Tax the Doing, Subsidize the Telling: Optimal AI Policy and the Codifiability of Skills*

Zhongchen Fan† Ruichi Xiong‡

July 17, 2026

Abstract

Artificial intelligence displaces the doing through which novices learn and teaches what can be told—opposing forces on formation. In an overlapping-generations model of expert–novice transmission, occupations indexed by expressibility, AI erodes tacit stocks and augments codifiable ones; the Pigouvian correction reverses sign once: tax the doing, subsidize the telling. A uniform AI tax is qualitatively wrong, not merely suboptimal. Calibrated to 682 U.S. occupations, the shadow schedule signs taxes on an eighth of employment, subsidies on the codifiable majority; the heaviest tax-side values fall not on routine work but on tacit, exposed work AI does but cannot teach.

Keywords: artificial intelligence, human capital, tacit knowledge, intergenerational externality, Pigouvian taxation

JEL Classification: J24, O33, H23, I20, J21

1 Introduction

An expert is made by doing the work, not by being told about it. A novice is formed by the same doing through which he contributes, so contribution and formation are a single act. For the tacit part of a skill there is no other route: it is acquired only by working beside someone who already has it.

*The authors contributed equally and are listed alphabetically by last name. We thank seminar participants for helpful comments. The Felten et al. (2021) AIOE dataset is publicly available at https://github.com/AIOE-Data/AIOE. Replication code accompanying this paper is available from the authors upon request and will be released publicly upon publication. All errors are our own.

†Xi'an Jiaotong University. Email: fanzc2022@xjtu.edu.cn.

‡Wuhan University. Email: ruichixiong@gmail.com.

Artificial intelligence acts on this formation in two opposing ways. First, it can do the work in the novice's place, delivering the contribution while forming no one—the displacing force that every prior automation also carried. Second, and unlike any prior automation, wherever the work can be put into words, it can supply the telling alongside the doing that remains. Through each hour of joint work, it teaches the novice what he would otherwise have learned only by doing. AI thus both thins the joint work and enriches what the remaining hours teach; which force prevails turns on how far the work can be codified.

Both forces are already visible in the field. Given a generative-AI assistant, customer-service novices raised their productivity by about a third while the most experienced agents barely moved (Brynjolfsson et al., 2025)—the telling force, in work that is largely tellable. The displacing force shows itself where the work is not. Robot-assisted surgery concentrates the operating in the senior surgeon, and residents reach independence having done far less of the work that forms them (Beane, 2019). In the labor market at large, employment in the occupations most exposed to generative AI has fallen for entry-level workers while holding up for experienced ones (Brynjolfsson et al., 2025). The doing disappears first at the rung where doing is how one learns. These observations sit on opposite sides of an old divide: where work can be transmitted by prescription, AI can both do it and teach it; where it cannot, AI can do the work but not teach it.

That divide is the cut Polanyi (1966) made between what we can tell and what we cannot, operationalized by Autor (2014) as the structural boundary of automation. Knowledge on its tacit side passes only through extended joint work with an expert, and the institutional response long predates artificial intelligence. Medical residencies, law-firm associate pyramids, doctoral advising, and the surviving craft apprenticeship all place a novice beside experienced practitioners for years. He is paid below what his developed skills will eventually earn, in exchange for expertise no other channel carries (Section 2).

The arrangements are a partial response to a long-standing externality: knowledge that forms only by doing carries value down a chain of future practitioners. But whoever directs the work is rewarded only over the horizon she is party to—a match, a contract, a career. The incomplete contract on the non-verifiable tacit channel cannot close the gap.[1] AI does not create this externality. It changes the stock on which the externality operates, and in both directions: it deepens the shortfall where it substitutes for doing it cannot teach back, and narrows it where the knowledge is codifiable enough for AI to teach.

The decisive terrain is a corner of the occupation space that no earlier technology could

[1] Emanuel et al. (2023) document the channel's fragility directly: software engineers seated in person with senior colleagues receive substantially more feedback than remote engineers, with the gap concentrated on junior engineers' learning. The transmission of tacit expertise survives only where expert–novice co-production survives.

open. Until now, whatever a machine could do, someone had first specified; what could not be told stayed beyond machines' reach. AI ends that identity. It learns the work from demonstration, as a novice would, and so performs what no one—itself included—can put into words: it escapes Polanyi's paradox, but only for itself. That AI reaches the cognitive work automation never touched is by now consistently documented across the exposure literature (Brynjolfsson et al., 2018; Webb, 2020; Eloundou et al., 2024); the dark side of that capability is our subject. Where AI does work it cannot teach back, it severs the channel through which the skill reproduces, and it re-supplies nothing. The corner is populated: a tenth of occupations and nine percent of U.S. employment sit in it, schoolteachers its largest member, and its depths carry the fastest erosion and the correction's deepest tax-side values. There the policy alternatives also run out: at the tacit floor the telling cannot be built, and only the doing can be protected.

We study the mechanism in its sharpest instance: a two-period overlapping-generations (OLG) economy in which the appropriation horizon is exactly one generation. Each cohort is a novice first, acquiring a single body of occupational knowledge by co-producing with an expert, then the expert who produces output and forms the next cohort's novice. Occupations are indexed by *expressibility*, a scalar measuring how far their knowledge can be codified; explicit and tacit are the two poles of the cross-occupation spectrum.

AI enters twice. In production it does the work, combining with human capital as a gross substitute; in transmission it augments formation with a reach that rises with expressibility and vanishes at the tacit floor, where what cannot be told cannot be taught. The appropriation extends Becker (1962) to overlapping generations: the expert captures her novice's contribution within the period they share but not the value his formation carries down the chain. The intergenerational wedge is therefore positive at every AI level, including none. AI moves the stock the wedge prices. It compresses the co-production that forms novices, in codifiable and tacit work alike, and it re-supplies the formation only where the knowledge is codifiable. We characterize the AI-marginal slice of the correction in closed form and show it is a continuous schedule across the expressibility spectrum, the proper target of any AI-pricing instrument.

The first result is a qualitative shift in the long-run distribution of human capital. AI deployment erodes the steady-state knowledge stock in tacit-type occupations and augments it in codifiable ones, with a single interior crossover in expressibility (Proposition 5). The sorting is anchored at the tacit pole by the Polanyi asymmetry: AI reaches into codifiable transmission but not into tacit demonstration. Across the spectrum, expressibility orders the contest between the formation AI re-supplies and the learning by doing it displaces.

Aggregate output keeps rising with AI under the baseline gross-substitutes calibration

even where the stock erodes. Over most of the AI range the erosion is silent, flagged by no output statistic. Output turns down only in the deepest tacit-and-exposed occupations, where the damage is already largest. The human-capital base thus shifts toward the explicit pole beneath a growing aggregate.

The second result is the shape of the optimal correction. The Pigouvian correction on AI is a continuous schedule in the occupation's expressibility: the discounted shadow value of the knowledge stock times the marginal damage AI does to it (Proposition 3). Because the shadow value is positive, the schedule carries the sign of the damage—positive, the tax direction, exactly where AI erodes the stock, and negative, the subsidy direction, where AI augments it: tax the doing, subsidize the telling. The schedule crosses zero once, at a *force-free* cut that balances AI's reach into transmission against the compression-weighted learning by doing it displaces (Proposition 6). The single crossing is a theorem of the calibrated class at the decentralized steady state and verified at the corrected one. Neither the appropriation nor the shadow value enters the cut condition—they scale the schedule's magnitude—so the sign turns on the AI-augmentation channel alone.

Within the baseline the schedule is a shadow object, a damage index; the planner's allocation is decentralized by the co-production subsidy, built from the same shadow value. Once experts also choose how much AI to use, the levied second-best price shares the schedule's sign structure at a force-free cut of its own. The first-best assignment leaves AI untaxed altogether and corrects the wedge on the co-production margin instead (Section 4). A uniform AI tax is therefore wrong in kind rather than degree: the index takes both signs across the spectrum, while any uniform instrument applies one.

The two results gain quantitative content from a calibration to the full cross-section of 682 U.S. occupations, with expressibility proxied by an O*NET task-content index, exposure by the AI Occupational Exposure index of Felten et al. (2021), and BLS employment as weights. The reversal lands at an empirically relevant place: the schedule crosses zero near the tacit pole. About an eighth of employment, and a similar share of occupations, sit on the tax side of the cut while the codifiable majority sit on the subsidy side. The employment-weighted index is negative, a net subsidy in direction, even as individual occupations range widely on both sides of zero. The index runs most positive not in routine work but in tacit, exposed occupations—under the baseline task-content proxy, counselors and clergy.

That is because exposure and expressibility are nearly independent in the data: how exposed an occupation is to AI is not aligned with how codifiable its work is (Brynjolfsson et al., 2018; Webb, 2020; Eloundou et al., 2024). Exposure sets how hard AI presses on an occupation; expressibility decides what the pressing does to formation. Exposure alone, the usual summary of whom AI affects, cannot say where correction is needed.

The magnitudes carry their ranges: the tax-side share runs from under one percent to about a third of employment across the sweep of the augmentation schedule's unidentified curvature, the widest of these ranges. The reversal and the cut survive every sweep (Online Appendix H). The theory is the structural content, and the calibration shows where its objects fall: a demonstration of the mechanism's reach rather than an estimate of its primitives.

The calibration also prices what the reversal implies for policy. The subsidy region is the creation of AI's re-supply channel alone. Replayed under automation—the polar case that does the work but cannot teach—the index turns positive in every one of the 682 occupations. AI's capacity to teach converts seven-eighths of employment to the subsidy side.

AI's own contribution to the correction is about a fifth of a large stake: the planner improves on the decentralized equilibrium by two-thirds of discounted welfare, employment-weighted. About four-fifths of that gap predates AI—AI sharpens an old externality rather than creating a new one. The AI rate is throughout a second-best instrument: even optimally targeted it recovers just under a tenth of the first-best gain, while the co-production subsidy recovers all of it.

The rate's case is also long-run. Evaluated through the transition at the calibrated generational discount, steady-state-optimal AI rates cost welfare—adjustment costs arrive a generation before formation gains. Rates re-solved for the transition keep the sign geography at half the size, recovering a small positive gain, while the co-production subsidy keeps a first-order one. On either horizon, what the rate must get right is the sign.

In an adoption-margin exercise with prohibition in the policy space, the optimum is a partition. Finite prices drive the taxed occupations to exactly zero use: the allocation a ban would pick, priced at the smallest rates that support it. They push the codifiable majority to the adoption frontier, the signs reversing across the spectrum. Policies blind to expressibility forfeit most of the gain. The automation model's rates, free to discriminate by occupation but seeing only erosion, recover essentially none of it; the best uniform rate, the subsidy the majority dictates, recovers three-fifths; ten brackets keyed to AI exposure do scarcely better, because each exposure bracket spans nearly the economy-wide range of expressibility. Two brackets on expressibility recover the full targeted gain.

Relation to Ide (2026).— Closely related contemporaneous work by Ide also studies how new technology bears on the intergenerational transmission of tacit knowledge, in an OLG economy with incomplete contracts built on the same Polanyi foundation. The mechanisms differ in what the technology may do to transmission. In Ide's economy transmission is human-to-human: automation reduces the span of control of the most-skilled experts and reallocates novices toward less-skilled mentors, and growth can slow through this matching

deterioration. The technology acts on who mentors whom, not on what a machine can teach.

