Artificial Intelligence Progress, the De-Skilling Trap, and Protectionism

John Gilbert Onur A. Koska Reza Oladi*

Preliminary Draft –July 2026

Abstract

We study how progress in artificial intelligence generates de-skilling in an economy that produces AI. Progress in AI displaces skilled workers from tasks while simultaneously competing for their labor in production. The resulting fall in the skill premium discourages education investment, contracting the skilled workforce and constraining both task performance and AI output–a self-reinforcing *de-skilling trap*. We then ask which policy instrument best escapes the trap. Among the three we consider–a tariff on high-tech imports, an education subsidy, and an AI production subsidy–the education subsidy is the most direct remedy. It relaxes the binding skilled-labor constraint without distorting the task allocation. A tariff on high-tech imports raises the skilled wage by Stolper–Samuelson and attenuates the skill premium collapse, acting as a complement to the education subsidy. An AI production subsidy, by contrast, distorts the task margin and does not relax the constraint.

1 Introduction

Artificial intelligence (AI) is reshaping labor markets at a pace that outstrips our theoretical understanding of its distributional consequences. A growing literature–building on Acemoglu and Restrepo (2018, 2019), Dornbusch et al. (1977), and Jones (1971)–models AI as a technology that substitutes for human labor across a range of tasks, potentially creating demand for new ones. Much of this work, however, abstracts from the fact that the production of AI–its

*Corresponding author. Department of Agricultural and Resource Economics, Utah State University. Email: reza.oladi@usu.edu.

engineering, data infrastructure, and computational services–is itself a major industry that competes directly for the very skilled labor it is said to threaten.

This paper asks a different question. When an economy produces AI and workers choose endogenously whether to acquire education, does AI progress generate *de-skilling*? The answer we find is yes, and the mechanism is self-reinforcing. AI progress lowers the skill premium, discouraging investment in education. The resulting contraction of the skilled workforce tightens the labor market for both task performance and AI production simultaneously–a double-squeeze that can shrink the high-tech sector even as AI makes it more productive. We call this the *de-skilling trap*. We then ask which policy instrument best escapes the trap: a tariff on high-tech imports, an education subsidy, or an AI production subsidy. The education subsidy emerges as the most direct remedy since it acts on the binding skilled-labor constraint itself, while a tariff on high-tech serves as a complementary instrument that operates through Stolper–Samuelson on the skill premium, and an AI production subsidy distorts the task margin without relaxing the constraint.

Our framework intersects three strands of the literature. First, the task-based approach to automation (Acemoglu and Restrepo, 2018, 2019), itself built on the Ricardian continuum of Dornbusch et al. (1977), provides the margin along which AI displaces labor. Second, the specific-factors tradition (Jones, 1971; Neary, 1978) presents the general-equilibrium machinery through which displacement is transmitted to factor prices. Third, Roy-type selection (Roy, 1951) provides the margin along which workers decide whether to become skilled. In most of the automation literature the skill composition of the workforce is taken as given, so the return to skill may fall without anyone's decision to acquire skill being disturbed. Closing that gap is what this paper does, and our closest antecedent is Gilbert et al. (2026)–hereafter GKO–which studies technical progress in AI in a open economy, its effects on wages and asks whether a trade tariff or a direct AI tax can counter it. GKO also incorporates AI production in the economy as an intermediate sector that builds AI from capital and skilled labor. Thus, the two-sided exposure of skilled workers–displaced from tasks by the very good they help produce–is present there as well. We depart in two respects, and the second gives the first its force. First, GKO holds the skilled and unskilled labor endowments fixed. There is no education margin, so a compressed premium has nowhere to propagate and de-skilling cannot arise. Second, GKO carries three final goods against three primary factors. Technically, this is an *even* structure in which the zero-profit conditions determine all factor prices from the cost side alone, a technology shock is absorbed into a determinate price vector, and a chain of factor-intensity rankings signs every response. Although, the substitutability between skilled labor and AI is present in that model, it cannot overturn any sign and it is masked by the structure itself. In our alternative formulation in this paper, reducing the goods dimension to two traded goods and three primary factors–a specific-factors (2×3) economy– makes the quantity side of the economy

inseparable to the price side. Hence, the factor bias of AI progress acts as an independent force. The sign of the skilled-wage response is then no longer settled by intensity rankings and it rather turns on an explicit substitution-elasticity condition, which we regard as the substantive content of the modeling choice rather than a nuisance. The difference is not merely formal. In GKO the skilled labor released from tasks is absorbed elsewhere and the unskilled wage *rises*, narrowing the wage gap. Here, in this paper, both wages fall, the premium compresses from above, and the education margin gives way. This is precisely what opens the de-skilling mechanism, the trap, and the policy comparisons that follow.

Our main contributions are the following. Under a plausible displacement condition, AI progress lowers the skilled wage, compresses the skill premium, and induces workers at the education margin to forgo schooling–a *de-skilling* mechanism that cannot arise where the skill composition of the workforce is taken as given. Endogenous skill acquisition dampens all factor price effects relative to a fixed-supply benchmark, consistent with the Le Châtelier principle (Samuelson, 1947; Silberberg, 1971), while amplifying sectoral reallocation. *De-skilling* moderates itself through the price side but reinforces itself through the quantity side. If the elasticity of skill supply exceeds a critical threshold, AI progress reduces output of the high-tech sector, the task composite, and domestic AI production simultaneously–what we coin *the de-skilling trap*. The trap is self-reinforcing: falling skill premia reduce skilled labor supply, which constrains both task performance and AI production, which further depresses returns to skill. Finally, we compare three instruments by their ability to mitigate the trap, as measured by their effect on the critical threshold. A tariff on high-tech imports raises the skilled wage through Stolper–Samuelson, attenuating the premium compression that drives de-skilling–though this direction depends on the economy's comparative-advantage pattern rather than on the mechanism itself. An education subsidy acts on the binding constraint directly, expanding skilled supply without disturbing the task margin, and the two instruments prove complementary: one works through the premium, the other through quantities. An AI production subsidy, by contrast, pulls the task threshold further toward AI and draws skilled labor into AI production without enlarging the pool it is drawn from, deepening the very squeeze it is meant to relieve.

The paper is organized as follows. Section 2 presents the model. Section 3 shocks the economy through progress in AI technology. Section 4 develops the de-skilling mechanism. It first establishes the mechanism under fixed skill supply (Section 4.1), then introduces endogenous skill acquisition, its self-limiting property, and the de-skilling trap (Section 4.2). Section 5 introduces policies and compares three instruments–a tariff on high-tech, an education subsidy, and an AI production subsidy–by their ability to mitigate the trap. Section 6 concludes.

2 The Setup

Consider a small open economy that produces two traded final goods and AI by using three primary factors: capital K, skilled labor S, and unskilled labor L. A high-tech sector, denoted by h, produces its traded final good using capital K_h and the task composite T under constant returns to scale technology $Y_h = F_h(K_h, T)$. The service sector s produces also a traded final good using capital K_s and unskilled labor L with constant returns to scale technology $Y_s = F_s(K_s, L)$. We maintain that good h is the import-competing good with domestic price equal to the world price, $p_h = p_h^*$, until trade policy is introduced in Section 5. Good s is the exportable, and we adopt it as the numéraire: $p_s = p_s^* = 1$. Sector a produces AI using capital K_a and skilled labor S_a with production technology $Q = F_a(K_a, S_a)$, where F_a exhibits constant returns to scale.

Following Gilbert et al. (2026), the task composite T aggregates a unit continuum of tasks $\omega \in [0, 1]$. Each task can be performed by skilled labor or by AI, with unit requirements $a_L(\omega)$ and $a_I(\omega, \delta)$ respectively, where δ parameterizes the state of AI technology. Define the relative efficiency of skilled labor as $\alpha(\omega, \delta) \equiv a_I(\omega, \delta)/a_L(\omega)$. We maintain that $\alpha_\omega(\omega, \delta) < 0$ and $\alpha_\delta(\omega, \delta) < 0$ for all (ω, δ). Skilled labor has comparative advantage in low-ω tasks, and AI progress reduces skilled labor's relative efficiency across the entire task spectrum.

Cost minimization implies a threshold $\bar{\omega}$ such that:

$$\alpha(\bar{\omega}, \delta) = \frac{w}{p_a}. \tag{1}$$

Skilled labor performs tasks $\omega \in [0, \bar{\omega}]$ while AI performs tasks $\omega \in (\bar{\omega}, 1]$. Then, the unit cost of the task composite is given by:

$$p_T = B(\bar{\omega})w + A(\bar{\omega}, \delta)p_a, \tag{2}$$

where $B(\bar{\omega}) = \int_0^{\bar{\omega}} a_L(\omega)\, d\omega$ and $A(\bar{\omega}, \delta) = \int_{\bar{\omega}}^1 a_I(\omega, \delta)\, d\omega$ are the skilled labor and AI requirements per unit of T.

