.[2]

2 Basic Model

Time is discrete, $t = 1, 2, \ldots$, and the researcher discounts with factor $0 \leqslant \delta < 1$. Human capital can be High (H) or Low (L), and AI may be available (A) or not (N). Although in principle there would be four states, HN, HA, LN and LA, we simplify by assuming that AI is not used when human capital is high and condense HN and HA to H. Each period the researcher chooses between two projects. The short-term project s yields a direct payoff of 1 upon success and no direct payoff next period. The long-term project t yields no direct

[2]The paper also shows that a correctly structured AI tutor can improve exam performance. Our model does not distinguish between different forms of AI, so it cannot capture this effect.

payoff in the current period, but generates a direct payoff $\Gamma > 1$ next period upon success. In state LA the researcher can choose whether to perform the task without AI (n) or else let the AI do it (a). The feasible actions are thus sn, tn, sa, and ta.

Success probabilities are as follows. In the high-human-capital state success is certain. In the low-human-capital state without AI the probability of success is p_N, and in the low-human-capital state with AI the probability of success is p_A, where $0 < p_N < p_A < 1$. If the researcher works on their own in a low-human-capital state, at the end of the period they can choose whether to remain in the low-human-capital state or move to the high-human-capital state next period. If the researcher uses AI in state LA, then next period's state is again LA. In the high-human-capital state the researcher can choose whether to retain their human capital or go back to the low human capital state.

The researcher incurs a quadratic self-control cost when resisting the temptation of a direct payoff. Let τ denote the highest achievable expected direct payoff in the period. Thus $\tau = p_N$ in state LN, $\tau = p_A$ in state LA, and $\tau = 1$ in state H. If an action yields direct payoff r_t in the current period and direct payoff r_{t+1} next period, then its *period payoff* is

$$u_t = r_t - \alpha(\tau - r_t)^2 + \delta r_{t+1},$$

where $\alpha \geqslant 0$ measures the severity of the self-control problem. Thus the self-control cost is determined by the gap between the direct payoff of the chosen action and the highest direct payoff available in that state.

Let $G := \delta\Gamma$ be the discounted direct payoff from receiving Γ next period. We assume throughout that $G > 1$; this implies that in the absence of self control costs action doing the long-term project without AI, action tn, is preferred to doing the short-term project without AI, action sn.

The resulting period payoffs are shown in Table 1.

Period payoff

State	Action			
	u^{sn}	u^{tn}	u^{sa}	u^{ta}
H	1	$G-\alpha$	–	–
LN	p_N	$p_N G-\alpha(p_N)^2$	–	–
LA	$p_N-\alpha(p_A-p_N)^2$	$p_N G-\alpha(p_A)^2$	p_A	$p_A G-\alpha(p_A)^2$

Table 1: Table entries are the period payoff by state and action. Rows are the three states H, LN, LA; columns correspond to the four actions sn, tn, sa, ta.

2.1 Benchmark: no AI available

To understand the role of AI, it is useful to first analyze the benchmark case in which AI is unavailable. In that case the only relevant states are LN and H, and the feasible actions are sn and tn.

In state H, the short-term project yields payoff 1, while the long-term project yields $G-\alpha$. Thus the long-term project is chosen if and only if $\alpha < \bar{\alpha}_H := G-1$. In state LN, the short-term project yields payoff p_N, while the long-term project yields $p_N G-\alpha p_N^2$. Thus the long-term project is chosen if and only if $\alpha < \bar{\alpha}_L := (G-1)/p_N$. Since $p_N < 1$, we have $\bar{\alpha}_L > \bar{\alpha}_H$: the long-term project is chosen for a wider range of self-control costs when human capital is low because the tempting direct payoff is smaller.

The remaining issue in state LN is whether the researcher prefers to remain in LN or move to H. The next proposition characterizes this choice and the resulting no-AI benchmark.

Proposition 1 (No-AI benchmark)**.** *If $G > 1+1/p_N$, there are cutoffs $\alpha^- < \alpha^+$ such that, starting from state LN, optimal behavior has four regions:*

(i) If $\alpha < \alpha^-$, the researcher chooses tn in LN, moves to H, and chooses tn in H.

(ii) If $\alpha^- < \alpha < \alpha^+$, the researcher chooses tn in LN and remains in LN.

(iii) If $\alpha^+ < \alpha < \bar{\alpha}_L$, the researcher chooses tn in LN, moves to H, and chooses sn in H.

(iv) If $\alpha > \bar{\alpha}_L$, the researcher chooses sn in LN, moves to H, and chooses sn in H.

If $1 < G \leqslant 1 + 1/p_N$, the researcher always moves from LN to H, and optimal behavior has three regions:

(i) If $\alpha < \bar{\alpha}_H$, the researcher chooses tn in LN and tn in H.

(ii) If $\bar{\alpha}_H < \alpha < \bar{\alpha}_L$, the researcher chooses tn in LN and sn in H.

(iii) If $\alpha > \bar{\alpha}_L$, the researcher chooses sn in LN and sn in H.

At each cutoff the researcher is indifferent between the corresponding choices.

When $G > 1 + 1/p_N$, the intuition for the middle two regions is simple. For intermediate values of α, the cost of resisting temptation makes H worse than LN, so the researcher prefers to remain in the low-human-capital state. This is the only case in which there researcher does not acquire human capital. For larger values of α, the researcher instead chooses the short-term project in H, eliminating that self-control cost and making H attractive again. Thus the effect of self-control costs on human-capital acquisition is nonmonotone.

When $1 < G \leqslant 1+1/p_N$, the researcher never strictly prefers to remain in LN. The long-term payoff is not large enough for the lower temptation in LN to make remaining there more attractive than moving to H. Self-control costs then affect which project is chosen in each state, but do not prevent human-capital acquisition. At $G = 1+1/p_N$ there is a single value of α at which the researcher is indifferent between remaining in LN and moving to H. Thus even without AI, self-control costs alone can prevent human-capital acquisition, and their effect on human-capital acquisition is nonmonotone. In particular when G is large, in the intermediate range of self-control cost described in (ii) results in non-acquisition of human capital. This observation is central for understanding the encouragement effect once AI becomes available.

3 Overview of the Results

Before characterizing the optimal decision, we will explain three key properties of the solution. First, conditional on the state, the value of G, and whether or not AI is used, the choice of which project to choose is purely myopic and does not depend on the discount

factor δ. Second, with self-control costs, the high-human-capital state can be worse than the low-human-capital state, even when AI is unavailable. Finally, conditional on G, the discount factor matters only for the choice of using AI in the low human capital state.

The project-choice comparison is myopic because, once the state and the technology are fixed, the future consequences of the short-term and long-term projects are the same. Whether the short-term or long-term project is preferred depends solely on whether α is above or below a state-specific cutoff. These cutoffs are shown in Table 2.

Comparison	Condition	Cutoff
$u_{LN}^{sn} > u_{LN}^{tn}$	$\alpha > \bar{\alpha}_L$	$\bar{\alpha}_L = \dfrac{G-1}{p_N}$
$u_{LA}^{sa} > u_{LA}^{ta}$	$\alpha > \bar{\alpha}_A$	$\bar{\alpha}_A = \dfrac{G-1}{p_A}$
$u_{LA}^{sn} > u_{LA}^{tn}$	$\alpha > \bar{\alpha}_N$	$\bar{\alpha}_N = \dfrac{G-1}{2p_A - p_N}$
$u_H^{sn} > u_H^{tn}$	$\alpha > \bar{\alpha}_H$	$\bar{\alpha}_H = G-1$

Table 2: Static project comparisons and cutoff values.