In ours, occupations differ in expressibility, and where knowledge is codifiable AI re-supplies part of the formation as codified instruction. This teaching channel, outside Ide's mechanism, turns AI from an eroding force into an augmenting one and delivers the sign-reversing schedule. Each framework addresses a margin the other abstracts from, his the matching of novices across heterogeneous experts, ours the codifiability of what passes between them. The erosion side also has a commons-level counterpart in Acemoglu et al. (2026), where agentic AI substitutes for human learning and the shared knowledge stock can deplete; that force is our compression, read economy-wide, and the codifiability axis with its sign-reversing re-supply channel is what this paper adds.

Relation to the AI–automation literature.— The task-allocation framework of Acemoglu and Restrepo (2018, 2020, 2022) characterizes the reallocation of tasks between labor and capital as automation proceeds, and it reads codifiability as one thing: replaceability. Work that can be specified is work a machine can perform, and where human capital forms by performing it, displacement erodes the stock—an erosion real but not specific to AI. Codifiability has a second edge that only a language machine turns: know-how that can be written down can be written into a training corpus and comes back as instruction. The erosion channel is thus the AI-generic half of our mechanism; the re-supply channel, which reverses the policy sign, is the AI-specific half. That framework describes the within-generation wage and employment incidence; ours identifies the between-generation consequences for transmission, where the structural externality lives.

Relation to the taxation of technology.— The robot-taxation literature prices technology from the distribution side, and its optimal rates are modest: transitional and vanishing in steady state in Guerreiro et al. (2022), a subsidy while robots are expensive turning to a tax as they cheapen in Thuemmel (2023), disciplined by the sufficient statistics of Costinot and Werning (2023). Korinek and Stiglitz (2019) carry the distributional case to AI itself, and Acemoglu et al. (2020) document a U.S. tax code that, if anything, already favors automation. Our schedule answers a corrective question instead. AI moves the formation of the knowledge stock—an externality that persists under complete instruments and carries a sign set by codifiability rather than a size set by redistributive weights. The two motives identify distinct externalities and can compound.

Recent analyses also find AI corrections that flip sign, but on other margins: over time on a redistribution margin in Growiec et al. (2026), across uses within a representative occupation in Afrouzi et al. (2026). Ours reverses across occupations: the same index signs a tax in tacit occupations and a subsidy in codifiable ones in the same period, a pattern no economy-wide or use-based rate can track. The instrument-assignment lesson survives the

change of motive: where a better-targeted margin can be reached—their income tax, our co-production subsidy—the technology itself goes untaxed (Section 4).

Relation to human-capital theory and the tacit-knowledge tradition.— The appropriation technology extends Becker (1962): firms will not pay for general training workers can walk away with, so trainees buy it with reduced wages. Labor-market imperfections restore part of the firm's incentive to pay (Acemoglu and Pischke, 1998). In contemporaneous work, Garicano and Rayo (2026) ask how AI bears on the viability of the training relationship itself—the career contract through which novices finance their formation. Our friction is complementary and operates however that contract is financed, since no career reaches past the generation it spans. Here the general–specific line maps onto expressibility: what can be told can be learned outside the match.

Jarosch et al. (2021) quantify the within-match counterpart, learning from coworkers partly priced into wages; the recursive tail across generations is the part no wage inside the match can price, and it is what the wedge measures. Ma et al. (2026) locate that learning in early careers—the novice phase the co-production margin here prices. The skill-investment scaffolding is closest to Ben-Porath (1967); relative to the within-life skill formation of Cunha and Heckman (2007); Cunha et al. (2010), our setting is between lives, with the explicit–tacit asymmetry in place of the cognitive–non-cognitive one.

The idea that tacit knowledge bounds automation traces to Polanyi (1966) and, for modern labor markets, to Autor (2014); Nelson and Winter (1982) root evolutionary growth theory in the persistence of tacit organizational routines. Jovanovic and Nyarko (1996) formalize how adopting a better technology can idle the expertise built on the old; in our economy the loss runs through training rather than obsolescence—the stock erodes not because the knowledge stops mattering but because its reproduction stops happening. The labor-market premium on social skills (Deming, 2017) prices the capacities at our tacit pole. Autor and Thompson (2025) rank occupations by the expertise their tasks demand; expressibility is a distinct axis—how codifiable the knowledge is, not how much of it a task requires. A skill can be highly expert yet largely tacit (surgery) or highly expert and largely codifiable (actuarial work). Against this background the paper contributes three things: the first formal OLG model in which expressibility indexes occupations and shapes intergenerational human-capital dynamics; the re-supply channel through which AI, unlike any prior automation, returns the codifiable portion of formation; and the optimal AI correction characterized as a continuous, sign-reversing schedule.

The remainder of the paper proceeds as follows. Section 2 develops the conceptual framework. Section 3 sets up the model and solves its decentralized equilibrium. Section 4 turns to the social planner: the wedge, the damage index, and instrument assignment.

Section 5 establishes the sign reversal at the force-free cut. Section 6 calibrates the model to the U.S. occupational cross-section. Section 7 concludes; the Online Appendix collects the proofs, the robustness analyses, and the data construction.

2 Conceptual Framework

Expertise spans a spectrum of expressibility. The foundational claim of Polanyi (1966)—"we can know more than we can tell"—fixes the spectrum's tacit end: knowledge there resists articulation and passes only through sustained joint work with those who command it. At the explicit end, knowledge travels through manuals, lectures, and code. Most expertise mixes the two, and where an occupation's knowledge sits on this axis is the single primitive our model turns on. Polanyi's own phrasing anticipates the mechanism (Polanyi, 1958, p. 53):

> An art which cannot be specified in detail cannot be transmitted by prescription, since no prescription for it exists. It can be passed on only by example from master to apprentice. … By watching the master and emulating his efforts in the presence of his example, the apprentice unconsciously picks up the rules of the art, including those which are not explicitly known to the master himself.

The prescription and the example are the two transmission technologies of this paper: the telling, which travels without the teller, and the doing, which requires expert and novice side by side.

Where knowledge passes only by example, formation must happen inside production. The novice learns the occupation's real work by doing it under an expert's direction, so his doing is at once contribution (output realized today) and formation (the skill he carries forward): one act, not two (Arrow, 1962). The institutional response is uniform across professions whose explicit content could hardly differ more: the medical residency, the surgical fellowship, the law-firm associate-to-partner pyramid, the academic advisor–student relationship, the engineering rotation in technology firms, and the surviving craft apprenticeship. All impose a multi-year period during which a novice works alongside experienced practitioners at a wage below the eventual marginal product of his developed skills. In exchange he receives expertise that cannot be acquired through any other channel.[2]

Nelson and Winter (1982) build evolutionary growth theory around tacit organizational routines that survive through this mechanism. The form predates artificial intelligence by

[2]The length of the below-marginal-product period varies considerably: an investment-banking analyst's track typically runs three years, a surgical residency with subspecialty fellowship seven, a doctoral advisor–student relationship in economics five to six. The contractual logic is the same in each: deferred compensation, partial appropriation by the expert, and structural barriers to early exit.

centuries: de la Croix et al. (2018) document that European apprenticeship institutions contributed measurably to technological creativity and per-capita growth before the Industrial Revolution. Our model abstracts from the specifics of any one institution to the structural feature all of them share: a contract-driven, two-period co-production between an expert and a novice, in which the novice's contribution and his formation are welded.

These arrangements only partially correct a long-standing externality. Becker (1962) shows that firms will not finance general training because trained workers can leave and take it with them, so the trainee finances his own. The transmission studied here carries an analogous incompleteness across generations rather than across employers, and a more severe one. Tacit human capital is transmitted recursively: today's novice becomes tomorrow's expert, whose own stock is the input to training the next generation's novice, and so on without end. A unit of the expert's training effort sets off a chain of social benefits that propagates across all subsequent generations, with discounted value $1/(1-\delta\gamma) > 1$. Here γ is the persistence of human capital across generations and δ the planner's intergenerational discount factor (both formalized in Section 3). The expert's incentive reaches only her novice's contribution during the period they share; the recursive tail accrues to generations with whom she never coexists, and she captures none of it.

That the tail is uncontractible to some degree is all the wedge requires, and the generational horizon already supplies it: the later generations who reap the tail arrive only after she is gone, so no contract can run to them. Tacitness only closes the remaining route—even a long-lived firm, guild, or partnership cannot condition on a non-articulable stock it can never verify. Institutions blunt the wedge: long training contracts, professional credentialing, accreditation, and deferred-compensation tracks in the spirit of Lazear (1979) fold part of the tail into the expert's contemporaneous reward. None closes it, and the residual gap is the object of the apprenticeship and professional-training literature (e.g., Acemoglu and Pischke, 1998; Wolter and Ryan, 2011).

We take this shortfall of the private below the social return as the model's central friction (Assumption 4). Our aim is to identify the channel and characterize the *shape* of its optimal correction: the cross-occupation sign pattern and its comparative statics. The calibrated magnitudes of Section 6 illustrate that shape.

Artificial intelligence is not the first general-purpose technology to move the boundary this friction sits on. The printing press, formal schooling, and the technical manual each widened the explicit side: they codified more of what practitioners knew and carried it more cheaply. Automation, when it came, pushed on the doing instead; it performed codified work and taught nothing. AI is the first to push on both sides at once.

On the doing side it displaces, as automation always has, but now on both sides of the

cut. Autor (2014) drew the automation boundary along Polanyi's distinction: machines substitute for the rules they are given, not for the situational judgment of an experienced practitioner. Yet even where an occupation's mature expertise is deeply tacit, the routine subtasks that fill a novice's early years are codifiable enough for AI to absorb. As it absorbs them, the joint work through which tacit judgment passed as a by-product disappears with them: the operating room where residents watch procedures they would once have performed (Beane, 2019).

On the telling side, and only where the knowledge itself is codifiable, AI restores. To codify a skill is to put its know-how into words; what can be put into words can be fed to a training corpus; and a model trained on that corpus returns it as instruction. So AI can carry part of the formation itself, teaching the novice through instruction what he would otherwise have acquired only by doing. The deployments in which novices gain most from AI assistance are this force at work, strongest where the task is most tellable (Brynjolfsson et al., 2025). Early survey evidence points the same way from the worker's side: U.S. workers report setting aside more time for career-relevant learning since generative AI arrived, with the increase concentrated among those who find AI a useful learning tool (Kim et al., 2026).

The restoring force vanishes at the tacit pole, where what forms the novice is exactly what cannot be told; there AI removes the doing without replacing the formation. The floor is definitional on the language channel: however capable the system, knowledge that cannot be put into words cannot be supplied as words. The results ask of the floor only an inequality (Section 5).

Whether AI can do an occupation's work and whether it can teach its knowledge are therefore different questions, and they need not share an answer. Under automation they coincided: a machine's capability *was* a codification (whatever it could do, someone had first specified). What could be codified was what a machine could perform, and work that resisted telling was safe from doing. AI separates the two because it acquires capability from demonstration rather than specification, and so can do what no one, the machine included, can articulate.