Lemma 1. *By the envelope theorem, $\partial p_T/\partial \bar{\omega} = 0$ at the cost-minimizing $\bar{\omega}$.*

Proof. $\partial p_T/\partial \bar{\omega} = a_L(\bar{\omega})\, w - a_I(\bar{\omega}, \delta)\, p_a = a_L(\bar{\omega})[w - \alpha(\bar{\omega}, \delta)\, p_a] = 0$ by (1). □

The intuition is the standard envelope logic applied to the cost-minimizing task margin. The threshold $\bar{\omega}$ is chosen precisely so that the marginal task is one for which the firm is indifferent between the two ways of performing it: at $\bar{\omega}$, the skilled-labor cost $a_L(\bar{\omega})\, w$ exactly equals the AI cost $a_I(\bar{\omega}, \delta)\, p_a$, which is what condition (1) states. Because the two methods cost the same at that boundary, shifting $\bar{\omega}$ by an infinitesimal amount–reassigning the marginal task from skilled labor to AI or vice versa–saves on one input exactly what it adds on the other,

leaving the unit cost of the task composite p_T unchanged to first order. The reallocation is, at the margin, a pure wash. The practical payoff is that when we evaluate how a change in AI quality δ affects p_T, we may treat $\bar{\omega}$ as fixed and ignore the induced movement of tasks across the threshold: only the direct effect of δ on the AI input requirement matters to first order. This is the task-margin analogue of Shephard's lemma–the indirect effect operating through the optimized choice variable vanishes at the optimum.

There is a continuum of workers with mass N. In the spirit of Roy (1951), each worker draws education cost κ from a continuous CDF $G(\kappa)$ with density $g(\kappa) > 0$ on $[0, \infty)$. A worker paying κ becomes skilled and earns w, otherwise he earns the unskilled wage w^U. He acquires skill whenever the skill premium $\pi \equiv w - w^U$ covers his cost, so the marginal worker is indifferent at

$$\kappa = \pi. \tag{3}$$

Skilled and unskilled supplies are $S = N \cdot G(\pi)$ and $L = N[1 - G(\pi)]$, since a worker educates exactly when $\kappa \leq \pi$. The elasticity of skill supply is given by:

$$\eta \equiv \frac{g(\pi)\,\pi}{G(\pi)} > 0. \tag{4}$$

Capital is perfectly mobile across sectors, implying:

$$K_h + K_s + K_a = \bar{K}. \tag{5}$$

Skilled labor is allocated between tasks and AI production. This, along with full employment of skilled labor, imply:

$$S_T + S_a = S \tag{6}$$

where $S_T = B(\bar{\omega}) \cdot T$. Given market clearing and full employment in unskilled labor market, we have:

$$L_s = L \tag{7}$$

Finally, the market clearing condition in AI market implies AI production equals its demand:

$$Q = A(\bar{\omega}, \delta) \cdot T. \tag{8}$$

We maintain that all markets are perfectly competitive. Hence, zero-profit implies:

$$p_h = c_h(r, p_T) \tag{9}$$

$$p_s = 1 = c_s(r, w^U) \tag{10}$$

$$p_a = c_a(r, w). \tag{11}$$

An equilibrium consists of a vector of factor prices (r, w, w^U), price vector (p_a, p_T), labor supply vector (S, L), allocation vector $(K_h, K_s, K_a, S_T, S_a)$, output vector (Y_h, Y_s, Q, T), the marginal task $\bar{\omega}$, and the education cutoff π.

3 AI Progress

Before we characterize the impacts of progress in AI on this economy, we shall reduce our multi-sector economy to an effectively compact two-sector specific-factor economy, consisting of (what we refer to as) the *high-tech complex*[1] and the service sector.

Since AI is produced solely as an input to the task composite, which in turn is used solely in sector h, the *effective unit cost* of good h is:

$$c_h^*(r, w; \delta) \equiv c_h\big(r, p_T^*(r, w; \delta)\big), \tag{12}$$

embedding the optimal $\bar{\omega}^*(r, w; \delta)$, the AI price $p_a = c_a(r, w)$, and the task pricing equation (2).

Lemma 2. *The effective capital and skilled-labor cost shares of good h are:*

$$\theta_{Kh}^* = \theta_{Kh} + (1 - \theta_{Kh})(1 - \theta_S)\,\theta_{Ka},$$
$$\theta_{Sw}^* = (1 - \theta_{Kh})\big[\theta_S + (1 - \theta_S)(1 - \theta_{Ka})\big] = 1 - \theta_{Kh}^*,$$

where θ_{Kh}, θ_S, and θ_{Ka} are the capital cost share in sector h, the skilled labor cost share in tasks, and the capital cost share in AI production, respectively.

Proof. One dollar of good h decomposes into: capital in h (share θ_{Kh}); and the task composite (share $1 - \theta_{Kh}$), which itself splits into skilled labor in tasks (share θ_S) and AI inputs (share $1 - \theta_S$), the latter splitting further into capital in AI (share θ_{Ka}) and skilled labor in AI (share $1 - \theta_{Ka}$). Summing capital and skilled-labor components delivers (10)–(11). they exhaust all primary factor costs, so $\theta_{Kh}^* + \theta_{Sw}^* = 1$. □

The economy reduces to a two-sector specific-factors model. The high-tech complex (H) uses capital and skilled labor with effective cost $c_h^*(r, w; \delta)$ at traded price p_h. Services (s) uses capital and unskilled labor with cost $c_s(r, w^U)$ at world price $p_s = 1$. Capital is the mobile factor while skilled and unskilled labor are each sector-specific. With two traded goods and three primary factors, the price system is underdetermined and the additional relationship comes from the quantity side–the allocation of mobile capital across sectors.

Let $\hat{x} \equiv dx/x$ denote proportional changes. With s as numéraire, $\hat{p}_s = 0$ identically. With h traded at the fixed world price $p_h = p_h^*$, we likewise have $\hat{p}_h = 0$

[1] We use the term "high-tech complex" to denote the aggregated production unit that combines domestic AI production and the high-tech good–formally, the sector absorbing effective capital $K_h^* = K_h + K_a$ and skilled labor $S = S_T + S_a$ (see Appendix Appendix A). This is an analytical aggregation device and is not to be confused with the established use of "complex" in economic geography, where an industrial (or high-tech) complex denotes a spatial concentration of interrelated firms (Isard and Vietorisz, 1955; Gordon and McCann, 2000).

throughout our positive analysis. AI progress alters domestic costs and factor prices, but not p_h. The zero-profit condition $p_h = c_h^*(r, w; \delta)$ adjusts through $\hat{r}$ and $\hat{w}$. The price-side equations are:

$$\hat{p}_h = \theta_{Kh}\,\hat{r} + (1 - \theta_{Kh})\,\hat{p}_T = 0, \tag{13}$$

$$\hat{p}_s = \theta_{Ks}\,\hat{r} + (1 - \theta_{Ks})\,\hat{w}^U = 0, \tag{14}$$

$$\hat{p}_a = \theta_{Ka}\,\hat{r} + (1 - \theta_{Ka})\,\hat{w}, \tag{15}$$

$$\hat{p}_T = \theta_S\,\hat{w} + (1 - \theta_S)\,\hat{p}_a + \Phi, \tag{16}$$

where

$$\Phi \equiv \frac{(1 - \theta_S)\,p_a}{p_T} \int_{\bar{\omega}}^{1} \frac{\partial a_I(\omega, \delta)}{\partial \delta}\, d\omega \cdot d\delta < 0$$

is the direct cost-reducing effect of AI progress on the task composite (Lemma 1).

The price-side equations (12)-(16) can be reduce to:

$$\theta_{Kh}^* \hat{r} + \theta_{Sw}^* \hat{w} = -(1 - \theta_{Kh})\Phi > 0 \tag{17}$$

$$\hat{w}^U = -\frac{\theta_{Ks}}{1 - \theta_{Ks}}\,\hat{r}. \tag{18}$$

where, we have substituted (15) into (16) and then using (13) with $\hat{p}_h = 0$ to eliminate $\hat{p}_T$ gives (17) after applying Lemma 2. Equation (18) follows directly from (14) with $\hat{p}_s = 0$.

Equation (17) is the high-tech locus, driven entirely by AI progress through Φ. The services equation (18) ties $\hat{w}^U$ to $\hat{r}$ alone: with s as numéraire, the only relative-price movement is forced into the high-tech complex through the changes in factor prices that clear the zero-profit condition for h. The equilibrium point on (17) is pinned by the quantity side, to which we shall turn below. Section 5 returns to (17) to introduce a tariff on h, which shifts its right-hand side by $\hat{p}_h > 0$.

Two price-side loci alone do not pin down $(\hat{r}, \hat{w})$: equation (17) provides one relation, but a second is needed. That second relation comes from the quantity side–factor market clearing combined with cost-minimization in each sector.