To see why the high human capital state can be worse even when AI is unavailable, consider the period payoffs in the high- and low-human-capital states. The high state raises productivity, but it also raises temptation: in state H the tempting direct payoff is 1, while in a low-human-capital state the tempting direct payoff is p_N without AI and p_A with AI. Let $\hat{u}_H := \max\{1, G-\alpha\}$, $\hat{u}_L := \max\{p_N, p_N G - \alpha(p_N)^2\}$, $\hat{u}_A := \max\{u_{LA}^{sa}, u_{LA}^{ta}\}$, and $\hat{u}_N := \max\{u_{LA}^{sn}, u_{LA}^{tn}\}$. Since temptation lowers period payoff, $\hat{u}_H$ can be lower than the corresponding low-state period payoff, and in this case the researcher may prefer low human capital despite its lower productivity.

Finally, conditional on G, the discount factor matters only for the choice of using AI in the low-human-capital state because it affects only the average present value of the continuation. In particular, δ changes the average present value of the no-AI path in the low-human-capital state, since that path may lead to high human capital. Thus a higher δ makes the researcher more willing to give up the AI period payoff in order to move toward H, while a lower δ makes AI more attractive.

Having laid out the basic considerations, we now give our main results. We first give the basic properties of the equilibria, then lay out the comparative statics of when each of

the four effects can occur.

Proposition 2 (Basic properties)**.**

(i) If $\alpha < \min\{\bar{\alpha}_H, \bar{\alpha}_N\}$, only the long-term project is chosen in every state.

(ii) If $\alpha > \bar{\alpha}_L$, only the short-term project is chosen in every state.

(iii) In the low state with AI available, for any fixed $G > 1$ there is a discount factor $\hat{\delta} \in (0, 1]$ such that for $\delta < \hat{\delta}$ AI is used, and for $\delta > \hat{\delta}$ AI is not used. If G is sufficiently large, then $\hat{\delta} < 1$.

The logic is straightforward and follows the discussion above. Parts (i) and (ii) come from the static project rankings. Part (iii) concerns the low-human-capital state with AI: using AI raises the period payoff, while not using AI can lead to H, so as δ rises the average present value of working without AI rises, and AI use falls whenever the transition to H is valuable enough.

Recall the four effects identified in the Introduction: **blowing off**, where the researcher uses AI to complete the long-term project while in the low-human-capital state, thereby forgoing the accumulation of human capital; **discouragement**, where the researcher uses AI to complete the short-term project in the low-human-capital state, even though they would have chosen the long-term project had AI been unavailable; **reversal**, where the arrival of AI causes the researcher to switch from the long-term project to the short-term project in the low-human-capital state, even though AI is not actually used; and **encouragement**, where the availability of AI induces the researcher to acquire human capital, even though they would have remained in the low-human-capital state without it. The next table gives a rough sense of when each effect occurs.

Effect	α-range	Discount factor	AI used?
Blowing off	$\alpha < \bar{\alpha}_A$	low δ	Yes
Discouragement	$\bar{\alpha}_A < \alpha < \bar{\alpha}_L$	low δ	Yes
Reversal	intermediate range	high δ	No
Encouragement	intermediate range	high δ	No

Table 3: Summary of the four effects.

The four effects come from two margins. One is project choice within the current state: AI can make short-term work more attractive, or make using AI to do long-term work more attractive than doing long-term work without AI. The other is the average present value of remaining in the low-human-capital state relative to moving to H. Blowing off and discouragement arise when AI makes behavior more myopic, reversal arises when AI changes the ranking of the two non-AI actions, and encouragement arises when AI changes the average present value of remaining in the low state relative to acquiring high human capital.

The first two rows are the cases in which AI is used in the low-human-capital state. For blowing off, $\alpha < \bar{\alpha}_A$ implies that ta is the best AI action in state LA; for sufficiently low δ, the period payoff from using AI—and thus its conditional average present value—dominates the average present value from working without AI and moving to H, so ta is optimal. Thus the researcher still chooses the long-term project, but delegates it to AI and remains in the low-human-capital state.

For discouragement, $\bar{\alpha}_A < \alpha < \bar{\alpha}_L$ means that tn is chosen in state LN, while sa beats ta among the AI actions in state LA. Let $\hat{u}_N := \max\{u_{LA}^{sn}, u_{LA}^{tn}\}$. If, in addition, $u_{LA}^{sa} > \hat{u}_N$, then for sufficiently low δ the short-term AI action also beats the best no-AI action followed by transition to H. Hence sa is optimal in state LA, even though tn would have been chosen if AI were unavailable. The formal proof of the two AI-use effects is given in Proposition 8 in Appendix B.

In the last two rows, the availability of AI changes behavior even though AI is not used in the relevant low-human-capital state. In reversal, the self-control cost created by the AI option makes the short-term non-AI action more attractive than the long-term non-AI action in state LA, even though the long-term project would have been chosen in state LN. In encouragement, AI changes the average present value of remaining low-skilled in a way that can make transition to H more attractive. These effects require intermediate self-control costs and are favored by high δ, because a patient researcher puts more weight on the average present value associated with working without AI. The next two propositions give the precise conditions for reversal and encouragement.

Proposition 3 (Reversal)**.**

(i) Reversal occurs only if

$$\bar{\alpha}_N < \alpha < \bar{\alpha}_L.$$

(ii) If $\bar{\alpha}_N < \alpha < \bar{\alpha}_L$ *and*

$$(1-\delta)\big(p_N - \alpha(p_A - p_N)^2\big) + \delta\hat{u}_H > \max\{p_A,\, p_A G - \alpha(p_A)^2\},$$

then reversal occurs.

The key to Proposition 3 is that the availability of AI raises the temptation benchmark enough to reverse the ranking of the two non-AI actions in state LA. The researcher then chooses the short-term project without AI, even though the long-term project would have been chosen if AI were unavailable.

Proposition 4 (Encouragement)**.** *A sufficient condition for encouragement for all* δ *sufficiently close to* 1 *is*

$$Gp_N - \alpha(p_N)^2 > 1 > Gp_A - \alpha(p_A)^2.$$

Such parameter values exist if and only if

$$G > \frac{1}{p_N} + \frac{1}{p_A}.$$

Proposition 4 gives the opposite kind of effect. Its condition says that, without AI, remaining in the low-human-capital state is attractive, while with AI the average present value of the best AI-assisted low-state action is not high enough to dominate moving to H. For sufficiently high discount factors, the researcher therefore chooses a no-AI path to high human capital. In this case, AI encourages human-capital acquisition by lowering the relative average present value of staying low-skilled.

4 Discussion of the Four Effects

This section discusses the economic mechanisms behind the four effects.

4.1 Blowing off

Blowing off occurs when the researcher uses AI for the difficult long-term task in the low-human-capital state, thereby forgoing human-capital acquisition. Thus the researcher still wants the direct payoff from the long-term project, but prefers to obtain it through AI rather than through working without AI.

Economically, the key force is that AI raises current productivity on the long-term task while eliminating the learning by doing that would otherwise come from doing that task unaided. This is the most direct way in which AI can weaken skill formation: the researcher does not switch away from the long-term project, but switches away from the skill-building version of it.