Polanyi's paradox is thereby displaced rather than resolved: defeated for production, it stands for transmission, escaped by AI only for itself. What the machine learned without telling, it cannot tell back. The asymmetry lies in the channel: what AI takes in need not pass through words, but what it can hand to a person must. Tacit skill moves between humans by shared doing rather than telling, so it survives the passage into AI but not the passage back out. Machine capability now reproduces by copying, while human capital still reproduces by co-production.

The two capacities are accordingly free to come apart across occupations, and in the data

they do: how exposed a job is to AI and how far its knowledge can be codified prove nearly independent (Section 6). A job can sit well within AI's reach to perform yet beyond its reach to teach, and it is there, exposed but tacit, that formation erodes without re-supply. In such jobs, AI performs a share of the work but cannot deliver the occupation's service without the expert. Displacing the doing that formed her thins the supply of a skill that remains essential—failed reproduction, not obsolescence.

Expressibility—not exposure, and not the routineness that organized the automation era's incidence—is therefore the axis on which AI policy must be signed. Exposure measures how much of an occupation's work AI can do, and so how hard AI presses on it; expressibility governs whether, having taken the work over, AI can restore the formation the doing would have built. In tacit occupations the displacing force runs unopposed and the pre-existing wedge widens; in codifiable ones the restoring force can match or outrun it, and the same wedge can narrow. The externality does not change in kind under AI; it is redrawn along the codifiable boundary, and the correction must be signed accordingly. The question the model answers is this balance: where on the axis erosion gives way to augmentation, and what correction each side calls for. Section 3 makes the two forces formal.

3 The Decentralized Equilibrium

We now formalize the model and solve its decentralized equilibrium. A single occupation passes its human capital down an overlapping-generations chain: in each period an expert produces, sourcing the work either from AI or from the novice who will succeed her, whose doing forms him. We first set out the environment and the learning by doing that forms the novice, then the role of AI, the technologies of the two routes the doing can take, and the expert's output and objective. We then solve the expert's problem and characterize the steady state. There AI's net effect on the knowledge stock emerges as a contest between its reach into transmission and the learning by doing it displaces.

3.1 Environment, Learning by Doing, and AI's Two Roles

Time is discrete and indexed by $k \in \{0, 1, 2, \ldots\}$. We use *period* and *generation* interchangeably for one step k—the interval over which a single expert forms a single novice, and the unit across which the intergenerational externality operates.[3] In each period k, two *cohorts* overlap. The *period-k expert* entered one period earlier as the novice of period $k-1$ and holds a scalar stock of human capital $h_k \geq 0$. The *novice*, formed during period k, becomes

[3] A period is a training-reproduction cycle, not a demographic generation.

the period-$(k+1)$ expert. Each cohort is thus economically active for two periods: as a novice in the first, then as an expert in the second.

Human capital needs a human to make it: the novice is formed by doing the occupation's real work under the expert's guidance. That work is at once his *contribution*—output realized within the period—and his *formation*—the human capital he carries into the next period as the expert he becomes. These are not two activities but one: he produces by becoming and becomes by producing (Arrow, 1962). The two are welded because the tacit part of the skill admits no route around the doing: it is acquired only in the work itself, alongside an expert who holds it (Section 2). This holds whether the work is a resident's surgery, an associate's brief, or an artisan's commission. It holds whether one reads its structure as a junior tier executing under a senior's judgment (Garicano, 2000) or simply as practice under supervision.

The same work can be done without the novice. AI performs the tasks, so the expert can source the doing from it instead. Novice and AI are substitutes in getting the work done but differ in what the doing leaves behind: when the novice does it, he is formed by it; when AI does it, the contribution is delivered and no one is formed.

The expert's one decision is therefore how to source the doing—a fraction $t_k \in [0,1]$ of her unit time endowment co-producing with the novice, the remainder $1-t_k$ producing with AI. Both yield the period's output; only the novice's share also forms the next expert, which is why the sourcing decision t_k governs the next generation's formation as a by-product. The novice makes no choice: he supplies his time to the match and takes the formation it delivers. At the end of period k the expert retires, and the novice graduates to the period-$(k+1)$ expert.

The expert works alongside an exogenous level of AI capability A. Its principal drivers—compute capacity, algorithmic progress, and public research funding—operate on time scales and through channels outside the human-capital decisions modeled here. The theory holds A fixed and characterizes the equilibrium and its comparative statics in the level of A.

AI bears on the work in two ways. First, it does the work, substituting for the doing: this is the channel automation already had, and on its own it severs the bundle—contribution without formation. Second, and unlike any prior automation, it can supply the formation itself: where the knowledge is codifiable, the novice can acquire it apart from the doing, from codified instruction with AI among its sources. Whether AI's substitution for the doing or its re-supply of the formation prevails—erosion or augmentation—is the contest of Section 3.4. The production and transmission technologies that follow formalize each channel in turn: the substitution for the doing in production, the re-supply of the formation in transmission.

3.2 Technologies and the Expert's Objective

The two routes the doing can take carry different technologies. On the share $1-t_k$ of her time the expert does the work with AI, producing $\Phi(h_k, A)$, her human capital and AI capability entering as joint inputs. AI needs no preparation: it arrives capable, and she deploys it directly, so this output turns on her own stock and the AI level alone. We leave Φ a *general* aggregator (Assumption 1).[4]

Assumption 1 (Production aggregator)**.** $\Phi(h, A)$ is strictly increasing and concave in h and increasing in A, with marginal products $\Phi_h \equiv \partial\Phi/\partial h > 0$ and $\Phi_A \equiv \partial\Phi/\partial A \geq 0$.

On the remaining share t_k the expert co-produces with the novice, who does the work under her direction and is thereby formed: what he carries into the next period is built here, through the transmission technology

$$h_{k+1} \;=\; H(h_k,\, t_k,\, A), \tag{1}$$

with A the AI level. This is where AI's second role takes formal shape: it enters transmission, not only production, supplying formation rather than merely doing the work. The channel operates through the match: transmission runs on co-production, and AI raises what each hour of it transmits—instruction alongside the doing, which is where the field evidence places the novice gains (Brynjolfsson et al., 2025). Online Appendix H quantifies the alternative, letting a share of the instruction bypass the match: the sign reversal of Section 5 and its single crossing survive, the cut moving modestly. As with production, we leave H a *general* technology, characterized by Assumption 2.

Assumption 2 (Transmission technology)**.** H is strictly increasing in h and t, strictly concave in t, and vanishes at a zero stock, $H(0, t, A) = 0$: nothing transmits without inheritance.[5] Its elasticity in the inherited stock, $\partial \ln H/\partial \ln h$, lies in $(0, 1)$ at every state, so transmission has diminishing returns in inheritance.

Co-production also yields current output, $T(h_{k+1})$, the novice-route counterpart of Φ—the same occupation work, distinguished only by the input that performs it, AI on the one

[4]More expert knowledge always raises output, whatever the level of AI. The stronger polar property is *strict essentiality*, $\Phi(0, A) = 0$: AI cannot deliver the occupation's service at all without human capital. Strict essentiality is needed only where AI can drive output down (Online Appendix Proposition G.1); the erosion and corrective-policy results below require only $\Phi_h > 0$. (An occupation whose stock fell toward zero would lose the capacity to deliver the service that defines it, regardless of available AI: surgery without surgeons, primary education without teachers, legal practice without judgment-experienced lawyers.)

[5]The zero-stock boundary does its work at the tacit pole, where formation has no route but the doing. At the explicit pole codified knowledge outlives its holder, and a vanished stock might be re-seeded from the corpus; the steady states studied below are interior, so the boundary is never engaged there.

share and the novice on the other. The novice, unlike AI, is not capable in advance: he can contribute only what his formation has brought him to. This route therefore runs on the stock h_{k+1} the co-production builds in him, whereas the AI route runs on the expert's own stock and a tool that arrives ready.

What this route produces accrues to her, appropriated within the period the two cohorts share. The *appropriation technology* T takes only the formed stock h_{k+1} as its argument: how much the expert can extract is settled within the shared period, independent of the payoffs the novice earns after the expert retires—the contracting boundary the overlapping-generations demography draws.[6] In the Becker (1962) reading the novice is paid in human capital: he surrenders the value he creates today for the expertise he carries into the next period as its expert.

Assumption 3 (Appropriation technology)**.** $T(h)$ is strictly increasing and concave in h, with the normalization $T(0) = 0$.

The expert's output sums the two ways of getting the work done: the share done with AI yields $(1 - t_k)\Phi(h_k, A)$, and the share co-produced with the novice yields the appropriated $T(h_{k+1})$, so occupation output is

$$Y_k \;=\; (1 - t_k)\,\Phi(h_k, A) \;+\; T(h_{k+1}). \tag{2}$$

The additive form embeds three assumptions worth stating.

First, the two routes' outputs are perfect substitutes in the period's output: a unit produced with AI and a unit appropriated from the novice are worth the same and simply sum. The routes therefore contend only for the expert's single unit of time, split t_k against $1 - t_k$.

Second, within that competition output is linear in the AI share and concave in co-production. An hour with AI returns the same $\Phi(h_k, A)$ as every other, a capable tool at constant return, so the AI term $(1 - t_k)\Phi$ is linear in t_k. An hour with the novice adds to $T(h_{k+1})$ at a diminishing rate (H concave in t_k) and, unlike the AI hour, builds the stock carried into the next period. This quasi-linearity lets Φ serve as the *price* of co-production: the output the expert forgoes on each hour spent with the novice rather than with AI. That

[6] We take the appropriation to depend on the novice's stock alone, $T(h_{k+1})$. A natural enrichment is immaterial to the results below: the expert's own stock could also enter the appropriation directly, $T(h_k, h_{k+1})$, as production does—the expert co-producing rather than only forming. But h_k is predetermined when she chooses t_k, so this touches neither the first-order condition nor the force-free threshold (Section 4.2), only the shadow value. Her stock already bears on the appropriated contribution through transmission, $h_{k+1} = H(h_k, \cdot)$. Proposition C.1 in Online Appendix C establishes this formally. A level wage paid to the novice inside the match is likewise immaterial: a transfer between the parties, it moves no margin for the expert or the planner (Online Appendix C).

price rises with A: as AI improves, co-production costs more. Where AI re-supplies nothing through transmission, the expert shifts the doing toward AI, draining the very activity through which tacit skill transmits.

Third, A does not enter the co-production output directly. AI bears on the period through exactly two channels: pure production, where it enters Φ, and transmission, where it enters H. The appropriated output T depends on A only through the formed stock h_{k+1}, never on its own. Holding any direct AI term out of T keeps these two roles separate and exhaustive, and it reflects that within the period the expert, not the novice, commands AI. The exclusion is also load-bearing for the force-free threshold of Section 4.2. An appropriation that rose with AI directly—AI-assisted monitoring of the novice's contribution, say—would place a term in the threshold condition itself, rather than only scaling the correction's magnitude (Online Appendix C).

The appropriation and the inheritance are one object: h_{k+1} is at once what the expert appropriates this period, through $T(h_{k+1})$, and what the next generation inherits as its stock—the bundle, now visibly the same h_{k+1}. Co-production yields both; AI yields only the contribution. Figure 1 draws the period in full: the expert's sourcing decision, the two routes the doing can take, and the chain that co-production alone carries to the next generation.

The period payoff Y_k of (2) values the novice's contribution but not the formation he carries beyond the shared period. Formation sets off a recursive chain: today's novice forms tomorrow's, and so on. The expert, rewarded only within the period she shares with her novice, does not capture the chain's discounted social value.