Under the benchmark case with $(\bar{K}, \bar{S}, \bar{L})$ exogenous, full employment of the three primary factors in hat form requires:

$$\lambda_{Kh}\,\hat{K}_h + \lambda_{Ks}\,\hat{K}_s + \lambda_{Ka}\,\hat{K}_a = 0, \qquad \lambda_{ST}\,\hat{S}_T + \lambda_{Sa}\,\hat{S}_a = 0, \qquad \hat{L}_s = 0,$$

where $\lambda_{Kh}, \lambda_{Ks}, \lambda_{Ka}$ are the capital allocation shares (summing to one) and $\lambda_{ST}, \lambda_{Sa}$ are the skilled-labor allocation shares between tasks and AI production (also summing to one). Cost-minimization at given factor prices, together with Shephard's lemma, gives the unit-input responses in each of the three production blocks (high-tech assembly, the task composite, and AI), parameterized by the elasticities of substitution σ_h, σ_a, and σ_s. The marginal-task condition (Lemma 1)

governs how the task boundary $\bar{\omega}$ reallocates skilled labor and AI in response to changes in w/p_a.

Substituting the cost-minimization expressions into the three full-employment equations and eliminating $\hat{w}^U$ via (18) reduces the quantity side to a single relation between $\hat{r}$ and $\hat{w}$:

$$\Omega_r \, \hat{r} + \Omega_w \, \hat{w} = \Omega_\delta, \tag{19}$$

where Ω_r, Ω_w, and Ω_δ are composite coefficients in the cost shares $\{\theta_{Ki}, \theta^*_{Sw}\}$, the substitution elasticities $\{\sigma_h, \sigma_a, \sigma_s\}$, and the allocation shares $\{\lambda_{Ki}, \lambda_{ST}, \lambda_{Sa}\}$. Unconditionally, $\Omega_w > 0$ and $\Omega_r < 0$. However, the sign of Ω_δ depends on the bias of the AI shock across factor demands in the high-tech complex. Under the displacement direction in which AI saves skilled labor more than capital per unit effective output (the bias the paper hangs on), $\Omega_\delta < 0$ when $\Phi < 0$. Equations (17) and (19) together constitute a 2×2 linear system that pins down $(\hat{r}, \hat{w})$ in terms of the AI shock Φ. Appendix A derives the explicit forms of Ω_r, Ω_w, and Ω_δ. The next section solves this system and derives the de-skilling result.

It is worth pausing on why the underdetermination is desirable. Were there a second good using capital and skilled labor–a third traded sector restoring equality between the number of traded goods and primary factors–its zero-profit condition would tie $\hat{w}$ to $\hat{r}$ directly, and AI progress would map into factor prices through cost shares alone, with its factor bias subsumed as in GKO. By dispensing with that sector we let the quantity locus (19) carry the bias explicitly through Ω_δ. The two loci then divide the shock cleanly: the price locus (17) transmits the *level* (productivity) content, Φ, while the quantity locus transmits the *bias* content, Ω_δ. Separating the two is what lets the model address the factor-bias question that motivates the paper–at the price that an intensity ranking alone no longer signs the skilled-wage response, which now turns on the relative size of Φ and Ω_δ and hence on a substitution-elasticity condition made explicit in Assumption 2.

4 The De-Skilling Mechanism

4.1 Fixed skill supply

We shall first establish the conditions under which AI progress drives de-skilling under the benchmark of fixed factor endowments. Combining the price-side locus (17) and the quantity-side relation (19) derived at the end of Section 3 gives the 2×2 system that pins down $(\hat{r}, \hat{w})$:

$$\mathbf{M} \begin{pmatrix} \hat{r} \\ \hat{w} \end{pmatrix} = \begin{pmatrix} -(1 - \theta_{Kh}) \, \Phi \\ \Omega_\delta \end{pmatrix}, \tag{20}$$

where $\mathrm{M} = \begin{pmatrix} \theta^*_{Kh} & \theta^*_{Sw} \\ \Omega_r & \Omega_w \end{pmatrix}$.

Assumption 1. *The effective capital intensity of the high-tech complex exceeds that of services:* $\theta^*_{Kh}/\theta^*_{Sw} > \theta_{Ks}/(1-\theta_{Ks})$.

This is the analogue of the Jones–Neary condition: the capital-heavy high-tech complex (including AI infrastructure) is more capital intensive than unskilled-labor-intensive services.

Through this system, AI progress raises the rental; but its effect on the skilled wage is a tension between the productivity gain, which pushes $\hat{w}$ up, and the displacement of skilled labor from tasks, which pushes it down. The sign is therefore not settled by intensity rankings alone–it turns on substitution elasticities, which we define before stating the condition. Let

$$\sigma_{SA} \equiv \frac{1}{|E_\omega|} = \left|\frac{\partial \ln \omega}{\partial \ln \alpha}\right|_{\bar{\omega}} > 0, \qquad E_\omega \equiv \left.\frac{\partial \ln \alpha}{\partial \ln \omega}\right|_{\bar{\omega}} < 0, \tag{21}$$

denote the elasticity of substitution between skilled labor and AI *at the task margin*. It measures how strongly the skilled-labor/AI input mix in the task composite responds to the relative price w/p_a. A flatter $\alpha(\cdot)$ schedule at neighborhood of $\bar{\omega}$ (larger σ_{SA}) means more tasks are reassigned per unit change in relative cost. Recall that we have two other within-sector elasticities already already introduced in the model: σ^*_h, between capital and skilled labor in the high-tech complex, and σ_s, between capital and unskilled labor in services (Appendix A).

AI progress also shifts the complex's unit input demands at constant factor prices. Write β^*_K and β^*_S for these proportional shifts, normalized by the cost-reducing term Φ, so that $\hat{a}^*_{Kh}\big|_{\hat{r}=\hat{w}=0} = \beta^*_K\,\Phi$ and $\hat{a}^*_{Sh}\big|_{\hat{r}=\hat{w}=0} = \beta^*_S\,\Phi$, where $a^*_{Kh} = \partial c^*_h/\partial r$ and $a^*_{Sh} = \partial c^*_h/\partial w$ are the complex's unit demands for capital and skilled labor. Clearly these two are not independent. Cost minimization fixes their cost-share-weighted average, $\theta^*_{Kh}\beta^*_K + \theta^*_{Sw}\beta^*_S = 1 - \theta_{Kh}$, so only the difference $b \equiv \beta^*_S - \beta^*_K$–the *factor bias* of AI progress–is left free by the cost shares. Because $\Phi < 0$, a larger β^*_K or β^*_S is a larger proportional reduction in the demand for that factor, so $b > 0$ means AI saves skilled labor more than capital. Appendix A derives these parameters and the constraint in full.

The sign of the bias, however, is not arbitrary. AI progress is skilled-labor-saving by construction. At constant factor prices the marginal-task condition (1) requires $\hat{\bar{\omega}}\big|_{\hat{r}=\hat{w}=0} = E_\delta\,\sigma_{SA}\,\hat{\delta} < 0$, where $E_\delta \equiv (\partial \ln \alpha/\partial \ln \delta)|_{\bar{\omega}} < 0$, so AI progress sheds skilled labor from the top tasks ($B(\bar{\omega})$ falls) and thus saves proportionally more on skilled labor than on capital. Hence, $b > 0$ for any positive AI progress–an inherent feature of the model, not an convenient assumption. What the cost shares leave open is only its magnitude. b is increasing in σ_{SA}, since the more

substitutable skilled labor and AI are at the margin, the larger the constant-price displacement.

The displacement bias dominates within-sector substitution when

$$\theta^*_{Kh}\,\lambda^*_{Kh}\,(\beta^*_S-\beta^*_K)\;>\;(1-\theta_{Kh})\left[\lambda^*_{Kh}\,\sigma^*_h+\frac{\lambda_{Ks}\,\sigma_s}{1-\theta_{Ks}}\right]. \tag{22}$$

Its left-hand side is increasing in σ_{SA} (through $b=\beta^*_S-\beta^*_K$) and its right-hand side is independent of σ_{SA}, so–taking the bias to range widely enough for a crossing–(22) holding with equality defines a unique threshold $\bar\sigma_{SA}$, itself increasing in σ^*_h and σ_s.

Assumption 2. *Skilled labor and AI are sufficiently substitutable at the task margin relative to factor substitution within sectors; in particular,* $\sigma_{SA}>\bar\sigma_{SA}$.