This effect arises when using AI on long-term work is attractive enough both relative to using AI for short-term projects and relative to the best no-AI path that leads to high human capital. In terms of the model, this requires that ta beat both sa and the best non-AI action followed by H. So the economic tradeoff is between higher current performance with AI and the future value of learning. The consequence is that the researcher remains in the low-human-capital state even while continuing to pursue the long-term project. Because any optimal non-AI action in state LA is followed by transition to H, whereas ta keeps the researcher in LA, blowing off replaces a path with learning by doing by one without it. The distortion is not that the researcher stops aiming high, but that they pursue the same long-term objective in a way that prevents skill accumulation.

4.2 Discouragement

Discouragement occurs when without the availability of AI the long-term task is chosen in the low-human-capital state, but when AI is made available there is a switch to the short-term task with AI. This differs from blowing off because the availability of AI now changes the project itself, not just the technology used to complete it. Under discouragement, the researcher gives up on long-term work and moves to the short-term project. The economic mechanism is that AI raises the direct payoffs of options that pay immediately, and this makes the short-term project with AI more attractive than the long-term project without AI.

For discouragement to occur, the long-term project must be preferred in state LN, which requires $\alpha < \bar{\alpha}_L$, while in state LA the short-term AI action must beat the long-term AI action, which requires $\alpha > \bar{\alpha}_A$. Thus discouragement occurs only for the intermediate range $\bar{\alpha}_A < \alpha < \bar{\alpha}_L$. On that interval, the self-control problem is strong enough that the short-term AI action becomes more attractive than the long-term AI action, but not so strong that the short-term project would already have been chosen even without AI. The remaining condition is that sa also beat using the best non-AI action and transitioning to H.

Economically, discouragement is a distortion on the project-choice margin that redirects effort toward the short-term task. This lowers eventual human capital only if working without AI in LN would otherwise have led to transition to H. If the no-AI choice in LN would already have been followed by remaining in LN, discouragement changes project choice but not the eventual level of human capital.

4.3 Reversal

Reversal occurs when without the availability of AI the long-term task is chosen in the low-human-capital state, but when AI is made available there is a switch to the short-term task without AI, so that AI changes behavior even though it is not used. The logic is similar to the violations of WARP discussed in Fudenberg and Levine [2006]: increasing the highest direct payoff increases self-control costs. Here this changes the ranking of the two non-AI actions.

For reversal to occur, the long-term project must be chosen in state LN, so $\alpha < \bar{\alpha}_L$, but conditional on not using AI in state LA, the short-term project must beat the long-term one, so $\alpha > \bar{\alpha}_N$. Thus reversal requires the intermediate range $\bar{\alpha}_N < \alpha < \bar{\alpha}_L$. The economic content of this interval is that self-control costs are not large enough to make the researcher avoid the long-term project in the no-AI environment, but become large enough to overturn that ranking once AI is added to the menu.

It is important to understand that reversal depends heavily on the assumption that self-control cost is quadratic. If self-control cost were to be linear, that is $\alpha(\tau - r_t)$ instead of $\alpha(\tau - r_t)^2$, then reversal is impossible. This is because introducing stronger temptation raises the self-control costs of all other alternatives by exactly the same amount for all other alternatives, and so does not change their ranking. Hence, with linear self-control, the independence of irrelevant alternatives is satisfied, and there can be no reversal.

4.4 Encouragement

Encouragement occurs when no human capital is acquired in the low state without AI, but human capital is acquired in the low state when AI is available. This is the most surprising effect, because it goes against the simple intuition that AI must always weaken incentives to learn. The key point is that AI changes not only the attractiveness of current actions, but also the average present value of remaining in the low-human-capital state. Without AI,

the researcher may prefer to remain in LN because moving to H raises future temptation too much. With AI, that low-human-capital state becomes less attractive as a place to remain, and this can make transition to H optimal.

Formally, encouragement requires that without AI the agent chooses tn and remains in LN. By Lemma 2, this means

$$\frac{G}{1+p_N} < \alpha < \frac{p_N G - 1}{(p_N)^2}.$$

So, absent AI, the long-term project is attractive enough that the researcher does not switch to the short-term project, but the average present value of remaining in the low state still exceeds the average present value of moving to H. Encouragement requires that taking the best non-AI action and moving to H is better than having the AI do either action. Thus AI must make the low state unattractive enough that the researcher now prefers skill acquisition. The encouragement effect is welfare-reducing for the researcher: AI induces human-capital acquisition by lowering the average present value of remaining in the low-human-capital state, not by raising the average present value of acquiring human capital.

A useful sufficient condition for encouragement for δ close to 1 is

$$Gp_N - \alpha(p_N)^2 > 1 > Gp_A - \alpha(p_A)^2.$$

The left inequality says that without AI the low state is attractive enough that the researcher prefers to remain there. The right inequality says that once AI is available, the AI-assisted low-state option is no longer attractive enough to dominate transition to H. This sufficient condition can hold only if

$$G > \frac{1}{p_N} + \frac{1}{p_A},$$

so encouragement requires the delayed direct payoff from long-term work to be large.

5 The General Case

This section shows that the main forces identified in the basic model do not depend on its stark transition structure. In the basic model, working without AI allows the researcher to

choose whether to acquire high human capital, whereas using AI leaves the researcher in the low-human-capital state. Here, each action instead induces a potentially interior probability of high human capital next period. We also allow AI to be used in the high-human-capital state and replace the quadratic self-control cost with a general weakly convex cost.

We let $I \in \{L, H\}$ index the agent's human capital and $J \in \{N, A\}$ track whether AI is available, so that the state is a pair (I, J). The feasible actions are denoted (k, ℓ) where $k \in \{s, t\}$ and $\ell \in \{n, a\}$. The short-term (s) project continues to provide only an immediate return of 1 with some probability and the long-term (t) project continues to provide a return in the next period with a return of Γ with some probability and $G = \delta\Gamma > 1$. The success rates are given by $p_{IJ}^{k\ell}$ and the probabilities of high human capital next period are $h_{IJ}^{k\ell}$. We make the following assumptions, which preserve the qualitative structure of the basic model:

1. If AI is not used then availability makes no difference: $p_{IN}^{kn} = p_{IA}^{kn}$,$h_{IN}^{kn} = h_{IA}^{kn}$.
2. The long-term task is better than the short-term task: $p_{IJ}^{tk} G > p_{IJ}^{sk}$, $h_{IJ}^{tk} \geqslant h_{IJ}^{sk}$.
3. The high state is better than the low state: $p_{HJ}^{k\ell} > p_{LJ}^{k\ell}$, $h_{HJ}^{k\ell} \geqslant h_{LJ}^{k\ell}$.
4. AI is better in the low state for success: $p_{LA}^{ka} > p_{LA}^{kn}$, but worse for human capital $h_{LA}^{ka} < h_{LA}^{kn}$.
5. AI worse in the high state: $p_{HA}^{kn} > p_{HA}^{ka}$, $h_{HA}^{ka} < h_{HA}^{kn}$.
6. All of the $p_{IJ}^{k\ell}$ are strictly positive.

Let T_{IJ} denote the highest achievable current period payoff given the current state. Then with low human capital we have $T_{LN} = p_{LN}^{sn}$ and $T_{LA} = p_{LA}^{sa}$, while in the high human capital state $T_{HN} = T_{HA} = p_{HN}^{sn}$.

For each state IJ and action $k\ell$, let $(r_{IJ}^{k\ell}, d_{IJ}^{k\ell})$ denote current expected payoff and discounted delayed payoff: $(r_{IJ}^{s\ell}, d_{IJ}^{s\ell}) = (p_{IJ}^{s\ell}, 0)$ and $(r_{IJ}^{t\ell}, d_{IJ}^{t\ell}) = (0, p_{IJ}^{t\ell} G)$, The period utility from the action (k, ℓ) in state IJ is $u_{IJ}(k\ell) = r_{IJ}^{k\ell} - \alpha c(T_{IJ} - r_{IJ}^{k\ell}) + d_{IJ}^{k\ell}$, where $c(\cdot)$ is strictly increasing, weakly convex, smooth, and satisfies $c(0) = 0$ and $c(1) = 1$. The decision maker chooses projects and AI to maximize the expected present value of utility.