Assumption 4 (Private objective)**.** The decentralized expert maximizes the period payoff Y_k of (2). She does not internalize the value her transmitted human capital creates for subsequent generations, so the private return to transmission falls short of its social return.

We remain agnostic about the precise contractual or institutional source of this friction.[7] Its magnitude is what the planner's problem of Section 4 takes as given and corrects.

3.3 Expert's Optimal Co-Production Time

The environment is now specified: the general production aggregator Φ, transmission technology H, and appropriation technology T. We solve the model for a single occupation

[7]The two candidate microfoundations deliver the same reduced form: the finite contracting horizon of overlapping generations (the recursive tail accrues to later generations who cannot be party to the expert's contract), and the non-verifiability of the tacit stock transmitted. Partial corrections—deferred compensation, credentialing, accreditation—blunt the wedge without closing it. Online Appendix B shows that neither Beckerian altruism nor the novice financing his own formation closes it.

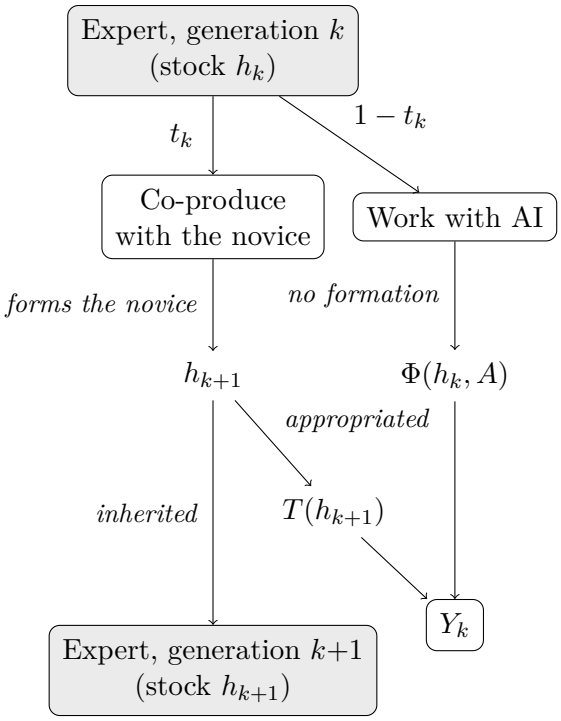


Figure 1: The Bundle and Disembodiment in a Single Period

Note: The generational chain runs down the center (shaded): the period-k expert, holding stock h_k, becomes—through the novice she forms—the period-$(k{+}1)$ expert. She splits her unit of time, a share t_k co-producing with the novice and the rest $1 - t_k$ working with AI. Co-production forms the novice's stock h_{k+1}, a single object with two roles: appropriated within the period as the expert's contribution $T(h_{k+1})$, and inherited as the next expert's stock. Working with AI instead delivers the contribution $\Phi(h_k, A)$ but forms no one—the doing disembodied from the formation that rides on it (Arrow, 1962). Period output sums the two contributions, $Y_k = (1 - t_k)\,\Phi(h_k, A) + T(h_{k+1})$; only co-production, on the central axis, carries the chain forward to the next generation.

and ask how AI moves the decentralized equilibrium. The expert's optimum fixes how she sources the doing; the steady state then determines AI's net effect on the knowledge stock.

The expert chooses how to source the doing—the fraction $t_k \in [0, 1]$ she co-produces with the novice rather than executing with AI—to maximize $Y_k = (1 - t_k)\Phi(h_k, A) + T(H(h_k, t_k, A))$ defined in (2). Write $T_h \equiv \partial T/\partial h$ for the marginal appropriation, evaluated at the novice's stock h_{k+1}, and $H_t \equiv \partial H/\partial t_k$ for the marginal product of co-production time. The marginal hour of co-production raises the appropriation by $T_h\, H_t$ and forgoes the output Φ she would have produced executing that hour with AI instead; balancing the two gives the optimum.[8]

[8]The primitives Φ, H, and T are twice continuously differentiable throughout. An interior optimum is *regular* when its second-order condition holds strictly, so the relevant Jacobian is nonsingular. The implicit function theorem then makes the policy $t^*(h_k; A)$ and, at the stable steady state (Lemma 2), the stock $h^*(A)$ continuously differentiable in A (proofs of Lemma 1 and Proposition 1). The derived steady-state objects (V', τ^*, μ, Y^*) inherit continuity as compositions. The value function's differentiability is taken up in Section 4.1.

Lemma 1 (Expert's optimal co-production time)**.** *The expert's interior optimum* $t^*(h_k; A)$ *solves*

$$T_h(h_{k+1})\, H_t(h_k, t^*, A) \;=\; \Phi(h_k, A), \qquad h_{k+1} = H(h_k, t^*, A), \tag{3}$$

unique because the appropriation is strictly concave in t_k *(Assumptions 2 and 3).*[9] *AI compresses co-production (*$\partial t^*/\partial A < 0$*, the* substitute regime*) exactly when it raises production faster than it raises the appropriated marginal return to co-production,*

$$\frac{\partial \ln \Phi}{\partial A} \;>\; \frac{\partial \ln(T_h H_t)}{\partial A};$$

the converse is the complement regime*.*

The substitute–complement boundary is the locus $\partial \ln \Phi/\partial A = \partial \ln(T_h H_t)/\partial A$, where $\partial t^*/\partial A = 0$, and two effects of AI oppose each other across it. On the production side, AI performs the execution the expert would otherwise source from the novice, raising the output she forgoes by co-producing. On the co-production side, its augmentation raises the appropriated return $T_h H_t$ to working with the novice. Co-production falls with AI when the first effect outweighs the second, and rises otherwise. In the substitute regime AI displaces the novice from the doing: the expert routes execution to the cheaper assistant and co-produces less.

The appropriated return $T_h H_t$ moves through two channels. AI's augmentation raises the marginal product of co-production H_t where it scales transmission, lifting the return. But its reach also raises the novice's stock, and because appropriation is concave, the larger stock moves the expert down the marginal-appropriation curve, weakly lowering T_h. This second channel is an additional compression force.

We write

$$\mu \equiv -\frac{d \ln t^*}{dA} \tag{4}$$

for the *co-production compression rate*: the proportional decline in equilibrium co-production per unit of AI, positive exactly where AI compresses co-production on net. Lemma 1 signs the response at fixed h_k; in the steady-state analysis below, μ is the total response along the fixed point, and its sign is an equilibrium property. By (3) and the chain rule, μ depends on (h_k, A) but is a single scalar at any given state.

[9] The optimum is interior under mild conditions. The upper corner $t^* = 1$ binds when co-production is worthwhile throughout. The lower corner $t^* = 0$ is ruled out wherever the first co-production hour pays, $\lim_{t\to 0} T_h H_t > \Phi$. This holds whenever that first hour's marginal product is unbounded—a co-production Inada condition $\lim_{t\to 0} H_t = \infty$ for $h > 0$—together with a positive marginal appropriation $T_h > 0$ (Assumption 3). Because the novice's stock stays bounded as $t \to 0$, T_h stays bounded away from zero there, which is all the argument uses. We do not impose the Inada condition as a primitive; the results below take co-production interior.

3.4 AI's Impact on the Steady-State Stock

The steady state solves $h^*(A) = H(h^*(A), t^*(A), A)$ with $t^*(A)$ from (3)—a positive fixed point of the equilibrium transition map $F(h) \equiv H(h, t^*(h; A), A)$ that folds the expert's optimal co-production into the transition. Such a fixed point exists, is stable, and is unique under mild regularity conditions.

Lemma 2 (Steady state)**.** *Fix $A \geq 0$. The map F is continuous, and under Assumption 2 it vanishes at the origin and has diminishing returns to the inherited stock. The conditions below restrict this equilibrium map, which embeds the optimal policy t^*. Then:*

(i) ***Existence.*** *If F rises above the 45-degree line at some stock, and the transmission's elasticity in the inherited stock stays bounded away from one at large stocks, F has a positive steady state $h^*(A)$.*

(ii) ***Stability.*** *If the map's slope at a steady state is less than one in absolute value, that steady state is locally stable.*

(iii) ***Uniqueness.*** *If the map's elasticity is strictly decreasing in h, the stable steady state is unique.*

Assumption 2 does the structural work; the further conditions are regularity restrictions, and the calibrated class of Section 6 satisfies them all. The proof is in Online Appendix A.

The comparative statics below are read at this stable steady state. There, three local elasticities of H govern them, each evaluated at $(h^*(A), t^*(A), A)$: the *persistence* $\gamma \equiv \partial \ln H / \partial \ln h$, below one by Assumption 2; the *learning-by-doing elasticity* $\epsilon \equiv \partial \ln H / \partial \ln t$; and the *reach* $\nu \equiv \partial \ln H / \partial A$, weakly positive since AI does not impair transmission.[10] At the steady state the persistence coincides with the marginal product of inherited capital, $\gamma = H_h$, since $H = h^*$ there. Throughout, we display the AI level only on objects whose movement with A the analysis tracks, chief among them the steady-state stock $h^*(A)$ and the reach $\nu(A)$. We write the rest bare.

Proposition 1 (AI's impact on the equilibrium)**.** *At the stable steady state of Lemma 2, $(h^*(A), t^*(A))$,*

$$\frac{dh^*(A)}{dA} \;=\; \frac{h^*(A)\,\big(\nu(A) - \epsilon\,\mu\big)}{1-\gamma}, \tag{5}$$

[10] The reach is a *semi*-elasticity (the log is in the numerator only), unlike the full elasticities γ and ϵ. AI capability enters in levels because the no-AI baseline $A = 0$ is admissible and central—the intergenerational wedge is read there—and $\partial \ln H / \partial \ln A$ is undefined at $A = 0$. The level convention also keeps the threshold $\nu = \epsilon\mu$ (Section 4.2) dimensionally homogeneous, the compression $\mu = -\,d \ln t^* / dA$ being a semi-elasticity in A as well.

with $\gamma < 1$ there (Assumption 2), $\nu(A)$ the reach, ϵ the learning-by-doing elasticity, and μ the compression (4). The stock erodes *($dh^*(A)/dA < 0$) where $\nu(A) < \epsilon\mu$ and is* augmented *($dh^*(A)/dA > 0$) where $\nu(A) > \epsilon\mu$: AI's net effect on the stock is a contest between its reach into transmission, $\nu(A)$, and the compression-weighted learning by doing it displaces, $\epsilon\mu$. Both contestants are* total *steady-state responses, not partials at a fixed inherited stock: each is read along the fixed point as the stock adjusts. At the boundary $\nu(A) = \epsilon\mu$ the distinction dissolves. A total response differs from its fixed-h_k counterpart only by a term proportional to the stock displacement $dh^*(A)/dA$, and the boundary is exactly where that displacement vanishes.*

The two contestants of equation (5) act on the same formation. The first, $\epsilon\mu$, is the learning by doing AI *displaces*. As work done with AI grows more productive, co-producing with the novice costs more and t^* falls; each unit of co-production lost carries ϵ of formation with it, lowering the stock through transmission. The second is the reach $\nu(A)$, the formation AI *re-supplies*: codified instruction that raises what each unit of co-production transmits, wherever the knowledge is codifiable. Where the reach dominates, the stock grows; where it does not, it shrinks.