Assumption 2 is, by construction, equivalent to (22), and therefore determines the sign of the skilled wage: it falls, i.e., $\hat w<0$ (Proposition 1). However, it does not deliver de-skilling on its own. The education margin responds to the skill *premium* $\pi=w-w^U$, a level gap, and AI progress also lowers the unskilled wage ($\hat w^U<0$). The premium compresses (i.e, $d\pi=w\hat w-w^U\hat w^U<0$) only if the skilled wage falls by more in absolute terms than the unskilled wage, $w|\hat w|>w^U|\hat w^U|$. Since $w>w^U$, this is strictly stronger than $\hat w<0$. It requires σ_{SA} above a weakly larger threshold $\bar\sigma^{\pi}_{SA}\geq\bar\sigma_{SA}$, the premium threshold established in Proposition 3. De-skilling is thus not merely a falling skilled wage, but a skilled wage that falls faster, in levels, than the unskilled wage. The following proposition addresses the factor price effects of AI progress.

Proposition 1. *AI progress raises the return to capital, lowers the task-composite price, and lowers the unskilled wage. Under Assumption 2, it also lowers the skilled wage.*

Proof. AI progress is a positive technology shock to the high-tech complex, raising the return to the mobile factor. Because it is skilled-labor-saving at the task margin, the displacement bias is positive, $b>0$ (established following (21)); the explicit solution (A17) then has a negative numerator over a negative denominator, so $\hat r>0$ (see Appendix A). From (12), we conclude that $\hat p_T<0$. From (18), $\hat w^U=-[\theta_{Ks}/(1-\theta_{Ks})]\,\hat r<0$–the factor specific to the non-improving sector loses. For the skilled wage, solve the 2×2 system (19), $\hat w=(\theta^*_{Kh}\,\Omega_\delta-\Omega_r\,R_1)/\det\mathbf{M}$, with $R_1\equiv-(1-\theta_{Kh})\Phi>0$ and $\det\mathbf{M}=\theta^*_{Kh}\Omega_w-\theta^*_{Sw}\Omega_r>0$ (all terms positive). Substituting $\Omega_\delta=\lambda^*_{Kh}\,b\,\Phi$, $\Omega_w=\lambda^*_{Kh}\sigma^*_h$, and $\Omega_r=-[\lambda^*_{Kh}\sigma^*_h+\lambda_{Ks}\sigma_s/(1-\theta_{Ks})]$ (see Appendix A), with $b=\beta^*_S-\beta^*_K$,

$$\hat w=\frac{\Phi}{\det\mathbf{M}}\left[\theta^*_{Kh}\lambda^*_{Kh}b-(1-\theta_{Kh})\big(\lambda^*_{Kh}\sigma^*_h+\tfrac{\lambda_{Ks}\sigma_s}{1-\theta_{Ks}}\big)\right].$$

Since $\Phi<0$ and $\det\mathbf{M}>0$, $\hat w<0$ if and only if the bracket is positive, i.e. if and only if (22) holds, which is Assumption 2. □

The skilled-wage result marks a departure from the standard specific-factors logic. There, a productivity gain in a sector raises the return to that sector's own specific factor–so skilled labor, specific to the high-tech complex, would gain. AI progress overturns this because it is not a neutral productivity shock but skilled-labor-saving technical change that displaces skilled labor from tasks within the very sector it improves. Condition (22) is exactly the requirement that this within-sector displacement dominate the productivity gain, so that the specific factor of the improving sector is harmed, $\hat{w} < 0$–contrary to the textbook prediction.

This non-neutrality also recasts the rental result. Its sign coincides with the specific-factors benchmark but no longer rests on it: $\hat{r} > 0$ follows from the skilled-labor-saving direction of the bias ($b > 0$) rather than from neutrality, and would fail under a strongly capital-saving bias (Appendix A). Correspondingly, the intensity ranking of Assumption 1–which governs factor-price movements in the textbook model–plays no role in Proposition 1; it enters only in the premium-compression result of Proposition 3.

Recall that $E_\omega \equiv (\partial \ln \alpha / \partial \ln \omega)\big|_{\bar{\omega}} < 0$ is the elasticity of $\alpha(\bar{\omega}, \delta)$ at the margin with respect to $\bar{\omega}$. Let also $E_\delta \equiv (\partial \ln \alpha / \partial \ln \delta)\big|_{\bar{\omega}} < 0$ be the elasticity of $\alpha(\bar{\omega}, \delta)$ at the margin with respect to δ. Maintaining Assumption 2, Proposition 1 gives $\hat{w} < 0 < \hat{r}$, so $\hat{w} - \hat{r} < 0$ throughout. The following result addresses the AI-induced task displacement.

Proposition 2. *The change in the task threshold due to AI progress decomposes into two opposing effects, a negative direct effect and a positive substitution effect:*

$$\hat{\bar{\omega}} = \frac{-E_\delta \, \hat{\delta}}{E_\omega} + \frac{\theta_{Ka}(\hat{w} - \hat{r})}{E_\omega}$$

Proof. Differentiation (in logarithm) of the marginal-task condition (1) gives $E_\omega \, \hat{\bar{\omega}} + E_\delta \, \hat{\delta} = \hat{w} - \hat{p}_a$. From (15), $\hat{p}_a = \theta_{Ka} \, \hat{r} + (1 - \theta_{Ka}) \, \hat{w}$, hence $\hat{w} - \hat{p}_a = \theta_{Ka}(\hat{w} - \hat{r})$. Substituting and solving for $\hat{\bar{\omega}}$ yields the decomposition. The direct-effect term is negative since $E_\delta, E_\omega < 0$ and $\hat{\delta} > 0$. The substitution-effect term is positive since by Proposition 1 the numerator $\theta_{Ka}(\hat{w} - \hat{r})$ is negative, and the denominator E_ω is also negative. The net sign of $\hat{\bar{\omega}}$ follows by comparing these magnitudes. □

The direct effect of AI progress is the displacement of skilled labor at constant factor prices. AI progress makes AI relatively more efficient at every task, so the boundary moves toward more AI tasks. The substitution effect arises because $\hat{w} - \hat{r} < 0$ implies $\hat{w} - \hat{p}_a = \theta_{Ka}(\hat{w} - \hat{r}) < 0$, so the relative skilled-labor price w/p_a falls. With skilled labor now relatively cheaper than AI services, humans substitute back for AI at the margin, pulling $\bar{\omega}$ upward. The net displacement is negative, $\hat{\bar{\omega}} < 0$, whenever the direct effect dominates the substitution effect, i.e., $|E_\delta \, \hat{\delta}| > \theta_{Ka} \, |\hat{w} - \hat{r}|$.

The proposition also consistent with Lemma 1. Both p_T and $\bar{\omega}$ move as AI progress unfolds–$\hat{p}_T < 0$ by Proposition 1 and $\hat{\bar{\omega}}$ is generally non-zero by Proposition 2–but the envelope result says that $\bar{\omega}$'s own equilibrium movement contributes nothing to $\hat{p}_T$ at first order. the total change in p_T in (16) runs entirely through $\hat{w}$, $\hat{p}_a$, and Φ, with no separate $\hat{\bar{\omega}}$ term. These two results are complements. Displacement matters for *quantities*–task allocation and the bias of factor demands in the high-tech complex (Appendix A)–but drops out of the price-side equations at first order.

Proposition 3. *Under Assumption 1, there exists a threshold $\bar{\sigma}^{\pi}_{SA} > \bar{\sigma}_{SA}$ such that if $\sigma_{SA} > \bar{\sigma}^{\pi}_{SA}$, AI progress reduces the skill premium, i.e., $d\pi = d(w - w^U) < 0$.*

Proof. Write $d\pi = w\hat{w} - w^U\hat{w}^U$ and substitute the price-side relations (17) and (18):

$$d\pi = -\frac{w(1-\theta_{Kh})\Phi}{\theta^*_{Sw}} + \left[\frac{\theta_{Ks}\, w^U}{1-\theta_{Ks}} - \frac{w\,\theta^*_{Kh}}{\theta^*_{Sw}}\right]\hat{r}.$$

The first term is positive sincde $\Phi < 0$. The pure productivity gain, acting alone, would *raise* the premium. The coefficient of $\hat{r}$ is negative under Assumption 1 and $w > w^U$ (which give $w\,\theta^*_{Kh}/\theta^*_{Sw} > w^U\,\theta_{Ks}/(1-\theta_{Ks})$), so the equilibrium rise in the rental works to compress the premium. Hence, $d\pi < 0$ if and only if $\hat{r}$ exceeds the threshold given by:

$$\hat{r}_\pi \;\equiv\; \frac{w\,(1-\theta_{Kh})\,|\Phi|}{w\,\theta^*_{Kh} - w^U\,\theta^*_{Sw}\,\theta_{Ks}/(1-\theta_{Ks})} \;>\; 0. \tag{23}$$