Note that if $c(\cdot)$ is linear, the no-reversal result of section 4.3 remains unchanged as the argument applies generally. Note also that the Bellman equation is

$$V_{IJ} = \max_{(k,\ell)} \left\{ (1-\delta)u_{IJ}^{k\ell} + \delta \left[(1-h_{IJ}^{k\ell}) V_{LJ} + h_{IJ}^{k\ell} \max\{V_{LJ}, V_{HJ}\} \right] \right\}, \qquad I \in \{L, H\}, \quad J \in \{N, A\}$$

5.1 Results

We first show that allowing AI use in the high-human-capital state does not change behavior there. We then establish that the main comparative statics of the basic model continue to hold with probabilistic, action-dependent human-capital transitions. Finally, the richer transition structure yields an additional comparative-static result concerning the human-capital return to the short-term project.

Proposition 5. *AI is never used in the high human capital state.*

This conclusion follows from the dominance of working without AI in the high-human-capital state. For either project, working without AI yields both a higher probability of success and a higher probability of retaining high human capital. Moreover, AI does not alter the temptation benchmark in this state, because the most tempting action is already the short-term project without AI. Thus AI offers neither a current-payoff advantage nor a continuation-value advantage when human capital is high.

The next proposition extends the basic comparative statics in the severity of the self-control problem and in the discount factor.

Proposition 6. *[Proposition 1 for the general model.]*

(i) There is an $\underline{\alpha} > 0$ such that if $\alpha < \underline{\alpha}$, then only the long-term project is chosen in every state.

(ii) There is an $\overline{\alpha}$ such that if $\alpha > \overline{\alpha}$, then only the short-term project is chosen in every state.

(iii) In the low state with AI available, for any fixed $G > 1$ there is a discount factor $\underline{\delta} \in (0, 1]$ such that, for $\delta < \underline{\delta}$, AI is used.

(iv) If in the low state AI does not generate human capital ($h_{LA}^{ka} = 0$), human capital does not depreciate in the high state if the long-term project without AI is chosen ($h_{HA}^{tn} = 1$), and $p_{HA}^{tn} > p_{LA}^{ta}$, then if Γ is sufficiently large, there is an $\overline{\delta} < 1$ such that, for $\delta > \overline{\delta}$, AI is not used.

Parts (i) and (ii) show that the basic project-choice comparative statics survive the introduction of action-dependent human-capital transitions. When α is small, the self-control cost of forgoing the immediate payoff is small. The long-term project then has both a period-payoff advantage, through its larger delayed return, and a weak continuation-value advantage, through its higher probability of producing high human capital. Consequently, the long-term project is chosen in every state. When α is sufficiently large, by contrast, the cost of resisting the immediate payoff dominates both of these advantages, and the short-term project is chosen everywhere.

Part (iii) concerns the case where the discount factor is small, so the researcher puts most weight on what succeeds now. In state LA, AI raises current success, so the AI-assisted short-term action is optimal for sufficiently low δ. The point is simply that AI improves present performance enough to dominate the continuation value when the future is discounted heavily.

The extra assumptions in part (iv) are needed. Example 1 shows that if AI does not generate human capital in the low state but human capital depreciates in the high state, then the low-state AI path can still dominate for patient researchers, because the researcher revisits the low state often enough that the current AI advantage keeps accumulating. Example 2 shows that even when the High state is absorbing, AI can still remain optimal if its current success advantage in the low state is large enough.

Example 1. *Take $\alpha = 0$, so that the short-term project will never be chosen, and suppose that AI does not produce human capital at all, so that $h_{LA}^{\ell a} = 0$. If AI is available, we know that it is not used in the high human capital state, so the choice at HA is tn. If AI is used in the low human capital state, the average present value is $V_{LA}^{ta} = p_{LA}^{ta} G$.*

It remains to find the average present values when AI is not used in the low human capital state. For illustrative purposes, take $h_{LA}^{tn} = h_{HA}^{tn} = 1/2$. As $\delta \to 1$ we have $V_{LA}^{tn} \to \left(p_{LA}^{tn} + p_{HA}^{tn}\right) G/2$, and this is an upper bound on V_{LA}^{tn} for all δ. Hence, AI will be used in LA if $p_{LA}^{ta} > \left(p_{LA}^{tn} + p_{HA}^{tn}\right)/2$, which is not ruled out. The point is that if human capital depreciates in the high human capital state, then the low human capital state must occur frequently. Hence, the disadvantage of not using AI in the low human capital state may be so great that it drags down the advantage of the high human capital state below what can be obtained with AI in the low human capital state.

Example 2. *Set $\alpha = 0$ and let the high state be absorbing: $h_{HJ}^{k\ell} = 1$ for all J, k, ℓ. In*

the low state let $h^{tn}_{LA} = 1$ *and* $h^{ta}_{LA} = \frac{99}{100}$ *and choose*

$$p^{tn}_{LA} = \frac{1}{10}, \qquad p^{ta}_{LA} = \frac{9}{10}, \qquad p^{tn}_{HN} = 1 = p^{tn}_{HA}.$$

Since the high state is absorbing, $V_H = G$*. In state LA,* $V^{tn}_L = (1-\delta)\frac{G}{10} + \delta G$*. For ta,*

$$V^{ta}_L = (1-\delta)\frac{9G}{10} + \delta\left(\frac{1}{100}V^{ta}_L + \frac{99}{100}G\right),$$

so

$$V^{ta}_L = \frac{(1-\delta)\frac{9G}{10} + \delta\frac{99G}{100}}{1-\delta/100}.$$

Hence

$$V^{ta}_L - V^{tn}_L = \frac{G(1-\delta)(800-9\delta)}{10(100-\delta)} > 0 \qquad \textit{for all } \delta < 1.$$

Thus AI is used in LA even when the high state is absorbing and AI in the low state allows nearly as much human-capital accumulation as no-AI.

In addition to letting us explore the robustness of our earlier results, the general model also permits a comparative static that has no meaningful counterpart in the basic model.

Proposition 7. *Suppose that in the low human capital state, the long-term project is chosen. If, for any short-term action, the probability of reaching the high-human-capital state is reduced, the long-term project is still chosen.*

This result isolates the role of learning by doing from the direct-payoff and temptation effects. Reducing the human-capital transition probability of a short-term action leaves its current payoff and its self-control cost unchanged. It only lowers the continuation value generated by that action. The payoff and transition probability associated with the long-term action are unchanged. Therefore, a reduction in learning from short-term work cannot make the short-term project more attractive relative to a long-term project that was already optimal. The proposition is one-sided: Reducing the human-capital return to the long-term project may overturn the researcher's choice, because it removes one of the advantages that supports long-term work. By contrast, reducing learning from the short-term project reinforces the original ranking. This distinction highlights that the model's conclusions are not driven solely by the direct payoffs of the two projects. Differences in

the amount of human capital generated by the projects provide an independent reason to choose long-term work, especially for patient researchers.

Taken together, the propositions show that the mechanisms of the basic model are robust to probabilistic and action-dependent human-capital accumulation. AI is not used once human capital is high, greater self-control costs shift choices toward short-term projects, and greater patience can shift choices away from AI by increasing the value of learning.