Steady-state output $Y^*(A) = (1 - t^*(A))\,\Phi(h^*(A), A) + T(h^*(A))$ need not decline even where the stock erodes, since the production aggregator gains directly from AI. Where erosion is deep enough, it can pull output down with the stock; whether output can *collapse* is governed by essentiality (Online Appendix Proposition G.1).

4 The Social Planner

The decentralized equilibrium of Section 3 is inefficient. By Assumption 4 the expert counts her novice's contemporaneous contribution but not the value his stock carries into all later generations, so she underinvests in co-production relative to a planner who internalizes the full transmission chain. This section studies that planner. We derive the shadow value of knowledge and the resulting intergenerational wedge. We then price AI's marginal effect on the stock with a Pigouvian damage index whose sign follows from that effect and whose zero is governed by a *force-free* condition. Once experts also choose how much AI to use, the correction becomes a problem of instrument assignment.

4.1 The Planner's Problem and the Intergenerational Wedge

The social planner discounts future generations at $\delta \in (0, 1)$ and values each period at the realized output Y_k of (2), the expert's production plus the novice's contemporaneous

contribution. The planner solves

$$V(h_k, A) = \max_{t_k \in [0,1]} \Big\{ (1 - t_k)\Phi(h_k, A) + T(h_{k+1}) + \delta V(h_{k+1}, A) \Big\}. \tag{6}$$

This problem has a unique bounded value function V, increasing in the stock and differentiable where the planner's policy is regular.[11] By Assumption 4, the decentralized expert's objective Y_k excludes the continuation value $\delta V(h_{k+1}, A)$. The planner's first-order condition adds the discounted shadow value of h_{k+1} to the expert's marginal appropriation:

$$-\Phi(h_k, A) + \big(T_h(h_{k+1}) + \delta\, V'\big)\, H_t \;=\; 0, \qquad V' \equiv \partial V / \partial h. \tag{7}$$

The shadow value V' follows from the envelope theorem applied to (6) (the response of the planner's optimal co-production to h_k drops out by the first-order condition (7)):

$$V'(h_k, A) = (1 - t_k)\, \Phi_h + \big(T_h(h_{k+1}) + \delta\, V'(h_{k+1}, A)\big)\, H_h.$$

The shadow value has two parts: the stock's marginal product in current output, and its effect on the next-generation stock weighted by what that stock is worth. That value is itself the marginal appropriation plus the discounted shadow value of the next generation, so the relation is recursive.

The shadow value is precisely what the decentralized expert omits. A comparison of the planner's first-order condition (7) with the decentralized expert's (3) shows that the planner values the marginal hour of co-production more highly.

Proposition 2 (Intergenerational wedge)**.** *The planner's marginal benefit of co-production*

[11] Standard dynamic-programming arguments deliver existence, uniqueness, and boundedness (Stokey et al., 1989): for *any* feasible co-production choice the transition is bounded by its full-intensity value, $H(h_k, t_k, A) \le H(h_k, 1, A)$. Past a finite stock, $H(h_k, 1, A)$ falls below the 45-degree line (proof of Lemma 2). A compact interval therefore absorbs every feasible path—the planner's as well as the decentralized one—and on it the continuous period return is bounded. The value function V increases because the flow and the transition both rise with the stock. Differentiability is the less immediate property. The period return need not be jointly concave in (h_k, t_k): its $(1 - t_k)\Phi$ term alone has an indefinite Hessian. Differentiability therefore rests on the envelope theorem of Milgrom and Segal (2002) rather than on concavity of V (Benveniste and Scheinkman, 1979), and that theorem yields a single derivative only where the maximizer is unique. We maintain, as a regularity condition wherever a derivative of V is read, that the planner's optimal co-production at the point of evaluation is unique, and interior or at a locally persistent corner; every normative statement below is conditional on exactly this regularity. Where a statement reads the derivative of a value evaluated along a privately chosen path—W in Lemma A.1, $\widehat{W}$ in Proposition 4—its positivity at a steady state follows from non-amplifying stability (Online Appendix A); away from a steady state, it is read as part of the same regularity. In the calibration the damped fixed-point computation of Online Appendix F converges to a single policy from dispersed starting values at every point where these objects are evaluated.

exceeds the decentralized expert's by

$$\textit{WEDGE}_k \;=\; \delta\, V'(h_{k+1}, A)\, H_t \;>\; 0, \tag{8}$$

strictly positive because V increases in the stock (so $V' > 0$) and $H_t > 0$. The wedge is the discounted continuation value the marginal hour of co-production creates through the next-generation stock: the hour raises h_{k+1} by H_t, and the shadow value capitalizes the recursive transmission chain $h_{k+1} \to h_{k+2} \to \cdots$ that the private objective omits (Assumption 4).

The expert and the novice need not be distinct people: the expert may stand for an occupation that forms whoever does its work rather than delegating it to AI, and the novice for that worker; the wedge is unchanged. What the wedge requires is only that whoever sources the doing be rewarded over a span shorter than the chain the formation feeds, here a single generation. Were that span to cover the whole chain, private and social returns would coincide and the wedge would vanish. That is the case of a single long-lived agent who internalizes every later cohort.

This is an intergenerational externality, not the dynamic inefficiency of the overlapping-generations tradition. In the Diamond (1965) economy markets are complete and no private return departs from its social counterpart; the competitive equilibrium is inefficient only when it overaccumulates capital, saving past the golden rule, and whether it does so turns on the interest rate against the growth rate. Ours is the reverse. A single missing market, the uncontractible tail of the transmission chain, holds the private return below the social one. The economy therefore transmits too little, at every discount factor and every AI level, with no golden-rule knife-edge to cross. Even the corrections point opposite ways: transfers that decumulate an excess stock there, a subsidy that encourages the underprovided co-production here.

4.2 The Pigouvian Damage Index and Its Force-Free Threshold

Proposition 2 gave the wedge (8) as a per-period quantity: the gap between the planner's and the expert's marginal benefit of co-production. The wedge prices an hour of co-production, not AI itself. AI policy needs a second price: a sign for AI's net effect on formation, a magnitude for the correction. A one-period gap cannot be that price, because AI moves a stock. AI's compression of co-production lowers the stock, while its reach into transmission raises it; whichever prevails, the change persists and compounds down the transmission chain.

That compounding change is exactly what the expert's objective omits. Pricing it is Pigou's prescription (Pigou, 1920): a correction signed and sized by marginal damage, read at

the allocation the correction itself supports. Proposition 3 builds that price for AI capability: the *Pigouvian damage index*. The construction takes two objects, the shadow value in closed form and the allocation at which every term is read.

The shadow value comes first. At a steady state ($h_{k+1} = h_k$, so $H_h = \gamma$), the envelope recursion closes and the shadow value admits a closed form.

Lemma 3 (Steady-state shadow value)**.** *The steady-state shadow value is*

$$V'(A) \;=\; \frac{(1 - t^*(A))\,\Phi_h + T_h\,\gamma}{1 - \delta\,\gamma} \;>\; 0. \tag{9}$$

The numerator is the stock's per-period marginal value, capitalized over the transmission chain by the geometric sum $1/(1-\delta\gamma)$. Holding the allocation fixed, the shadow value rises with the marginal appropriation and the persistence.

The allocation completes the construction. The index is read at the *corrected* steady state: allocation and shadow value jointly consistent under the planner's condition (7).[12]

Proposition 3 (The Pigouvian damage index)**.** *With co-production interior, the Pigouvian damage index is the marginal steady-state stock displacement valued at the discounted shadow price:*

$$\tau^*(A) \;=\; \delta\,V'(A)\left(-\frac{dh^*(A)}{dA}\right) \;=\; \underbrace{\frac{\delta\,V'(A)\,h^*(A)}{1-\gamma}}_{>\,0} \cdot \big(\epsilon\mu - \nu(A)\big), \tag{10}$$

with $V'(A)$ the shadow value (9), μ the compression (4), ϵ the learning-by-doing elasticity, and $\nu(A)$ the reach. Then:

(i) ***Sign.*** *Since the magnitude prefactor is strictly positive,* $\operatorname{sign}\tau^*(A) = \operatorname{sign}\big(\epsilon\mu - \nu(A)\big) = -\operatorname{sign}\big(dh^*(A)/dA\big)$*, every term read at the corrected steady state: the index is a marginal social cost where AI erodes the stock and a marginal social benefit where AI augments it.*

(ii) ***Force-free threshold.*** *The index vanishes, $\tau^*(A) = 0$, exactly where $\nu(A) = \epsilon\mu$. This threshold is technological, in transmission and production jointly: the compression μ*

[12]This corrected steady state exists. Given a conjectured shadow value $v \geq 0$, the effective marginal appropriation $T_h + \delta v$ induces a stable steady state by the argument of Lemma 2. The lemma's hypotheses survive the conjecture: the effective appropriation $T(h) + \delta v\,h$ inherits Assumption 3; the transition's bound $H(h, 1, A)$ and persistence are free of the conjecture; and a higher conjecture only raises co-production, preserving the crossing. The shadow value (9) implied at that steady state is the conjecture's update. The update is continuous in the conjecture, by regularity of the interior optimum and the implicit function theorem. It is bounded: as the conjecture grows, co-production approaches its upper corner and the update's numerator stays bounded. The update therefore sends a compact interval into itself, and Brouwer's theorem gives a fixed point. The damped iteration that computes it (Online Appendix F) converges to the same point from dispersed starting values throughout the calibration.

carries the production margin $\partial \ln \Phi / \partial A$. It contains no appropriation, shadow value, or discount term. These enter only the positive prefactor; the appropriation enters it through the shadow value V'. They therefore scale $|\tau^*|$ and can shift the equilibrium at which the contest $\epsilon\mu - \nu(A)$ is read, but they cannot set its sign or its zero.

The index is anchored to an exact welfare object. Online Appendix A derives the *marginal externality of AI capability* (Lemma A.1): the uninternalized part of the welfare derivative of a permanent increment in the AI level, read along the path the experts steer. It is negative exactly where AI erodes the stock. The index is that externality re-expressed as a price, through three conversions. It reverses the orientation: a damage where welfare falls. It compounds the one-generation displacement through the inheritance chain into the permanent shift of the steady state: each displacement moves the next inheritance, and the rounds accumulate until the stock settles. It is read at the corrected allocation, while the externality is read at the decentralized one. The first two conversions are a sign flip and a positive scaling, so the index inherits the externality's contest and its algebraic zero $\nu = \epsilon\mu$ exactly. The third moves only where the terms are read: the shared zero is a shared form, not a shared location. Re-read at the corrected allocation, the contest can come out the other way: erosion at the decentralized one need not place the corrected index on the tax side.[13]

The force-free property splits the policy problem in two. The intergenerational friction sets *whether* to intervene and *how much*: the wedge is positive, so a correction is warranted, and its size scales with the shadow value. The technologies set the *direction*: whether the correction runs tax-side or subsidy-side turns only on how far AI reaches into transmission against the learning by doing it displaces.

Not even the size of the wedge bears on the direction: optimal AI policy is oriented by technological primitives. The valuation parameters retain one channel, relocating where the contest is read, and the calibration bounds it tightly (Online Appendix H). Behind the property is a cancellation, the echo of the bundle. The expert's within-period return runs through the very stock that carries the formation, so at her optimum the appropriation drops out of the threshold (Online Appendix A). The occupational sign reversal of Section 5 builds on exactly this property.