By Appendix A, the equilibrium rental $\hat{r}$ is linear and strictly increasing in the displacement bias $b = \beta^*_S - \beta^*_K$, with slope independent of b; and b is itself increasing in σ_{SA}. The condition $\hat{r} > \hat{r}_\pi$ therefore holds exactly when σ_{SA} exceeds the value $\bar{\sigma}^{\pi}_{SA}$ at which $\hat{r} = \hat{r}_\pi$. Finally, $\hat{r}_\pi$ strictly exceeds the skilled-wage cutoff of Proposition 1 (their difference is proportional to $w^U\theta^*_{Sw}\theta_{Ks}/(1-\theta_{Ks}) > 0$), so $\bar{\sigma}^{\pi}_{SA} > \bar{\sigma}_{SA}$, establishing the claim. □

The result of Proposition 3 holds when skilled labor and AI are substitutable *enough*–more than Assumption 2 already requires. Assumption 2 asks for σ_{SA} large enough to make the skilled wage *fall* ($\hat{w} < 0$). Proposition 3 asks for it to be larger still, so that the skilled wage falls fast enough to drag the whole skilled–unskilled *premium* down ($d\pi < 0$). Both are thresholds on the same primitive, σ_{SA}, and the premium threshold sits strictly above the wage threshold, $\bar{\sigma}^{\pi}_{SA} > \bar{\sigma}_{SA}$. In particular the hypothesis of Proposition 3 already implies $\sigma_{SA} > \bar{\sigma}_{SA}$, i.e. Assumption 2, and hence $\hat{w} < 0$. The gap between the two thresholds makes economic sense, rather than a technicality. AI progress pushes *both* wages down–the skilled wage directly through displacement, the unskilled wage through the general-equilibrium rise in the rental–so a falling skilled wage is necessary but not sufficient for premium compression. What is required is that it fall by more,

in levels, than the unskilled wage, $w|\hat{w}| > w^U|\hat{w}^U|$; since $w > w^U$, this is a strictly higher bar than $\hat{w} < 0$ alone.

The de-skilling mechanism works as follows. AI progress lowers the returns to education by squeezing the skilled wage from above while also compressing the unskilled wage from below–but the skilled wage falls by more. The premium is what governs skill acquisition: a lower return to education weakens the incentive to become skilled, so the compression in $d\pi$ matters not only as a price effect but as a force acting on the future supply of skilled workers. The fixed-supply benchmark holds that supply constant and so cannot capture this feedback. To assess it, we must let skill acquisition respond to the premium–the task of the next subsection.

4.2 Endogenous skill acquisition

We now allow the supply of skilled workers to respond to the skill premium via $S = N \cdot G(\pi)$, so that $\hat{S} = \eta \cdot (d\pi/\pi)$. Throughout the remainder of the paper, a subscript e denotes an effect derived under endogenous skill supply, while the benchmark effect under fixed skill supply carries no subscript, as in the notation used up to this point.

Proposition 4. *With endogenous skill acquisition, all factor price effects of AI progress are dampened relative to the fixed-supply benchmark:*

$$|\hat{r}|_e < |\hat{r}|, \quad |\hat{w}|_e < |\hat{w}|, \quad |\hat{w}^U|_e < |\hat{w}^U|.$$

Dampening is monotonically increasing in η. However, sectoral reallocation is amplified: de-skilling contracts the high-tech complex and expands services by more than under fixed supply.

Proof. By Proposition 3, $d\pi < 0$, so $\hat{S} = \eta(d\pi/\pi) < 0$. Moreover, the right-hand side of the quantity-side equation in (19) gains a term $\Omega_S \hat{S}$ with $\Omega_S > 0$ and $\hat{S} < 0$. Hence the solution is $\hat{r}_e = \hat{r} + (\Omega_S/\det \mathbf{M})\,\hat{S}$, where $\det \mathbf{M} > 0$ unconditionally (all terms positive). Therefore. $|\hat{r}|$ is dampened, and since $\hat{w}^U$ and $\hat{p}_T$ are monotone in $\hat{r}$, they are dampened as well. Dampening increases in η because a larger η produces a larger $|\hat{S}|$. Sectoral amplification follows directly from (A8): $\hat{Y}_h$ falls with $\hat{S}$ and $\hat{Y}_s$ rises with $\hat{L} = -[S/(N-S)]\,\hat{S} > 0$. □

Intuitively, $d\pi < 0$, hence deskilling $\hat{S} = \eta(d\pi/\pi) < 0$, loosens the capital market for the high-tech complex (fewer skilled workers reduces demand for capital in H) and expands the unskilled workforce in services, both of which moderate the rise in r.

De-skilling is *self-limiting on the price side* but *self-reinforcing on the quantity side*. The fall in S moderates the factor price changes that caused it (dampening),

yet the contraction of the skilled workforce directly shrinks the high-tech complex through the quantity channel (amplification). These two forces set the stage for what we coin the *de-skilling trap*.

Corollary 1. *As $\eta \to 0$, the endogenous model replicates the fixed-supply benchmark. As $\eta \to \infty$, π is pinned approximately constant and all adjustment is through quantities.*

The quantity consequences of de-skilling can be severe enough to contract the very sector that AI was meant to enhance. This is the de-skilling trap, which we address formally by the following proposition.

Proposition 5. *There exists a critical elasticity $\eta^* > 0$ such that:*

(i) If $\eta < \eta^*$, AI progress expands Y_h, T, and Q; i.e., efficiency gains dominate.

(ii) If $\eta > \eta^*$, AI progress contracts Y_h, T, and Q; i.e., the de-skilling supply response dominates.

Proof. The total effect on the task composite is:

$$\hat{T} = \hat{T}\big|_{\hat{S}=0} + \frac{\partial \hat{T}}{\partial \hat{S}} \cdot \hat{S}. \tag{24}$$

The first term is positive (efficiency gains with fixed supply). The second is negative ($\partial \hat{T}/\partial \hat{S} > 0$, $\hat{S} < 0$), and since $\hat{S} = \eta(d\pi/\pi)$ it grows in magnitude with η while the first term is independent of η. At η^* the terms cancel. Hence, for $\eta > \eta^*$, $\hat{T} < 0$. The same logic applies to $Q = A \cdot T$ and $Y_h = F_h(K_h, T)$. □

The *double squeeze* makes the trap self-reinforcing through two simultaneous channels: (i) *task squeeze*, whereby fewer skilled workers perform tasks (S_T falls), hence reducing T. (ii) *AI squeeze* whereby fewer skilled workers produce AI (S_a falls), reducing Q, which further constrains T. The second channel exists only because AI is built from the same skilled labor it displaces: it is the domestic production of AI, together with the endogenous supply of skill, that closes the loop.

Corollary 2. *The critical threshold satisfies:*

$$\eta^* = -\frac{\hat{T}\big|_{\hat{S}=0}}{(\partial \hat{T}/\partial \hat{S}) \cdot (d\pi/\pi)\big|_{\eta=0}} > 0.$$

The trap is harder to trigger (larger η^*) when AI efficiency gains are large, when the supply response has a small direct effect on T, or when initial factor price changes are small.

The trap captures a fundamental paradox of AI production: the very workers that AI displaces from tasks are the same workers needed to produce AI. Economies with highly responsive education systems–where workers quickly abandon skill acquisition when returns fall–are most vulnerable, because the supply contraction feeds the trap faster than the efficiency gains can compensate.

5 Policies to Escape the De-Skilling Trap

We now compare three policy instruments–a tariff on high-tech imports, an education subsidy, and an AI production subsidy–and rank them by their ability to mitigate the de-skilling trap, as measured by their effect on the critical threshold η^*. Throughout this section we consider an economy undergoing AI progress in which the supply response is active, so that the compression of the premium contracts the skilled workforce, $\hat{S} = \eta\,(d\pi/\pi) \leq 0$, and the economy lies in the trap regime, $\eta > \eta^*$, of Proposition 5. This is the case the policies are meant to address, and we do not restate it in the results that follow.

So far, we have maintained that $p_h = p_h^*$, where p_h^* denotes the foreign price of h. Suppose the government imposes a specific tariff $t > 0$ on good h, so its domestic price becomes $p_h = p_h^* + t$. A proportional tariff change shows up in hat-form as $\hat{p}_h = \hat{t}\,[t/(p_h^* + t)]$, which is now non-zero.[2] The high-tech locus (17) picks up $\hat{p}_h$ on its right-hand side, while the services equation is unchanged since $\hat{p}_s = 0$ still holds (services remain the numéraire):

$$\theta_{Kh}^*\,\hat{r} + \theta_{Sw}^*\,\hat{w} = \hat{p}_h - (1 - \theta_{Kh})\,\Phi, \qquad \hat{w}^U = -\,\frac{\theta_{Ks}}{1 - \theta_{Ks}}\,\hat{r}. \tag{25}$$

The skill premium change under the combined AI shock and tariff is:

$$d\pi = w\,\hat{w} - w^U\,\hat{w}^U = w\,\hat{w} + \frac{\theta_{Ks}\,w^U}{1 - \theta_{Ks}}\,\hat{r}.$$

Solving $\hat{w}$ from (25) and separating the AI-only and tariff-only contributions yields:

$$d\pi = d\pi_{\text{AI}} + \frac{w\,\hat{p}_h}{\theta_{Sw}^*}. \tag{26}$$

The first term is the AI-only effect of Proposition 3 and the second is the trade-policy term.