6 Conclusion

AI does more than raise productivity. By increasing the best available direct payoff, it also changes the temptation benchmark faced by a researcher with costly self-control. This matters because the choices that determine current output also determine whether human capital is acquired: AI affects not only what can be produced today, but also whether the researcher has reason to acquire the skills that make future work productive. Without self-control costs, the availability of AI cannot lower the researcher's average present value unless AI is actually used. With costly self-control, by contrast, even unused AI can lower average present value by raising the cost of resisting immediate rewards, and it can distort project choice in ways that slow or prevent human-capital accumulation.

The model also distinguishes four ways that AI can affect human-capital formation. Blowing off occurs when the researcher uses AI on the long-term project and thereby avoids the learning by doing that would have raised future productivity. Discouragement occurs when AI induces a switch from non-assisted long-term work to AI-assisted short-term projects. Reversal occurs when AI changes the ranking of unassisted actions and induces a switch from the long-term project to the short-term project even though AI is never used. Encouragement means that the effect of AI goes in the opposite direction: it induces the acquisition of human capital that would not have occurred in its absence. This effect is not welfare-improving for the researcher; it occurs because AI lowers the relative average present value of remaining in the low-human-capital state. These effects are most relevant when AI substitutes for early-stage learning; if AI is strongly complementary with expertise, then AI can instead strengthen the incentive to acquire human capital. Our analysis shows that unrestricted early access to AI can weaken the incentives to acquire human capital. This suggests a role for training rules, staged access, and institutional designs that preserve learning-by-doing while still allowing workers to benefit from AI. The point is not that AI

use is undesirable in general. Once the relevant human capital has been acquired, AI can be valuable because as Agrawal, Gans, and Goldfarb [2025] emphasizes, it can complement judgment even as it substitutes for implementation. The key issue is therefore not AI's effect on current output alone, but its effect on the incentives to acquire the skills on which future output depends.

A Auxiliary Results

Lemma 1 (Static project rankings)**.** *The cutoffs satisfy*

$$\bar{\alpha}_L > \bar{\alpha}_A > \bar{\alpha}_N, \qquad \bar{\alpha}_N \gtrless \bar{\alpha}_H \iff 2p_A - p_N \lesseqgtr 1.$$

Proof. Since $p_N < p_A < 2p_A - p_N$, we have $\bar{\alpha}_L > \bar{\alpha}_A > \bar{\alpha}_N$. Also, $\bar{\alpha}_N \gtrless \bar{\alpha}_H$ is equivalent to $(G-1)/(2p_A - p_N) \gtrless G - 1$, which is equivalent to $2p_A - p_N \lesseqgtr 1$. □

Let W_{LN} and W_{LA} denote the average present values starting from the low-human-capital state when AI is unavailable and available, respectively. The average present value of high human capital can depend on whether AI would be available after a later return to low human capital, so we write $W_H N$ for the average present value of H in the no-AI environment and $W_H A$ for the average present value of H in the AI-available environment.

Lemma 2 (No-AI average present values and transition from LN)**.** *Let*

$$\hat{u}_H := \max\{1, G - \alpha\}, \qquad \hat{u}_L := \max\{p_N, p_N G - \alpha(p_N)^2\}.$$

Then

$$W_{HN} = (1-\delta)\hat{u}_H + \delta \max\{W_{HN}, W_{LN}\},$$

and

$$W_{LN} = (1-\delta)\hat{u}_L + \delta \max\{W_{HN}, W_{LN}\}.$$

Hence $W_{HN} \gtrless W_{LN}$ if and only if $\hat{u}_H \gtrless \hat{u}_L$. After a non-AI action in LN, the researcher remains in LN if and only if

$$\hat{u}_L > \hat{u}_H,$$

which is equivalent to

$$\frac{G}{1+p_N} < \alpha < \frac{p_N G - 1}{(p_N)^2}.$$

Proof. In state H, the researcher first obtains the best period payoff $\hat{u}_H = \max\{1, G-\alpha\}$ and then chooses whether to retain high human capital or return to the low-human-capital state. Hence

$$W_{HN} = (1-\delta)\hat{u}_H + \delta \max\{W_{HN}, W_{LN}\}.$$

Similarly, in state LN the researcher obtains the best no-AI period payoff $\hat{u}_L = \max\{p_N, p_N G - \alpha(p_N)^2\}$ and then chooses whether to remain in LN or move to H, so

$$W_{LN} = (1-\delta)\hat{u}_L + \delta \max\{W_{HN}, W_{LN}\}.$$

Subtracting the two equations gives

$$W_{HN} - W_{LN} = (1-\delta)(\hat{u}_H - \hat{u}_L).$$

Thus $W_{HN} \geqslant W_{LN}$ if and only if $\hat{u}_H \geqslant \hat{u}_L$, and remaining in LN is optimal if and only if $\hat{u}_L > \hat{u}_H$.

Because $p_N < 1$, this inequality is equivalent to

$$\frac{G}{1+p_N} < \alpha < \frac{p_N G - 1}{(p_N)^2}.$$

□

Lemma 3 (AI-available low state)**.** *In state LA, it is never optimal to choose SN or TN and remain in LA next period. Therefore*

$$\hat{u}_A := \max\{u_{LA}^{sa}, u_{LA}^{ta}\} > \hat{u}_N := \max\{u_{LA}^{sn}, u_{LA}^{tn}\},$$

and thus

$$W_{HA} = (1-\delta)\hat{u}_H + \delta \max\{W_{HA}, W_{LA}\}$$

and

$$W_{LA} = \max\left\{\hat{u}_A,\ (1-\delta)\hat{u}_N + \delta W_H A\right\}.$$

Proof. In state LA, the feasible (action, continuation state) pairs are (sn, H), (sn, LA),

(tn, H), (tn, LA), (sa, LA), and (ta, LA). Since $p_A > p_N$, if the continuation state is fixed at LA, the AI version of each project strictly dominates the corresponding non-AI version. Hence no optimal policy chooses sn or tn and remains in LA, so $\hat{u}_A > \hat{u}_N$. Since any optimal no-AI action in state LA must be followed by transition to H, the average present value is $(1-\delta)\hat{u}_N + \delta W_H A$. If instead an AI action is optimal, then $W_{LA} = (1-\delta)\hat{u}_A + \delta W_{LA}$, so $W_{LA} = \hat{u}_A$. Combining the two cases gives $W_{LA} = \max\{\hat{u}_A, (1-\delta)\hat{u}_N + \delta W_{HA}$. □

Lemma 4 (AI use in the low-human-capital state)**.** *Suppose AI is available in state LA. An AI action is strictly optimal in state LA if and only if $\hat{u}_A > (1-\delta)\hat{u}_N + \delta\hat{u}_H$. Moreover, if $\hat{u}_A \geqslant \hat{u}_H$, an AI action is optimal for every δ, and if $\hat{u}_N < \hat{u}_A < \hat{u}_H$, there is a unique cutoff*

$$\hat{\delta} = \frac{\hat{u}_A - \hat{u}_N}{\hat{u}_H - \hat{u}_N} \in (0, 1)$$

such that an AI action is strictly optimal for $\delta < \hat{\delta}$, a no-AI action followed by transition to H is strictly optimal for $\delta > \hat{\delta}$, and the researcher is indifferent at $\delta = \hat{\delta}$.