Because A is exogenous (experts choose co-production, not AI use), the index is not a levied rate. It prices instead the externality's geography—sign, zero, and relative size. That

[13] From the fixed point $h^* = H(h^*, t^*(h^*; A), A)$, the settled displacement is the per-period impact times the stable multiplier: $dh^*/dA = (\partial h_{k+1}/\partial A)\big/\big(1 - dH/dh_k|_{t^*(\cdot)}\big)$, the denominator positive by stability. AI's direct production gain $(1 - t^*)\Phi_A$, which the expert internalizes, is in neither the externality nor the index. The per-period wedge on the use margin is the levied object $\hat{\tau}$ of Section 4.3.

geography is the proper target of any instrument that prices AI. The instruments themselves, and the assignment between them, are the subject of Section 4.3.

4.3 Implementation: Adoption and Instrument Assignment

The model has two *optimal rates*—the term we reserve for instruments derived from explicit planner problems—and which one the planner can deploy turns on what is contractible. If co-production can be instrumented, that instrument alone restores the optimum and the optimal AI tax is zero (Corollary 1). When co-production cannot be contracted on, the planner prices AI instead (Proposition 4); the non-verifiability behind Assumption 4 makes this the relevant case.

Pricing AI use—or showing it should go unpriced—needs a margin to act on. The period-k expert now chooses, alongside co-production, an AI use level $a \in [0, A]$; the AI level A of Section 3.1 becomes the *frontier* of feasible use. Use is priced at a resource cost $p > 0$ per unit, the market price of AI services. The chosen level enters production and transmission in place of the frontier: output is $(1 - t_k)\Phi(h_k, a)$, transmission is $h_{k+1} = H(h_k, t_k, a)$, and the expert's payoff is

$$Y_k \;=\; (1 - t_k)\,\Phi(h_k, a) \;+\; T(h_{k+1}) \;-\; p\,a. \tag{11}$$

The conventions of Section 3 carry over, with partial derivatives in a read at the chosen use level ($\Phi_a \equiv \partial\Phi/\partial a$, $H_a \equiv \partial H/\partial a$). At an interior, regular optimum the expert balances each margin,

$$T_h\,H_t \;=\; \Phi(h_k, a), \qquad\qquad (1 - t_k)\,\Phi_a \;+\; T_h\,H_a \;=\; p, \tag{12}$$

the first the co-production trade-off of (3), the second equating the marginal value of AI use—its production gain plus its appropriated transmission gain—to its price. The baseline is nested at a free frontier: at $p = 0$ the marginal value of use is positive wherever AI raises production, the expert adopts to the frontier, $a = A$, and Section 3 applies verbatim. A positive price puts the margin in play: use settles at zero, in the interior, or at the frontier, as the price compares with that value. The live margin is what gives price-side instruments their traction.

The planner's problem adds the continuation value, as in (6):

$$V(h_k, A) \;=\; \max_{t_k \in [0,1],\; a \in [0,A]} \Big\{ (1 - t_k)\Phi(h_k, a) + T(h_{k+1}) - p\,a + \delta\,V(h_{k+1}, A) \Big\}, \tag{13}$$

with first-order conditions $-\Phi + (T_h + \delta V')H_t = 0$ and $(1 - t_k)\Phi_a + (T_h + \delta V')H_a = p$. Set

against these, the expert's two conditions (12) miss the *same* term: the discounted shadow value $\delta V'$, applied to the marginal transmission—H_t on the co-production margin, H_a on the use margin. The externality operates through a single channel, the novice's stock h_{k+1}, however many choices feed it.

Suppose first that the co-production margin can be instrumented. The instrument pays the expert in proportion to the appropriation, at rate $s^*(A) = \delta V'(A)/T_h$, with T_h read at the corrected steady state. The payment raises the expert's effective marginal appropriation to $T_h + \delta V'$, the planner's marginal valuation of the novice's stock. Off the steady state the instrument generalizes as a payment schedule: the expert receives $S(h_{k+1})$ alongside the appropriation, with marginal payment $S'(h_{k+1}) = \delta V'(h_{k+1})$ read at the realized next-generation stock. Her effective marginal appropriation is then $T_h + \delta V'(h_{k+1})$ at every state—the planner's own—and the constant rate s^* is this schedule's steady-state restriction.

The rate's sign is a result, not part of the design. The decentralized economy underprovides co-production everywhere (Section 4.1), and a correction to an underprovided margin can only encourage it. The strictly positive wedge of Proposition 2 is that underprovision in marginal form: s^* is positive at every state, and the instrument is a subsidy throughout—the *co-production subsidy*. The damage index τ^* and the subsidy are two faces of one correction, both built from the discounted shadow value $\delta V'$: the former prices AI's effect on the stock, the latter restores that value to the co-production return.

The co-production subsidy therefore corrects both margins at once.

Corollary 1 (First-best instrument assignment)**.** *Suppose the planner pays the co-production subsidy $s^*(A) = \delta V'(A)/T_h$ constructed above, with both margins interior and regular at the corrected steady state. Then the subsidized expert's first-order conditions on* both *margins coincide with the planner's at that steady state: the subsidy alone supports it, and the optimal tax on AI use is zero,*

$$\tau^{FB}(A) \;\equiv\; 0. \tag{14}$$

The proof (Online Appendix A) is one observation: the subsidized expert's effective marginal appropriation $T_h + \delta V'$ multiplies H_t in the co-production condition and H_a in the use condition alike, reproducing the planner's pair. At the first best the planner never taxes AI, in either direction. Erosion is real, but it is a symptom of underpriced transmission, not of underpriced AI. Once the transmission externality is paid for, private AI use is efficient. This is the principle of targeting (Bhagwati and Ramaswami, 1963): the correction belongs at the distorted margin, co-production, leaving every other price undistorted.

The assignment above presumes the co-production margin can be instrumented. The friction that generates the wedge says otherwise: the subsidy must condition on transmission or on the value it generates within the period, the very objects whose non-verifiability

is a leading microfoundation of Assumption 4. When no co-production-side instrument is feasible, the planner's remaining lever is the price of AI use: experts face $p + \tau$, with the revenue rebated lump-sum. With the targeted instrument out of the feasible set, the constrained optimum prices the use margin away from its first-best level (Lipsey and Lancaster, 1956). Only here does the planner face AI's production value against formation erosion as a genuine trade-off: at the first best the subsidy carries the correction and the market price allocates use, while the constrained planner does both with one price.

Proposition 4 (Second-best AI pricing)**.** *Suppose the co-production margin cannot be instrumented and the planner corrects the price of AI use, the expert choosing (t_k, a) privately at the corrected price. Write $\widehat{W}$ for the constrained planner's value (use chosen to maximize discounted welfare subject to the expert's private co-production response) and $\widehat{W}'$ for its derivative. Then:*

(i) ***Rate (interior margin).*** *At states where the constrained optimum has an interior, regular use margin, the Markov schedule*

$$\hat{\tau}(h_k) \;=\; \delta\,\widehat{W}'(h_{k+1})\left(-\frac{\partial h_{k+1}}{\partial a}\right), \qquad \frac{\partial h_{k+1}}{\partial a} \;\equiv\; H_a + H_t\,\frac{\partial t^*}{\partial a}, \tag{15}$$

levied period by period, makes the expert's use condition coincide with the constrained planner's at every such state along the induced path. The damage it prices is the negative of the stock's response to use, which sets the direct transmission gain from use (H_a) against the co-production that higher use crowds out ($H_t\,\partial t^*/\partial a$). At the steady state that the schedule induces, every object is stationary and the schedule reduces to a constant rate: the fixed point of (15) with all terms read at the allocation that rate supports (Online Appendix A).

(ii) ***Sign.*** $\operatorname{sign}\hat{\tau} = -\operatorname{sign}\partial h_{k+1}/\partial a$ *at each such state: a tax where AI use erodes the novice's stock, a subsidy where it builds it. Use erodes the stock when the crowded-out learning by doing outweighs the reach.*

(iii) ***Force-free threshold.*** *The interior schedule vanishes where $\nu = \epsilon\,\mu_a$, with $\mu_a \equiv -\partial \ln t^*/\partial a$ the fixed-h_k use-compression, $\nu = \partial \ln H/\partial a$ the reach at the chosen use level, and ϵ read at the adoption equilibrium. This is the use-margin counterpart of the threshold in Proposition 3, and, like that one, it is technological: it contains no appropriation, shadow value, or discount term; those scale $|\hat{\tau}|$ but cannot set its sign or its zero.*

(iv) ***Corners.*** *Where the constrained optimum puts use at a corner, the supporting correction is a set. At zero use, every rate no smaller than the minimal supporting rate* $\tau_0(h_k) \equiv (1-t^*)\,\Phi_a + T_h H_a - p$*, read at* $a = 0$*, implements* $a = 0$*—the allocation that prohibition selects, at the same welfare. At the frontier, every rate no larger than the analogous* $\tau_A(h_k)$ *implements* $a = A$*. The interior formula of part (i) prices the band between.*

The result admits two readings. First, the friction works twice: having created the intergenerational wedge, it also removes the instrument aimed at that wedge, and what survives is the AI price. That price is therefore not one implementation among several but the one the friction itself selects. Its rate is Pigouvian in form, evaluated at the constrained allocation (Sandmo, 1975).

Second, $\hat{\tau}$ and the damage index τ^* are different magnitudes of the same structure: $\tau^*(A)$ prices the cumulated steady-state displacement from the AI level, while $\hat{\tau}$ prices the per-period damage of marginal use. The two share the sign rule and the force-free form of the threshold, each read at its own equilibrium.

The two optima also price the friction. The distance from the decentralized equilibrium to the constrained optimum is the value of internalization: what correction through the reachable margin recovers. The distance from the constrained optimum to the first best is the value of contractibility: what reaching the source would be worth. In the calibration the second distance dwarfs the first (Section 6.5).

The corners of part (iv) are not a technicality. Where use strongly builds the stock, the optimal use subsidy drives the expert to the frontier, and any deeper subsidy supports the same allocation. On the suppression side, a bounded marginal product of the first unit of use means that a finite price implements exactly zero use; any technology in which AI capability adds to an installed base of tools delivers the bound, as the calibration's tooling floor does (Section 6). Suppression requires no ban, only a rate no smaller than the minimal supporting level, and every such rate shares prohibition's allocation and welfare.

When the calibrated cross-section of Section 6 is read through this proposition, nearly every occupation sits at one of these corners. The policy content there is the partition into zero-use and frontier occupations together with its minimal supporting rates.

5 The Occupational Cross-Section

The mechanism of Sections 3–4 is stated for one occupation: AI erodes or augments the knowledge stock as its reach ν falls short of or exceeds the compression-weighted learning

by doing $\epsilon\mu$ (Proposition 1). The damage index follows the erosion, vanishing at the force-free threshold $\nu = \epsilon\mu$ (Proposition 3). The tacitness of knowledge is already inside this mechanism: the reach is the formation that AI re-supplies as codified instruction, so it extends only as far as the occupation's knowledge can be told. What this section adds is the cross-section. It names the primitive that sets the reach—the occupation's expressibility—and reads the mechanism across occupations that differ in that primitive alone. Carried across the spectrum, the threshold becomes a *cut* and the index a *schedule* that reverses sign: the tax direction where knowledge is tacit, the subsidy direction where it is codifiable.