[2]Under the opposite trade pattern–unskilled-intensive services importable and high-tech exportable–the results of this subsection reverse: a tariff on services would raise the unskilled wage by Stolper–Samuelson, compress the skill premium, and amplify de-skilling, while trade liberalization would attenuate it. The trade-policy direction we obtain here is thus conditional on the comparative-advantage pattern of the economy under study and not on the de-skilling mechanism itself, which goes through regardless.

Proposition 6. *A tariff on h combined with AI progress attenuates de-skilling relative to free trade. Equivalently, sharper import competition in high-tech ($\hat{p}_h < 0$) deepens de-skilling. Moreover, the tariff raises the critical threshold η^*, making the de-skilling trap harder to trigger.*

Proof. It directly follows from (26) that $d\pi - d\pi_{\text{AI}} = w\,\hat{p}_h/\theta^*_{Sw} > 0$ when $\hat{p}_h > 0$, since $w > 0$ and $\theta^*_{Sw} > 0$. The tariff thus makes $d\pi$ less negative. For η^*: protection shrinks $|d\pi|$ without affecting the efficiency gains $\hat{T}|_{\hat{S}=0}$, so the supply response $|\hat{S}| = \eta|d\pi/\pi|$ is smaller at any given η. From the expression for η^* in Corollary 2, a smaller $|d\pi/\pi|$ in the denominator raises η^*, making the trap harder to trigger. □

The mechanism runs through Stolper–Samuelson within the high-tech complex. A tariff on the skilled-intensive good h raises p_h, favoring the factor used intensively in h: skilled labor. The skilled wage w rises (and the capital return r rises somewhat, which by (18) pushes the unskilled wage w^U down). On net, the skill premium $\pi = w - w^U$ expands. When AI progress then hits and compresses w, the elevated premium provides a buffer: the education margin contracts by less than it would under free trade. By symmetry, exposure to cheaper high-tech imports–a fall in p^*_h relative to a fixed domestic cost structure–runs the same logic in reverse: skilled wages fall, the premium collapses further, and de-skilling intensifies. Import competition in high-tech is therefore a de-skilling shock in its own right, even before AI progress is layered on top.

Next, instead assume that the government offers a subsidy ς per worker who acquires education, financed by a lump-sum tax. The skill acquisition margin becomes $\pi = w - w^U + \varsigma$. The following result addresses the impact of this alternative policy.

Proposition 7. *An education subsidy $\varsigma > 0$ in the presence of AI progress: (i) increases the supply of skilled workers ($dS/d\varsigma > 0$); (ii) expands the high-tech sector, task composite, and AI output; (iii) contracts services (through the fall in L); (iv) dampens factor price effects further; and (v) raises η^*, expanding the non-trap regime.*

Proof. **(i)** $\partial\pi/\partial\varsigma = 1$, so $dS = Ng(\pi)\,d\varsigma > 0$. **(ii)** The increase in S directly expands the high-tech complex through (A8): more workers available for both tasks and AI production, reversing the supply channel of the trap. **(iii)** $dL = -dS < 0$. **(iv)** A larger S loosens the skilled labor constraint, moderating the rise in r (same mechanism as Proposition 4, reversed direction). **(v)** The subsidy reduces $|\hat{S}|$ at any η, raising η^* by the same argument as Corollary 2. □

Finally, consider a production subsidy τ to the AI sector, which lowers the effective user price to $p_a(1-\tau)$, shifting the marginal task condition to $\alpha(\bar{\omega}, \delta) = w/[p_a(1-\tau)]$. Note that τ enters the model only through this condition:

the education margin $\pi = w - w^U$ contains no term in τ, so the subsidy reaches the skilled-labor constraint $S = NG(\pi)$ only indirectly, through its general-equilibrium effect on (w, w^U). In this respect it differs from the education subsidy of Proposition 7, which enters π additively as $\pi = w - w^U + \varsigma$ and therefore shifts the education margin directly.

Proposition 8. *An AI production subsidy $\tau > 0$: (i) shifts the task threshold strictly toward AI, $\bar{\omega}^\tau < \bar{\omega}^*$, so that tasks are assigned to AI even where skilled labor is the lower-cost input at undistorted prices; (ii) crowds skilled labor out of task performance, $dS_T < 0$, and contracts the task composite, $dT < 0$, whenever the resulting loss in task labor outweighs the reallocation gain.*

Proof. **(i)** At undistorted prices the threshold $\bar{\omega}^*$ solves $\alpha(\bar{\omega}^*, \delta) = w/p_a$; under the subsidy it solves $\alpha(\bar{\omega}^\tau, \delta) = w/[p_a(1-\tau)]$. Since $\tau > 0$, the right-hand side rises, $w/[p_a(1-\tau)] > w/p_a$. Because $\alpha_\omega < 0$, $\alpha(\cdot, \delta)$ is strictly decreasing in its first argument, so a higher target value of α requires a strictly lower threshold: $\bar{\omega}^\tau < \bar{\omega}^*$. The interval of tasks performed by AI, $(\bar{\omega}, 1]$, therefore expands, and tasks in $(\bar{\omega}^\tau, \bar{\omega}^*]$ are shifted to AI even though skilled labor is the lower-cost input for them at the undistorted price p_a. **(ii)** Skilled labor is fully employed, $S_T + S_a = S$ by (6), and $dS \leq 0$ by the premise of this section. The threshold shift in (i) raises AI output Q and hence its skilled-labor requirement, $dS_a > 0$. The constraint then gives $dS_T = dS - dS_a < 0$: task performance loses skilled labor. The task composite T combines the marginal task reallocation (which raises measured productivity) with the fall in S_T (which lowers it), so $dT < 0$ whenever the latter dominates–formally, whenever the reduction in task labor $|dS_T|$ outweighs the efficiency gain from moving the marginal tasks to AI. In that case the subsidy deepens the double squeeze rather than relieving it. □

Corollary 2 organizes the comparison. The threshold η^* rises when the premium response $|d\pi/\pi|$ is smaller or when the supply contraction $|\hat{S}|$ is partly offset. The tariff works on the first: by Stolper–Samuelson it cushions the fall in w, shrinking $|d\pi/\pi|$ (Proposition 6). The education subsidy works on the second: it adds directly to S, damping $|\hat{S}|$ at any given premium response (Proposition 7). The AI production subsidy works on neither and instead lowers η^*, drawing tasks and skilled labor toward AI without enlarging the skilled pool (Proposition 8).

Corollary 3. *The tariff and the education subsidy raise η^* through distinct channels–the premium response and the supply level–so their combined effect exceeds that of either instrument alone.*

Proof. The tariff shrinks $|d\pi/\pi|$ (Proposition 6) and the education subsidy reduces $|\hat{S}|$ at a given premium response (Proposition 7). Both raise η^* by Corollary 2, and because they act on separate channels neither offsets the other; acting together they raise η^* by more than either alone. □

6 Conclusion

We began this paper with a puzzle about where AI leaves skilled workers. The usual worry is that automation hollows out routine, less-skilled work. We have argued that when an economy produces AI at home, and when workers choose whether to acquire skills, the pressure can fall precisely on the skilled workforce–and that this reversal is not incidental but structural, rooted in the fact that the same skilled labor AI displaces from tasks is also the labor that builds AI.

The mechanism runs as follows. AI progress displaces skilled labor from the tasks it used to perform, and the accompanying rise in the return to capital outweighs whatever productivity gain accrues to the skilled workforce, so the skilled wage falls. As the premium to education compresses, fewer workers at the margin choose to acquire skills, and the skilled labor supply contracts. This is where the education margin matters. When the skill composition of the workforce is taken as given, a falling premium is the end of the story. Here, it is the beginning of one, because the premium is precisely what governs who becomes skilled. Moreover, because the economy's capacity to produce AI is built from the very skilled labor whose returns are eroding, that contraction feeds back on the sector the technology was supposed to strengthen.

Once education responds strongly enough to those eroding returns, the mechanism can turn on itself. Fewer skilled workers are available both to perform tasks and to produce AI–a double squeeze that can shrink the high-tech sector even as AI makes it technically more productive. The trap is self-reinforcing: falling returns to education thin out the workforce that uses and builds AI, which depresses those returns further. What looks locally like a productivity improvement can, in general equilibrium, leave the high-tech complex smaller than it started.