Proof. If $\hat{u}_A > \hat{u}_H$ then AI is used with low human capital regardless of the discount factor. Otherwise, define $w_A = \hat{u}_A$ to be the optimal average present value conditional on using AI with low human capital. Similarly, let

$$w_N = \hat{u}_N + \delta(\hat{u}_H - \hat{u}_N)$$

be the average present value conditional on doing it yourself and acquiring human capital. Consequently, AI is strictly worse if

$$\delta > \frac{\hat{u}_A - \hat{u}_N}{\hat{u}_H - \hat{u}_N} = \hat{\delta}.$$

For G large, $\hat{u}_A < \hat{u}_H$ and the long-term project is always optimal. Hence, as $G \to \infty$,

$$\hat{\delta} \to \frac{p_A - p_N}{1 - p_N} < 1.$$

□

Lemma 5 (Sufficient condition for encouragement near $\delta = 1$)**.** *Suppose*

$$Gp_N - \alpha(p_N)^2 > 1 > Gp_A - \alpha(p_A)^2.$$

Then, for all δ sufficiently close to 1, encouragement occurs.

Proof. Let $f(x) := Gx - \alpha x^2$. The hypothesis says $f(p_N) > 1 > f(p_A)$. Since $p_N < p_A < 1$ and f is concave, $f(p_A) < f(p_N)$ implies that f is decreasing on $[p_A, 1]$, so $f(1) < f(p_A) < 1$.

Thus in state H the best period payoff is $\hat{u}_H = 1$, while in state LN the best period payoff is $\hat{u}_L = f(p_N) > 1$. By Lemma 2, without AI the researcher chooses tn and remains in LN.

Now consider state LA. The best period payoff using AI is $\hat{u}_A = \max\{p_A, f(p_A)\} < 1$. When AI is available, the researcher can always choose the short-term no-AI action in state H, which yields period payoff $\hat{u}_H = 1$ with no self-control cost (since sn is the temptation benchmark in H) and can be repeated forever by remaining in H. Therefore $W_{HA} \geqslant 1$. The average present value of the best no-AI action in state LA followed by transition to H is $(1-\delta)\hat{u}_N + \delta W_{HA}$, which converges to $W_{HA} \geqslant 1$ as $\delta \uparrow 1$. Since $\hat{u}_A < 1$, for all sufficiently high δ the best no-AI choice followed by transition to H is better than using AI. Thus, with AI available, the researcher acquires human capital, whereas without AI the researcher remains in LN. □

Proposition 8 (AI-use effects)**.** *Fix G.*

(i) If $\alpha < \bar{\alpha}_A$, then ta is the best AI action in state LA, and blowing off occurs for all sufficiently low δ.

(ii) If $\bar{\alpha}_A < \alpha < \bar{\alpha}_L$ and $u^{sa}_{LA} > \hat{u}_N$, then tn is chosen in state LN, while sa is optimal in state LA for all sufficiently low δ.

Proof. Part (i): By Lemma 3, blowing off occurs if ta is optimal in state LA: in that case the researcher chooses the long-term project using AI and remains in state LA. The condition $ta > sa$ is $\alpha < \bar{\alpha}_A$. We now show that this same inequality also implies that ta beats both no-AI actions in period payoff. The comparison with tn is immediate, since $u^{ta}_{LA} - u^{tn}_{LA} = (p_A - p_N)G > 0$. For sn, ta beats sn by direct algebra under $\alpha < \bar{\alpha}_A$. Hence if $\alpha < \bar{\alpha}_A$, then ta is the action with the highest period payoff in state LA, so $\hat{u}_A = u^{ta}_{LA}$ and $\hat{u}_A > \hat{u}_N$.

By Lemma 4, if $\hat{u}_A \geqslant \hat{u}_H$, an AI action is optimal for every $\delta < 1$. If instead $\hat{u}_N < \hat{u}_A < \hat{u}_H$, there is a cutoff

$$\hat{\delta} = \frac{\hat{u}_A - \hat{u}_N}{\hat{u}_H - \hat{u}_N} \in (0, 1)$$

such that an AI action is strictly optimal for $\delta < \hat{\delta}$. Since under $\alpha < \bar{\alpha}_A$ the optimal AI action is ta, blowing off occurs for all sufficiently low δ.

Part (ii): Discouragement means that tn is chosen in state LN but sa is optimal in state LA. When $\alpha < \bar{\alpha}_L$ tn is chosen in state LN, while $\alpha > \bar{\alpha}_A$ implies that sa beats ta among the AI actions in state LA. Thus on the interval $\bar{\alpha}_A < \alpha < \bar{\alpha}_L$, the best AI action in state LA is sa, so $\hat{u}_A = u^{sa}_{LA}$.

The condition $u^{sa}_{LA} > \hat{u}_N$ implies $\hat{u}_A > \hat{u}_N$. By Lemma 4, sa is optimal in state LA for all sufficiently low δ. Hence AI induces a switch from tn in state LN to sa in state LA, which is discouragement. □

B Proofs of the propositions

Proof of Proposition 1. In the no-AI benchmark, the only states are LN and H, the feasible actions are sn and tn, and after a no-AI action in LN the researcher can choose whether to remain in LN or move to H. These are exactly the choices analyzed in Lemma 2. Hence, by that lemma, the researcher prefers to remain in LN if and only if $\alpha^- < \alpha < \alpha^+$, where $\alpha^- := G/(1+p_N)$ and $\alpha^+ := (p_N G - 1)/p_N^2$.

The interval (α^-, α^+) is nonempty if and only if $G > 1 + 1/p_N$. In that case $\alpha^- < \bar{\alpha}_H < \alpha^+ < \bar{\alpha}_L$, and combining these inequalities with the project-choice cutoffs gives the four regions in the first part of the proposition. If instead $1 < G \leqslant 1 + 1/p_N$, the interval (α^-, α^+) is empty, so the researcher never strictly prefers to remain in LN. Since $\bar{\alpha}_H < \bar{\alpha}_L$, the project-choice cutoffs then give the three regions in the second part.

□

Proof of Proposition 2. Parts (i) and (ii) follow immediately from Lemma 1 and the ordering $\bar{\alpha}_L > \bar{\alpha}_A > \bar{\alpha}_N$, together with $\bar{\alpha}_L > \bar{\alpha}_H$. For part (iii), fix $G > 1$, holding the other primitives fixed. Recall

$$\hat{u}_A := \max\{u^{sa}_{LA}, u^{ta}_{LA}\}, \qquad \hat{u}_N := \max\{u^{sn}_{LA}, u^{tn}_{LA}\}, \qquad \hat{u}_H := \max\{1, G-\alpha\}.$$

Because G is fixed, these three quantities do not depend on δ. By Lemma 3, $\hat{u}_A > \hat{u}_N$. By Lemma 4, an AI action is optimal in state LA if and only if $\hat{u}_A > (1-\delta)\hat{u}_N + \delta\hat{u}_H$. If $\hat{u}_A \geqslant \hat{u}_H$, this inequality holds for every $\delta < 1$, since $\hat{u}_A > \hat{u}_N$. In this case we may set $\hat{\delta} = 1$.

Suppose instead that $\hat{u}_A < \hat{u}_H$. Since $\hat{u}_N < \hat{u}_A < \hat{u}_H$, the right-hand side of the preceding inequality is strictly increasing in δ, from $\hat{u}_N$ at $\delta = 0$ to $\hat{u}_H$ at $\delta = 1$. Hence there is a unique cutoff

$$\hat{\delta} = \frac{\hat{u}_A - \hat{u}_N}{\hat{u}_H - \hat{u}_N} \in (0, 1)$$

at which the researcher is indifferent. AI is strictly optimal for $\delta < \hat{\delta}$, while a no-AI action followed by transition to H is strictly optimal for $\delta > \hat{\delta}$. This proves the cutoff claim.