5.1 Expressibility and the Reach

Occupations differ in *what kind* of knowledge their stock is; the dimension that matters here is how far it can be codified. This is the occupation's *expressibility* $\rho \in [0, 1]$, the degree to which the knowledge can be transmitted by prescription rather than by example (Polanyi, 1966; Autor, 2014). Equivalently, it is the degree to which the knowledge can be acquired apart from the doing rather than only through it. Expressibility runs from the *tacit* pole ($\rho = 0$) to the *explicit* pole ($\rho = 1$).

An occupation here is what it has been throughout: one overlapping-generations transmission chain, now indexed by the expressibility of the knowledge it passes down. The occupations in the data enter with the calibration, as measured counterparts of these chains. The chains form a continuum in ρ, each evolving independently and each facing the same exogenous AI level A. Because A is one scalar shared by every occupation, expressibility is the sole *structural* source of cross-occupation variation in AI's effect: how any prediction responds to A differs across occupations only through ρ.[14]

We add the expressibility, after a semicolon, to the steady-state objects of Sections 3–4 whose movement across occupations the analysis tracks: the stock $h^*(A; \rho)$, co-production $t^*(A; \rho)$, shadow value $V'(A; \rho)$, damage index $\tau^*(A; \rho)$, and the reach $\nu(A; \rho)$ that drives the sign; the cut-point $\bar{\rho}(A)$ depends on the level alone. Expressibility enters the analysis through the reach: AI's re-supply of the formation widens as the knowledge becomes more codifiable.

Assumption 5 (AI-augmentation channel: the reach)**.** AI augments transmission at every state, $\partial H/\partial A \geq 0$. Write the strength of that augmentation as the semi-elasticity

[14]Occupations also differ in how much of their work AI can perform: an exposure intensity $\omega_j \in [0, 1]$ that the calibration adds as a second occupation characteristic, empirically near-independent of expressibility (Section 6). Occupation j then faces the effective level $A\omega_j$: exposure shifts where the occupation sits on the A-axis and rescales the magnitude of its response. But it operates through the same contest and the same force-free cut, refining the calibration without altering any structural result. Proposition D.1 in Online Appendix D establishes this formally.

$\partial \ln H(h_k, t_k, A; \rho)/\partial A$, whose steady-state value is the reach ν of Section 3.4. At every state (h_k, t_k, A), this strength is continuous and strictly increasing in ρ; at the tacit pole it is zero, $\partial H/\partial A = 0$ when $\rho = 0$. We assume the same pattern at the equilibrium: the steady-state reach $\nu(A; \rho)$, the semi-elasticity read at each occupation's own steady state, is continuous and strictly increasing in ρ.

Only the monotonicity needs the equilibrium clause. The zero at the tacit pole carries over on its own: a semi-elasticity that vanishes at every state vanishes at the steady state. The monotonicity does not, because the steady state itself moves with ρ. Whenever the augmentation enters transmission multiplicatively, $H = \widetilde{H}(h_k, t_k; \rho)\, \psi(A; \rho)$, the semi-elasticity depends on no state variable, and the fixed-state and equilibrium properties coincide. The parametric class the calibration uses is of this form.

The floor at the tacit pole encodes Polanyi's hard claim: purely tacit knowledge cannot be acquired by codification, however capable the AI or however much stock and co-production the match deploys—it has no route but the doing. The results ask less of the floor than the paradox supplies. The tax direction at the tacit pole requires only that the reach fall short there, $\nu(A; 0) < \epsilon\mu$, so the exact zero is a clean normalization of that inequality. Machine demonstration that trains without telling—a surgical simulator, say—moves no result until it lifts the reach past the doing it displaces.

5.2 Erosion and Augmentation across Occupations

With the reach rising in expressibility, the single-occupation contest of Proposition 1 divides the spectrum.

Proposition 5 (Cross-occupation erosion and augmentation)**.** *Under Assumption 5, suppose the contest $\nu(A; \rho) - \epsilon\mu$ (each term read at the occupation's own steady state) is continuous and monotone in ρ and changes sign in the interior of the spectrum.*[15] *Then AI's effect on the steady-state stock sorts by expressibility: $dh^*(A; \rho)/dA < 0$ at every expressibility below a single interior* crossover *and $dh^*(A; \rho)/dA > 0$ at every one above it—erosion in the tacit-type occupations, augmentation in the codifiable ones.*

[15] Assumption 5 orders only the reach. The contest's other side, the compression-weighted learning by doing $\epsilon\mu$, is an equilibrium object and moves with ρ as well. Monotonicity of the contest therefore adds exactly one condition: $\epsilon\mu$ does not rise faster than the reach. The two sides are coupled in the condition's favor: when the augmentation scales transmission, the same force that widens the reach raises the return to co-production, blunting the compression (Section 3.3). A broad primitive class delivers the condition outright. Multiplicative augmentation, isoelastic co-production, and power appropriation make the contest strictly increasing in ρ at a common-stock read, for any production technology; log-separable production carries the single crossing to each occupation's own steady state (Lemma A.2, Online Appendix A). The calibration meets the three conditions and reads the cross-section at a common stock, its normalized baseline.

At the tacit pole the contest is settled before it begins. The reach floor leaves the appropriated return to co-production flat in A. The substitute condition of Lemma 1 then collapses to $\Phi_A > 0$: wherever AI raises production at all, it compresses co-production and, through the stable transition of Lemma 2, erodes the stock. The erosion is in this sense structural: the signature of a technology that aids production without aiding transmission. At the explicit pole the re-supply route opens: AI supplies part of the formation directly, and where that reach outweighs the compression-weighted learning by doing, the stock expands. Augmentation there mirrors the erosion at the tacit pole—the signature of a technology that aids transmission, not only production.

5.3 The Sign-Reversing Schedule

The correction inherits this geometry. Read across the spectrum, the index of Proposition 3 is a continuous schedule, and its force-free threshold becomes an interior cut-point. Figure 2 draws the result: the reach rises from the Polanyi floor and crosses the compression term once, and the crossing splits the spectrum into a tax region and a subsidy region.

Proposition 6 (Sign reversal and the force-free cut)**.** *Under Assumption 5, with the contest $\epsilon\mu - \nu(A;\rho)$ continuous and monotone in ρ and changing sign in the interior of the spectrum:*

(i) ***Single sign reversal.*** *The shadow schedule $\tau^*(A;\rho)$ of (10) reverses sign exactly once: the tax direction at the tacit pole gives way to the subsidy direction at the explicit pole. Its* size*, scaled by the strictly positive prefactor of (10), varies across occupations through $V'(A;\rho)$, $h^*(A;\rho)$, and the persistence, so $|\tau^*|$ need not be monotone in ρ.*

(ii) ***Force-free cut.*** *There is a unique cut-point $\bar{\rho}(A) \in (0,1)$ at which $\tau^*(A;\bar{\rho}(A)) = 0$, characterized by*

$$\nu(A;\bar{\rho}(A)) = \epsilon\,\mu \tag{16}$$

with $\tau^*(A;\rho) > 0$ (the tax direction) for $\rho < \bar{\rho}(A)$ and $\tau^*(A;\rho) < 0$ (the subsidy direction) for $\rho > \bar{\rho}(A)$. The cut is force-free*: neither the appropriation technology T nor the shadow value V' enters the condition $\nu(A;\bar{\rho}) = \epsilon\mu$ (Proposition 3); the shadow value scales the schedule's magnitude, and both act on the cut's location only by relocating the equilibrium at which the condition is read.*

(iii) ***Reversal of the levied price.*** *Where the use margin is interior, with the use-margin contest $\epsilon\mu_a - \nu$ continuous and monotone in ρ and changing sign in the interior of the spectrum, the second-best schedule $\hat{\tau}(A;\rho)$ of (15) reverses sign at its own force-free cut, $\nu = \epsilon\mu_a$ (Proposition 4). The pattern is the same reversal—a tax on the tacit side of*

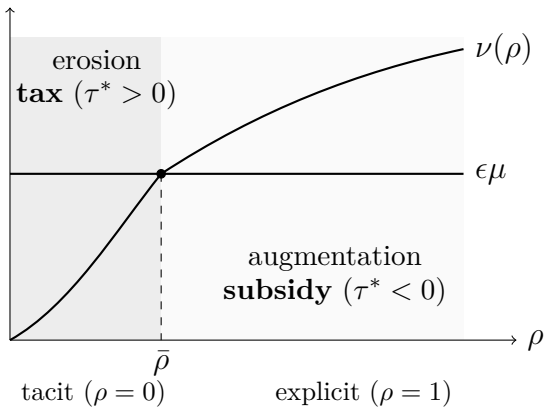


Figure 2: The Sign Reversal across Expressibility (Schematic)

Note: AI's reach into transmission at a fixed AI level, written $\nu(\rho)$, rises from the Polanyi floor $\nu(0) = 0$ at the tacit pole and crosses the compression term $\epsilon\mu$ once, at the force-free cut $\bar{\rho}$ where $\nu(\bar{\rho}) = \epsilon\mu$ (Proposition 6). To its left the reach falls short, AI erodes the stock, and the index is positive—the tax direction ($\tau^* > 0$); to its right the reach dominates, AI augments the stock, and the index is negative—the subsidy direction ($\tau^* < 0$). The term $\epsilon\mu$ is drawn flat for clarity; the argument needs only that the contest $\epsilon\mu - \nu(\rho)$ be monotone and cross zero once.

the cut, a subsidy on the explicit side—read on the adoption margin rather than the AI level, at a different magnitude.[16]

The reversal rests on three ingredients of different standing: a theorem, an assumption, and a hypothesis. The theorem is the force-free property of part (ii): it holds nonparametrically, by the cancellation behind Proposition 3. The assumption is the Polanyi floor (Assumption 5), which signs the tacit endpoint: at zero reach the contest $\epsilon\mu - \nu(A; \rho)$ is strictly positive wherever AI compresses co-production. The hypothesis is the monotone contest crossing zero in the interior. The calibrated class of Section 6 derives it from primitives at the normalized steady state, and the crossing at the corrected equilibrium is verified numerically. Were the reach never to overtake the compression term, the schedule would run in the tax direction throughout—the automation special case the calibration prices.

Automation, as the task-substitution literature models it, does the work but cannot teach: it is the case $\nu \equiv 0$. With the reach gone from the contest, its damage index reduces to the positive prefactor of (10) times $\epsilon\mu$—the learning by doing it displaces—and the sign is set by the compression μ alone. Automation therefore carries no reversal in expressibility. The reversal, a subsidy precisely where the knowledge is codifiable, requires $\nu > 0$: the transmission channel that sets AI apart from prior automation. In a codifiable occupation where both displace the doing, automation erodes the occupation's transmission while AI

[16] At corner allocations, where the calibrated cross-section puts nearly every occupation (Section 6), the levied object is the support set of Proposition 4(iv). The partition into zero-use and frontier occupations, with its smallest-supporting-rate selection, is a numerical result.

augments it; the two coincide only at the tacit pole, where neither can teach.