The policy analysis follows from the shape of the trap. Because the binding constraint is the supply of skilled labor, the instrument that acts on it directly is the most natural remedy: an education subsidy expands the skilled workforce without distorting the task margin, and its benefits cascade through both task performance and AI production. A tariff on high-tech imports helps for a different reason–by Stolper–Samuelson it raises the skilled wage and cushions the premium collapse–but it works on the price margin rather than the quantity constraint, so it complements an education subsidy without substituting for it. An AI production subsidy, by contrast, pulls in the wrong direction. It distorts the task allocation and draws skilled labor into AI production without enlarging the pool from which that labor is drawn.

This comparison carries a topical reading. An economy that lets its high-tech sector face intensifying import competition accelerates de-skilling at home, and protection can slow it. But the deeper lesson is that the trap is a supply-

side phenomenon, and the most direct response operates through the education margin rather than the border: trade policy treats a symptom, while the education subsidy addresses the cause. speak of a first-best–requires specifying preferences and the cost of public funds; we leave that comparison, together with the political economy of protection, for future work.

Appendix A Mathematical Derivation

This appendix derives the explicit forms of the composite coefficients Ω_r, Ω_w, and Ω_δ appearing in equation (19) as well as the solution for our system.

Effective two-sector reduction: Folding AI production into the high-tech complex gives an effective two-sector model: the high-tech complex absorbs effective capital $K_h^* \equiv K_h + K_a$ and skilled labor $S = S_T + S_a$; services absorbs K_s and unskilled labor L. The effective cost function for the complex is $c_h^*(r, w; \delta)$ with cost shares θ_{Kh}^* and $\theta_{Sw}^* = 1 - \theta_{Kh}^*$ (Lemma 2). Let $\sigma_h^* > 0$ denote the effective elasticity of substitution between K and S in the high-tech complex, and $\sigma_s > 0$ the elasticity in services. The capital allocation shares $\lambda_{Kh}^* = K_h^*/\bar{K}$ and $\lambda_{Ks} = K_s/\bar{K}$ sum to one.

Unit-input responses: By Shephard's lemma, differentiating $c_h^*(r, w; \delta)$ with respect to factor prices gives the unit input demands $a_{Kh}^* = \partial c_h^*/\partial r$ and $a_{Sh}^* = \partial c_h^*/\partial w$. Differentiating again and using the two-factor elasticity of substitution, the hat-form unit-input responses to factor-price changes and the AI shock are:

$$\hat{a}_{Kh}^* = -\,\theta_{Sw}^*\,\sigma_h^*\,(\hat{r} - \hat{w}) + \beta_K^*\,\Phi, \tag{A1}$$

$$\hat{a}_{Sh}^* = +\,\theta_{Kh}^*\,\sigma_h^*\,(\hat{r} - \hat{w}) + \beta_S^*\,\Phi, \tag{A2}$$

where β_K^* and β_S^* are the technology-bias parameters of the AI shock on capital and skilled-labor unit demands, respectively. Precisely, each is the proportional shift in the corresponding unit demand at constant factor prices, normalized by Φ:

$$\begin{aligned} \beta_K^* &\equiv \frac{\hat{a}_{Kh}^*\big|_{\hat{r}=\hat{w}=0}}{\Phi} = \frac{1}{\Phi}\frac{\partial \ln a_{Kh}^*}{\partial \ln \delta}\,\hat{\delta} = \frac{\delta}{a_{Kh}^*\,\Phi}\frac{\partial^2 c_h^*}{\partial r\,\partial \delta}\,\hat{\delta}, \\ \beta_S^* &\equiv \frac{\hat{a}_{Sh}^*\big|_{\hat{r}=\hat{w}=0}}{\Phi} = \frac{1}{\Phi}\frac{\partial \ln a_{Sh}^*}{\partial \ln \delta}\,\hat{\delta} = \frac{\delta}{a_{Sh}^*\,\Phi}\frac{\partial^2 c_h^*}{\partial w\,\partial \delta}\,\hat{\delta}, \end{aligned} \tag{A3}$$

the last equalities using $a_{Kh}^* = \partial c_h^*/\partial r$ and $a_{Sh}^* = \partial c_h^*/\partial w$. Setting $\hat{r} = \hat{w} = 0$ in the responses above returns $\hat{a}_{Kh}^* = \beta_K^*\Phi$ and $\hat{a}_{Sh}^* = \beta_S^*\Phi$, as required.

The two parameters are not independent. Since $c_h^* = r\,a_{Kh}^* + w\,a_{Sh}^*$, differentiating at constant factor prices gives:

$$\hat{c}_h^*\big|_{\hat{r}=\hat{w}=0} = \theta_{Kh}^*\,\hat{a}_{Kh}^*\big|_0 + \theta_{Sw}^*\,\hat{a}_{Sh}^*\big|_0.$$

The left-hand side is the complex's total proportional cost reduction from AI progress, which equals $(1 - \theta_{Kh})\Phi$: Φ is the cost diminution of the *task composite* (eq. (16)), and it enters good h's cost only through the task composite, whose share in h is $1 - \theta_{Kh}$ (eq. (13)). Hence the cost-share-weighted change in unit inputs satisfies:

$$\theta_{Kh}^*\hat{a}_{Kh}^* + \theta_{Sw}^*\hat{a}_{Sh}^* = (1 - \theta_{Kh})\,\Phi, \tag{A4}$$

which, substituting (A3) and using $\theta^*_{Kh} + \theta^*_{Sw} = 1$, gives the adding-up (Jones) constraint:

$$\theta^*_{Kh}\,\beta^*_K + \theta^*_{Sw}\,\beta^*_S = 1 - \theta_{Kh}. \tag{A5}$$

The two cost shares here are distinct objects: θ^*_{Kh} is capital's *effective* share in the complex, while θ_{Kh} on the right is capital's share in good h against the task composite, and the right-hand side $1 - \theta_{Kh}$ reflects the normalization of Φ to the task-composite cost.

Equation (A5) sets the cost-share-weighted *average* of (β^*_K, β^*_S) but not their relative magnitudes; the latter are determined by the underlying three-block structure (AI production technology, the marginal-task condition, and the relative intensities of K and S in AI production versus in direct task performance), which we do not unpack further here. The substantive economic content is encoded in the difference $b \equiv \beta^*_S - \beta^*_K$: with $\Phi < 0$, a larger β^*_K or β^*_S is a larger proportional reduction in that factor's unit demand, so $b > 0$ (equivalently $\beta^*_K - \beta^*_S < 0$) means the AI shock disproportionately reduces unit skilled-labor demand relative to unit capital demand–the skilled-labor-saving displacement direction maintained throughout the paper.

For services, with cost shares θ_{Ks} and $1 - \theta_{Ks}$ and elasticity σ_s:

$$\hat{a}_{Ks} = -\,(1 - \theta_{Ks})\,\sigma_s\,(\hat{r} - \hat{w}^U), \tag{A6}$$

$$\hat{a}_{Ls} = +\,\theta_{Ks}\,\sigma_s\,(\hat{r} - \hat{w}^U), \tag{A7}$$

with no technology shock.

Output and capital response; Full employment of the specific factors at fixed endowments gives $\hat{S} = 0$ and $\hat{L} = 0$. From $\hat{S} = \hat{a}^*_{Sh} + \hat{Y}^*_h = 0$ and $\hat{L} = \hat{a}_{Ls} + \hat{Y}_s = 0$:

$$\hat{Y}^*_h = -\hat{a}^*_{Sh} = -\theta^*_{Kh}\,\sigma^*_h\,(\hat{r} - \hat{w}) - \beta^*_S\,\Phi. \tag{A8}$$

$$\hat{Y}_s = -\hat{a}_{Ls} = -\theta_{Ks}\,\sigma_s\,(\hat{r} - \hat{w}^U). \tag{A9}$$

Combining unit-input responses with output changes gives the total capital demand changes in each sector:

$$\hat{K}^*_h = \hat{a}^*_{Kh} + \hat{Y}^*_h = -\sigma^*_h\,(\hat{r} - \hat{w}) + (\beta^*_K - \beta^*_S)\,\Phi, \tag{A10}$$

$$\hat{K}_s = \hat{a}_{Ks} + \hat{Y}_s = -\sigma_s\,(\hat{r} - \hat{w}^U). \tag{A11}$$

The Φ term cancels in $\hat{K}^*_h$ except through the bias difference $\beta^*_K - \beta^*_S$, which is the only structural way the shock enters capital demand at given factor prices.