It remains to show that $\hat{\delta} < 1$ when G is sufficiently large. For sufficiently large G, the long-term project gives the highest period payoff in each of the relevant comparisons, so

$$\hat{u}_A = p_A G - \alpha p_A^2, \qquad \hat{u}_N = p_N G - \alpha p_A^2, \qquad \hat{u}_H = G - \alpha.$$

Moreover, since $p_A < 1$,

$$\hat{u}_H - \hat{u}_A = (1 - p_A)G - \alpha(1 - p_A^2) = (1 - p_A)\big[G - \alpha(1 + p_A)\big] > 0$$

for all sufficiently large G. Thus the cutoff is interior, and

$$\hat{\delta} = \frac{(p_A - p_N)G}{(1 - p_N)G - \alpha(1 - p_A^2)} \longrightarrow \frac{p_A - p_N}{1 - p_N} < 1$$

as $G \to \infty$. Hence $\hat{\delta} < 1$ for all sufficiently large G.

□

Proof of Proposition 3. (i) Because tn is chosen in state LN if and only if $\alpha < \bar{\alpha}_L$, and sn beats tn if and only if $\alpha > \bar{\alpha}_N$, reversal requires $\bar{\alpha}_N < \alpha < \bar{\alpha}_L$.

(ii) On the interval $\bar{\alpha}_N < \alpha < \bar{\alpha}_L$, the best no-AI action in state LA is sn. By Lemma 3, any optimal no-AI action in state LA is followed by transition to H, so the average present value of choosing sn in state LA is at least $(1-\delta)\big(p_N - \alpha(p_A - p_N)^2\big) + \delta\hat{u}_H$. If this average present value exceeds both AI period payoffs, p_A and $p_A G - \alpha(p_A)^2$, then sn is optimal in state LA. Since tn is chosen in state LN, reversal occurs.

□

Proof of Proposition 4. The sufficient condition follows from Lemma 5. It remains to show when such parameter values exist. The inequalities $Gp_N - \alpha(p_N)^2 > 1 > Gp_A - \alpha(p_A)^2$ are equivalent to $\frac{Gp_A - 1}{(p_A)^2} < \alpha < \frac{Gp_N - 1}{(p_N)^2}$.

Finally, this interval is nonempty if and only if $\frac{Gp_A-1}{(p_A)^2} < \frac{Gp_N-1}{(p_N)^2}$, that is, if and only if $G > \frac{p_N+p_A}{p_N p_A} = \frac{1}{p_N} + \frac{1}{p_A}$. □

Proof of Proposition 5. By assumption, the short-term project without AI succeeds with a higher probability than the short-term project with AI in the high human capital states. This makes the short-term no-AI action the temptation benchmark, so it incurs zero self-control cost. Because the short-term no-AI action also yields a higher probability of retaining high human capital, it strictly dominates the short-term AI action. Similarly, both long-term actions yield a direct current payoff of zero and thus face the exact same self-control cost. However, by assumption, the long-term project without AI yields a higher delayed success probability and a higher probability of retaining high human capital. Thus, the long-term no-AI action strictly dominates the long-term AI action. Because every AI action is dominated by its non-AI counterpart, AI is never chosen in the high human capital states. □

Proof of Proposition 6. For each J, let $D_J := \max V_{LJ}, V_{HJ} - V_{LJ} \geqslant 0$.

(i) Fix a state IJ and a feasible technology ℓ, and compare the long-term action $t\ell$ with the short-term action $s\ell$. The difference between the values from choosing these actions and behaving optimally thereafter is

$$(1-\delta)\left[p_{IJ}^{t\ell}G - p_{IJ}^{s\ell} - \alpha\left(c(T_{IJ}) - c(T_{IJ} - p_{IJ}^{s\ell})\right)\right] + \delta\left(h_{IJ}^{t\ell} - h_{IJ}^{s\ell}\right) D_J.$$

The continuation term is nonnegative by Assumption 2. Moreover,

$$c(T_{IJ}) - c(T_{IJ} - p_{IJ}^{s\ell}) \leqslant c(T_{IJ}) \leqslant c(1) = 1.$$

It follows that $t\ell$ is strictly preferred to $s\ell$ whenever $\alpha < p_{IJ}^{t\ell}G - p_{IJ}^{s\ell}$.

Assumption 2 implies that every difference $p_{IJ}^{t\ell}G - p_{IJ}^{s\ell}$ is strictly positive, so their minimum, denoted by $\underline{\alpha}$, is strictly positive. If $\alpha < \underline{\alpha}$, the long-term project is strictly preferred to the short-term project for every technology in every state. Therefore only the long-term project is chosen.

(ii) Let $\underline{T} := \min_{IJ} T_{IJ} > 0$. Every period payoff is at most G: a short-term action yields at most $1 < G$, while a long-term action yields at most G, and self-control costs are nonnegative, so $V_{IJ} \leqslant G$ in every state.

Every long-term action has zero current payoff and therefore incurs a self-control cost of at least $\alpha c(\underline{T})$. The value from choosing any long-term action and behaving optimally thereafter is therefore at most $G - (1-\delta)\alpha c(\underline{T})$. This bound is negative whenever

$$\alpha > \overline{\alpha} := \frac{G}{(1-\delta)c(\underline{T})}.$$

By contrast, in every state the researcher can choose a short-term action that attains T_{IJ}. This action incurs no self-control cost and has a nonnegative value. Hence, if $\alpha > \overline{\alpha}$, every long-term action is strictly worse than some short-term action, so only the short-term project is chosen.

(iii) Consider state LA. As $\delta \to 0$ with G fixed, the weight on continuation values vanishes, and the agent's choice is governed strictly by the period payoffs $u^{k\ell}_{LA}$:

$$\begin{aligned}
u^{sa}_{LA} &= p^{sa}_{LA} \\
u^{sn}_{LA} &= p^{sn}_{LA} - \alpha c\left(T_{LA} - p^{sn}_{LA}\right) \\
u^{ta}_{LA} &= p^{ta}_{LA} G - \alpha c(T_{LA}) \\
u^{tn}_{LA} &= p^{tn}_{LA} G - \alpha c(T_{LA})
\end{aligned}$$

By Assumption 4, AI strictly increases the success probability for both project types: $p^{sa}_{LA} > p^{sn}_{LA}$ and $p^{ta}_{LA} > p^{tn}_{LA}$. For the short-term project, sa attains the temptation benchmark ($T_{LA} = p^{sa}_{LA}$) and incurs no self-control cost, so $u^{sa}_{LA} > u^{sn}_{LA}$. For the long-term project, both actions incur the same self-control cost, but ta provides a higher expected delayed return, so $u^{ta}_{LA} > u^{tn}_{LA}$. Thus, at $\delta = 0$, every no-AI action is strictly dominated in period payoff by its AI counterpart. Because the action set is finite, this strict payoff advantage is positive. Action values vary continuously with δ, so some AI action remains strictly optimal for all sufficiently small positive δ. Hence, there exists $\underline{\delta} \in (0, 1]$ such that AI is used in state LA whenever $\delta < \underline{\delta}$.

(iv) Assume that AI does not generate human capital in the low state and that the no-AI long-term action leaves high human capital unchanged.

Because current-payoff terms and self-control costs are bounded, while the delayed payoff from the long-term project grows with Γ, there exists Γ_0 such that, for every $\Gamma > \Gamma_0$ and all δ sufficiently close to 1, ta is the best AI action in LA and tn is the best no-AI action in H. Define $U_A := u^{ta}_{LA}$ and $U_H := u^{tn}_{HA}$. Since $U_H - U_A = \delta\Gamma(p^{tn}_{HA} - p^{ta}_{LA}) - \alpha[c(T_{HA}) - c(T_{LA})]$,

the assumption $p^{tn}_{HA} > p^{ta}_{LA}$ implies that $U_H > U_A$ for every sufficiently large Γ and all δ sufficiently close to 1.

Now consider the feasible no-AI policy that chooses tn in both L and H. Because $h^{tn}_{HA} = 1$, its value in H is U_H. Its value W_L from L satisfies $W_L = (1-\delta)u^{tn}_{LA} + \delta[(1 - h^{tn}_{LA})W_L + h^{tn}_{LA}U_H]$, so $W_L - U_H = (1-\delta)(u^{tn}_{LA} - U_H)/[1 - \delta(1 - h^{tn}_{LA})]$. For fixed Γ, the right-hand side converges to zero as $\delta \uparrow 1$, so $W_L - U_H \to 0$. Because $U_H > U_A$ for δ sufficiently close to 1, there exists $\overline{\delta} < 1$ such that $W_L > U_A$ for every $\delta > \overline{\delta}$.

It remains to show that $W_L > U_A$ rules out AI. Suppose, toward a contradiction, that an AI action is optimal in state LA. For sufficiently large Γ, ta is the best AI action, and by assumption every AI action in LA has $h^{ka}_{LA} = 0$. The Bellman equation would therefore imply $V_{LA} = (1-\delta)U_A + \delta V_{LA}$, and hence $V_{LA} = U_A$. On the other hand, the no-AI policy that chooses tn in both states is feasible, so $V_{LA} \geqslant W_L > U_A$, a contradiction. Therefore AI is not used in state LA.

Thus there exists Γ_0 such that, for every $\Gamma > \Gamma_0$, there is an $\overline{\delta}(\Gamma) < 1$ such that AI is not used whenever $\delta > \overline{\delta}(\Gamma)$.

□

Proof of Proposition 7. Fix a low-human-capital state LJ, and let V denote the original value function. Suppose that a long-term action $t\ell^*$ is optimal in state LJ. Define

$$D_J := \max\{V_{LJ}, V_{HJ}\} - V_{LJ} \geqslant 0.$$

Consider any short-term action $s\ell$, write

$$h := h^{s\ell}_{LJ} \qquad \text{and} \qquad M_J := \max\{V_{LJ}, V_{HJ}\},$$

and reduce h to $h - \varepsilon$, where $0 < \varepsilon \leqslant h$. Holding the original value function V fixed, this changes the continuation value of $s\ell$ by

$$\begin{aligned}&\delta\big[(1 - h + \varepsilon)V_{LJ} + (h - \varepsilon)M_J\big] - \delta\big[(1-h)V_{LJ} + hM_J\big]\\ &= -\delta\varepsilon(M_J - V_{LJ}) = -\delta\varepsilon D_J \leqslant 0.\end{aligned}$$

Thus the action value of $s\ell$, evaluated at the original value function, weakly falls by $\delta\varepsilon D_J$. Its period payoff is unchanged. The primitives of every other action are unchanged, so their action values, evaluated at the original value function, are unchanged.

In particular, the action value of $t\ell^*$ is unchanged. Since $t\ell^*$ was optimal before the reduction and the only affected competing action has become weakly less attractive, $t\ell^*$ remains optimal when the modified Bellman operator is evaluated at V. Consequently, the maximized Bellman value in state LJ is unchanged. The Bellman operator in every other state is also unchanged when evaluated at V. Therefore, if $\widetilde{T}$ denotes the modified Bellman operator,

$$\widetilde{T}V = TV = V.$$

Thus V is a fixed point of the modified Bellman operator. Because $\widetilde{T}$ is a contraction with modulus $\delta < 1$, it has a unique fixed point. Hence the value function after the reduction is also V. In particular, $t\ell^*$ remains optimal in state LJ, so the long-term project is still chosen. □

References

Acemoglu, D., D. Kong, and A. Ozdaglar (2026). *AI, Human Cognition and Knowledge Collapse*. Working Paper. National Bureau of Economic Research.

Acemoglu, D. and P. Restrepo (2018). "The race between man and machine: Implications of technology for growth, factor shares, and employment". *American economic review*, 108, 1488–1542.

Agrawal, A., J. Gans, and A. Goldfarb (2018). *Prediction Machines: The Simple Economics of Artificial Intelligence*. Boston, MA: Harvard Business Review Press.

Agrawal, A. K., J. S. Gans, and A. Goldfarb (2025). *The Economics of bicycles for the mind*. Tech. rep. National Bureau of Economic Research.

Allcott, H., M. Gentzkow, and L. Song (2022). "Digital addiction". *American Economic Review*, 112, 2424–2463.

Autor, D. H. (2015). "Why Are There Still So Many Jobs? The History and Future of Workplace Automation". *Journal of Economic Perspectives*, 29, 3–30.

Autor, D. H., F. Levy, and R. J. Murnane (2003). "The Skill Content of Recent Technological Change: An Empirical Exploration". *The Quarterly Journal of Economics*, 118, 1279–1333.

Bastani, H. et al. (2025). "Generative AI without guardrails can harm learning: Evidence from high school mathematics". *Proceedings of the National Academy of Sciences*, 122, e2422633122.

Beane, M. (2019). "Shadow learning: Building robotic surgical skill when approved means fail". *Administrative Science Quarterly*, 64, 87–123.

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

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

Davies, B. (2026). "Curriculum design in the age of AI". *arXiv preprint arXiv:2607.18735*.

Dell'Acqua, F. et al. (2023). "Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality". *Harvard Business School Technology & Operations Management Unit Working Paper*.

Fudenberg, D., G. Georgiadis, and L. Rayo (Oct. 2021). "Working to Learn". *Journal of Economic Theory*, 197, 105347.

Fudenberg, D. and D. K. Levine (2006). "A dual-self model of impulse control". *American economic review*, 96, 1449–1476.

Fudenberg, D. and L. Rayo (Nov. 2019). "Training and Effort Dynamics in Apprenticeship". *American Economic Review*, 109, 3780–3812.

Garicano, L. and L. Rayo (Sept. 2017). "Relational Knowledge Transfers". *American Economic Review*, 107, 2695–2730.

Gul, F. and W. Pesendorfer (2001). "Temptation and self-control". *Econometrica*, 69, 1403–1435.

Hogg, D. W. (2026). "Why do we do astrophysics?" *arXiv preprint arXiv:2602.10181*.

Ide, E. (2025). *Automation, AI, and the Intergenerational Transmission of Knowledge*. Discussion Paper. CEPR.

Karamanis, M. (2026). *The Machines Are Fine*. URL: https://ergosphere.blog/posts/the-machines-are-fine/ (visited on 04/26/2026).

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

O'Donoghue, T. and M. Rabin (1999). "Doing it now or later". *American economic review*, 89, 103–124.

Shaw, S. D. and G. Nave (2026). "Thinking-Fast, Slow, and Artificial: How AI is Reshaping Human Reasoning and the Rise of Cognitive Surrender". *Available at SSRN 6097646*.

Shen, J. H. and A. Tamkin (2026). "How AI impacts skill formation". *arXiv preprint arXiv:2601.20245*.

Snower, D. (1996). "The low-skill, bad-job trap". *Acquiring skills: Market failures, their symptoms and policy responses*, 111–24.