With $\hat{\bar{K}} = 0$, full employment of capital requires:

$$\lambda^*_{Kh}\,\hat{K}^*_h + \lambda_{Ks}\,\hat{K}_s = 0. \tag{A12}$$

Substituting the expressions above:

$$\lambda^*_{Kh}\big[-\sigma^*_h(\hat{r} - \hat{w}) + (\beta^*_K - \beta^*_S)\Phi\big] - \lambda_{Ks}\,\sigma_s\,(\hat{r} - \hat{w}^U) = 0. \tag{A13}$$

Using (18) to eliminate $\hat{w}^U$, $\hat{r} - \hat{w}^U = \hat{r}/(1-\theta_{Ks})$, and collecting terms in $\hat{r}$ and $\hat{w}$ yields:

$$\Omega_r\,\hat{r} + \Omega_w\,\hat{w} = \Omega_\delta, \tag{A14}$$

which is (19), where the composite coefficients are:

$$\Omega_r = -\lambda^*_{Kh}\,\sigma^*_h - \frac{\lambda_{Ks}\,\sigma_s}{1-\theta_{Ks}}, \quad \Omega_w = \lambda^*_{Kh}\,\sigma^*_h, \quad \Omega_\delta = -\lambda^*_{Kh}\,(\beta^*_K - \beta^*_S)\,\Phi. \tag{A15}$$

Since $\lambda^*_{Kh}, \lambda_{Ks}, \sigma^*_h, \sigma_s > 0$ and $\theta_{Ks} \in (0,1)$, we have unconditionally $\Omega_r < 0$ and $\Omega_w > 0$. The sign of Ω_δ depends on the bias direction. Under the displacement direction $\beta^*_K < \beta^*_S$–the AI shock saves on unit skilled-labor demand by more than on unit capital demand–we have $\beta^*_K - \beta^*_S < 0$ and, together with $\Phi < 0$, $\Omega_\delta < 0$. This is the direction maintained in Assumption 2 and underlying the de-skilling result of Proposition 3. It is what allows $\hat{w}$ to fall in equilibrium even as the locus (17) has a positive right-hand side.

Solution: Stacking (17) and the quantity equation, in matrix form, gives:

$$\begin{pmatrix} \theta^*_{Kh} & \theta^*_{Sw} \\ \Omega_r & \Omega_w \end{pmatrix} \begin{pmatrix} \hat{r} \\ \hat{w} \end{pmatrix} = \begin{pmatrix} -(1-\theta_{Kh})\,\Phi \\ \Omega_\delta \end{pmatrix}. \tag{A16}$$

Solving this system yields:

$$\hat{r} = \frac{\Omega_w\,(1-\theta_{Kh})\,\Phi + \theta^*_{Sw}\,\Omega_\delta}{\theta^*_{Sw}\,\Omega_r - \theta^*_{Kh}\,\Omega_w}, \tag{A17}$$

$$\hat{w} = \frac{-\theta^*_{Kh}\,\Omega_\delta - \Omega_r\,(1-\theta_{Kh})\,\Phi}{\theta^*_{Sw}\,\Omega_r - \theta^*_{Kh}\,\Omega_w}. \tag{A18}$$

The denominator $\theta^*_{Sw}\,\Omega_r - \theta^*_{Kh}\,\Omega_w$ is negative unconditionally (both terms are negative), so $\hat{r}$ carries the sign of the negation of the numerator. Substituting $\Omega_w = \lambda^*_{Kh}\sigma^*_h$ and $\Omega_\delta = \lambda^*_{Kh}\,b\,\Phi$, with $b \equiv \beta^*_S - \beta^*_K$ the factor bias of AI progress, the numerator factors as:

$$\Omega_w\,(1-\theta_{Kh})\,\Phi + \theta^*_{Sw}\,\Omega_\delta = \lambda^*_{Kh}\,\Phi\big[\sigma^*_h(1-\theta_{Kh}) + \theta^*_{Sw}\,b\big],$$

so, since $\lambda^*_{Kh} > 0$ and $\Phi < 0$,

$$\hat{r} > 0 \quad\Longleftrightarrow\quad \sigma^*_h(1-\theta_{Kh}) + \theta^*_{Sw}\,b > 0 \quad\Longleftrightarrow\quad b > -\frac{\sigma^*_h(1-\theta_{Kh})}{\theta^*_{Sw}}.$$

The intensity ranking of Assumption 1 does not enter this condition; the sign of $\hat{r}$ is governed by the direction of the AI bias b. In this model AI progress saves skilled labor by more than capital ($b > 0$): at constant factor prices it displaces skilled labor from the top tasks ($\hat{\bar{\omega}}\big|_{\hat{r}=\hat{w}=0} < 0$; see the discussion following (21)). The bracket is then a sum of two positive terms, so $\hat{r} > 0$ follows directly–without the substitution-dominance condition. The rental stays positive for every $b \geq 0$ and

would turn negative only under a *strongly* capital-saving bias, $b < -\sigma_h^*(1-\theta_{Kh})/\theta_{Sw}^*$, which the task structure excludes. The sign of $\hat{w}$, by contrast, is genuinely ambiguous–it turns on the relative magnitudes of the technology effect Φ and the bias term Ω_δ–and the substitution-dominance condition (Assumption 2, derived next) is what ensures $\hat{w} < 0$.

The substitution-dominance condition: Recall $b \equiv \beta_S^* - \beta_K^*$, so $\Omega_\delta = \lambda_{Kh}^* b\,\Phi$ by (A15), and let $\det \mathbf{M} = \theta_{Kh}^*\Omega_w - \theta_{Sw}^*\Omega_r > 0$ (every term positive). From (A18), we have:

$$\hat{w} = \frac{1}{\det \mathbf{M}}\Big[\theta_{Kh}^*\Omega_\delta - \Omega_r(1-\theta_{Kh})\Phi\Big] = \frac{\Phi}{\det \mathbf{M}}\Big[\theta_{Kh}^*\lambda_{Kh}^* b - (1-\theta_{Kh})\big(\lambda_{Kh}^*\sigma_h^* + \tfrac{\lambda_{Ks}\sigma_s}{1-\theta_{Ks}}\big)\Big], \quad \text{(A19)}$$

where the bracket uses $\Omega_r = -[\lambda_{Kh}^*\sigma_h^* + \lambda_{Ks}\sigma_s/(1-\theta_{Ks})]$. Since $\Phi < 0$ and $\det \mathbf{M} > 0$, $\hat{w} < 0 \iff$ the bracket is positive, which is exactly (22). The condition compares the skilled-labor-saving bias b (in the first term) against the within-sector substitution elasticities σ_h^*, σ_s (in the second term). De-skilling requires the task-margin bias (which itself increasing in σ_{SA}) to dominate the factor substitutability of the high-tech complex and of services.

For the skill *premium*, $\hat{r}$ from (A17) is:

$$\hat{r} = \frac{\Omega_w(1-\theta_{Kh})|\Phi| + \theta_{Sw}^*\lambda_{Kh}^* b\,|\Phi|}{\det \mathbf{M}} = \frac{\lambda_{Kh}^*|\Phi|\big[(1-\theta_{Kh})\sigma_h^* + \theta_{Sw}^* b\big]}{\det \mathbf{M}}, \quad \text{(A20)}$$

linear and increasing in b with slope $\theta_{Sw}^*\lambda_{Kh}^*|\Phi|/\det \mathbf{M}$ independent of b (note $\det \mathbf{M}$ does not involve b). Premium compression requires $\hat{r} > \hat{r}_\pi$ with $\hat{r}_\pi$ as in (23); since $\hat{r}_\pi > (1-\theta_{Kh})|\Phi|/\theta_{Kh}^*$, the value of the wage threshold (22), the premium threshold is binding and implies $\hat{w} < 0$. Both thresholds are met for b large enough, i.e. for σ_{SA} high relative to σ_h^* and σ_s.

References

Acemoglu, D., and Restrepo, P. (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, 108(6), 1488–1542.

Acemoglu, D., and Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, 33(2), 3–30.

Dornbusch, R., Fischer, S., and Samuelson, P.A. (1977). Comparative advantage, trade, and payments in a Ricardian model with a continuum of goods. *American Economic Review*, 67(5), 823–839.

Gilbert, J., Koska, O.A., and Oladi, R. (2026). Tariffs, taxes and wages in the age of artificial intelligence. *Economic Inquiry* 64(2): 395–408.

Gordon, I.R., and McCann, P. (2000). Industrial clusters: Complexes, agglomeration and/or social networks? *Urban Studies*, 37(3), 513–532.

Isard, W., and Vietorisz, T. (1955). Industrial complex analysis and regional development, with particular reference to Puerto Rico. *Papers and Proceedings of the Regional Science Association*, 1, 229–247.

Jones, R.W. (1971). A three-factor model in theory, trade, and history. In J.N. Bhagwati et al. (Eds.), *Trade, Balance of Payments, and Growth*. North-Holland, Amsterdam.

Neary, J.P. (1978). Short-run capital specificity and the pure theory of international trade. *Economic Journal*, 88(351), 488–510.

Roy, A.D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2), 135–146.

Samuelson, P.A. (1947). *Foundations of Economic Analysis*. Harvard University Press, Cambridge, MA.

Silberberg, E. (1971). The Le Châtelier principle as a corollary to a generalized envelope theorem. *Journal of Economic Theory*, 3(2), 146–155.