Two consequences follow for policy. First, a uniform AI instrument is qualitatively misspecified, not merely suboptimal: the index spans both signs across the spectrum. Any single-signed instrument therefore runs against the correction on one side of the cut—a tax where the corrective direction is a subsidy, or the reverse. The common instinct—that AI threatens skill formation and should be restrained to protect it—belongs to one side of the cut only. On the tacit side, where AI erodes, restraint is the correction. On the codifiable side, where AI augments, the instinct points the wrong way: it would curb the very technology the correction favors.

Second, the reversal separates cleanly into a structural part and an empirical one. *That* the schedule reverses—the tax direction on the tacit side of the cut, the subsidy direction on the explicit side—is the structural part, carried by the floor and the monotone contest. *Where* it reverses and how sharply—the location of $\bar{\rho}(A)$ and the steepness of the schedule around it—are empirical, set by the magnitudes of the reach, the compression-weighted learning by doing, the persistence, and the production technology. The next section specializes these to the parametric forms we calibrate and illustrates where the cut falls.

6 Calibration to the U.S. Occupational Cross-Section

We now take the model to the data. Section 6.1 specializes the general primitives of Sections 3–5 to a parametric form, preserving the sign reversal of Proposition 6. The remainder of the section calibrates that form to 682 U.S. occupations, on which both axes of the analysis, expressibility and exposure, are measured.

The theory is the structural content: the force-free cut holds for the general primitives, and the single reversal follows wherever the net contest is monotone with an interior crossing—the hypothesis this section derives within the calibrated class. The calibration asks where the cut lies: whether the reversal lands at an empirically relevant place in the occupation distribution, and how much of the workforce sits on each side. It is a demonstration of the mechanism's scope rather than an estimate of its primitives. Its ordinal findings—a single interior cut, the index reversing across it, the deepest tax-direction values in the tacit-and-exposed corner—ride on the two measured axes; the cardinal magnitudes are the illustrative tier.

The findings are robust in layers. The sign reversal and the interior cut survive the substitution sweep, every $\pm 30\%$ structural perturbation, the bending of the augmentation schedule, and rank-versus-cardinal measurement of both axes. Membership in the tax-side set is the measure-dependent layer: acquisition-based orderings preserve the reversal and the

cut, keeping the aggregate tax-side share within the structural-sweep range while relocating which occupations sit below the cut (Online Appendix H).

6.1 Model Parameterization

We adopt the standard multiplicative skill-formation form (Ben-Porath, 1967; Cunha and Heckman, 2007; Cunha et al., 2010), a Cobb–Douglas core in the inherited stock and co-production time, scaled by a separable AI factor $\psi(A;\rho)$ with $\psi(0;\rho)=1$:

$$h_{k+1} \;=\; B\, h_k^{\gamma(\rho)}\, t_k^{\alpha}\, \psi(A;\rho). \tag{17}$$

Here $B>0$ is the training productivity. The augmentation factor takes the one-parameter form $\psi(A;\rho)=(1+A)^{\eta(\rho)}$ with $\eta(\rho)=\eta_{\max}\rho$, giving the reach $\nu(A;\rho)=\eta(\rho)/(1+A)$ (Assumption 5). The reach is the semi-elasticity of transmission in AI, so its decline in A makes transmission log-concave in A: the proportional lift from an added unit of AI diminishes as the augmentation saturates. The parameter $\eta_{\max}$ sets how steeply the reach rises with expressibility and is disciplined by the novice-productivity evidence (Online Appendix E).

Production takes the CES form

$$\Phi(h_k,A) \;=\; \left[\,\theta_h\, h_k^{\sigma} \;+\; \theta_A\,(\underline{A}+A)^{\sigma}\,\right]^{1/\sigma}, \qquad \theta_h,\theta_A\geq 0, \quad \theta_h+\theta_A=1, \tag{18}$$

where the exponent σ indexes substitutability and $\sigma\in(0,1)$ places the two in the gross-substitutes region. The elasticity of substitution between the stock and AI is $1/(1-\sigma)$. The floor $\underline{A}$ is the codified tooling that production already runs on—software, references, procedures—so AI capability adds to an installed base. It is the production-side counterpart of the unit shift in the augmentation factor, and, like that shift, it bounds the marginal product of the first unit of AI capability. Under gross substitutes the aggregator is non-essential: AI alone yields positive output, $\Phi(0,A)>0$, so aggregate output rises in A even if the stock vanishes.[17] The gross-complements region $\sigma\leq 0$, where output can collapse with the stock, is treated in Online Appendix G.

The Cobb–Douglas core sets the learning-by-doing elasticity $\epsilon=\partial\ln H/\partial\ln t$ of Section 3.4 to the constant α; the elasticities $\gamma(\rho)$ and α are distinct economic primitives. The persistence $\gamma(\rho)$ measures how faithfully an expert's stock carries into the next genera-

[17] With $\sigma\in(0,1)$ the human-capital term vanishes at $h=0$ while the AI term survives: $\Phi(0,A)=\theta_A^{1/\sigma}(\underline{A}+A)>0$. Since $\theta_h h^{\sigma}\geq 0$, moreover, $\Phi(h,A)\geq\theta_A^{1/\sigma}(\underline{A}+A)$ for every h, so output is bounded below by a term that increases without bound in A; even in the eroded limit $h^{*}\to 0$ it stays positive. For $\sigma\leq 0$, by contrast, the aggregator is essential, $\Phi(0,A)=0$ (the Inada condition), so a vanishing stock collapses output.

tion; the learning-by-doing elasticity α measures how much an extra hour of co-production adds to the novice's. The two are pinned separately: $\gamma(\rho)$ by the persistence evidence, α by assumption. After the expert optimizes, the net intergenerational transmission is $(\gamma(\rho) - \alpha s_h)/(1-\alpha)$, where $s_h \equiv \partial \ln \Phi / \partial \ln h$.[18] It stays below one at the calibrated parameters, so the transmission chain converges monotonically to a unique stable steady state (Lemma 2).

The appropriation technology specializes to the linear $T(h) = \beta(\rho)\, h$, so the marginal appropriation $T_h = \beta(\rho)$ is constant in the stock. Like $\gamma(\rho)$ and $\eta(\rho)$, the appropriation coefficient $\beta(\rho)$ varies with expressibility; Section 6.2 calibrates the three ρ-functions together.

The expert's optimal co-production time takes the closed form $t^* = (\alpha \mathcal{W}_k / \Phi_k)^{1/(1-\alpha)}$ with $\mathcal{W}_k = \beta(\rho) B h_k^{\gamma(\rho)} \psi(A;\rho)$, and $\mu \equiv -d \ln t^* / dA$ as before. The steady-state stock and the comparative static in A of (5) specialize to

$$h^*(A;\rho) = \big[B (t^*)^{\alpha} \psi(A;\rho)\big]^{1/(1-\gamma(\rho))}, \qquad -\frac{dh^*(A;\rho)}{dA} = \frac{h^*(A;\rho)}{1-\gamma(\rho)} \big(\alpha\mu - \nu(A;\rho)\big), \tag{19}$$

and the shadow value (9), with the persistence shown explicitly, is

$$V'(A;\rho) \;=\; \frac{1}{1-\delta\gamma(\rho)} \cdot \left[(1-t^*)\, \frac{\partial \Phi}{\partial h}(A;\rho) \;+\; \gamma(\rho)\, \beta(\rho)\right], \tag{20}$$

with $\partial \Phi / \partial h(A;\rho) = \theta_h\, h^*(A;\rho)^{\sigma-1} (\Phi^*)^{1-\sigma}$ in the CES form, where $\Phi^* \equiv \Phi(h^*(A;\rho), A)$ is steady-state production. The Pigouvian schedule of Proposition 6 is then

$$\tau^*(A;\rho) \;=\; \delta\, V'(A;\rho) \Big(-\frac{dh^*(A;\rho)}{dA}\Big) \;=\; \underbrace{\frac{\delta\, V'(A;\rho)\, h^*(A;\rho)}{1-\gamma(\rho)}}_{>0} \cdot \big(\alpha\mu - \nu(A;\rho)\big), \tag{21}$$

whose force-free cut is the parametric form of (16),

$$\nu(A; \bar\rho(A)) \;=\; \alpha\mu, \tag{22}$$

with the learning-by-doing elasticity now the explicit parameter α. For $\rho < \bar\rho(A)$ the index is positive, $\tau^* > 0$, the tax direction; for $\rho > \bar\rho(A)$ it is negative, $\tau^* < 0$, the subsidy direction. The cut moves with whatever enters the contest $\alpha\mu - \nu$: the elasticity α and the augmentation slope $\eta_{\max}$ directly, and the persistence $\gamma(\rho)$, the production technology, and

[18] The net transmission is the elasticity of the equilibrium map $F(h) \equiv H(h, t^*(h;A), A)$ at the steady state, with the optimal time folded in. By the chain rule, $d\ln F / d\ln h = \gamma + \alpha\,(d\ln t^*/d\ln h)$, the coefficients being the elasticities of H in h and t. Since $t^* \propto (h^\gamma/\Phi)^{1/(1-\alpha)}$ gives $d\ln t^*/d\ln h = (\gamma - s_h)/(1-\alpha)$, substituting and canceling the $\alpha\gamma$ terms yields $(\gamma - \alpha s_h)/(1-\alpha)$.

the AI level through the compression μ.

6.2 Data and Calibration

We calibrate the model to the full cross-section of 682 U.S. occupations. Each occupation j enters with three measured quantities. The expressibility ρ_j is proxied by a routine-cognitive task index built from O*NET task content in the tradition of the task measures of Acemoglu and Autor (2011). It is high where work is codifiable and rule-based (court reporters, data-entry keyers, payroll clerks), low where it is tacit (counselors, the crafts, the performing arts). The exposure ω_j is the AI Occupational Exposure (AIOE) index of Felten et al. (2021). Both are rank-normalized to the unit interval. The employment weight n_j is national employment from the Bureau of Labor Statistics' Occupational Employment and Wage Statistics (BLS OEWS).

The two measures are nearly independent. Their rank correlation is 0.06—expressibility explains less than half a percent of the variance in exposure—and they scatter across the unit square with no discernible trend (Figure 3). This is not the pattern automation would leave. Routine-biased automation fell on codifiable, rule-based work, so an exposure measure built on it would rise with expressibility and trace the diagonal of the square. AI exposure is instead flat in expressibility: it reaches tacit work—counseling, teaching, the crafts—as readily as codifiable work. That AI exposure tracks the cognitive content of a job rather than its routineness is a pattern the exposure literature has recorded from its outset (Brynjolfsson et al., 2018; Webb, 2020; Eloundou et al., 2024). For this paper it is the empirical signature of the central claim of Section 2: AI acts on a different margin than automation.

The independence is what separates AI's two forces in the data. Exposure and expressibility are the empirical proxies for the two roles of Section 2: how much of the work AI can do, and how much of it AI can teach. The model imposes no relation between them, so their near-zero correlation is evidence, not assumption, that doing and teaching are distinct capacities. The orthogonality is foreshadowed in the exposure literature (Felten et al., 2021; Webb, 2020); its structural reading is what the model supplies.

In the model the two play separate parts: exposure sets the scale, entering as the effective AI level $A_j = \omega_j A$ atop the common tooling floor $\underline{A}$ (Online Appendix D). Expressibility sets the sign, whether AI erodes formation or augments it. The division is robust to relaxation: Online Appendix H lets expressibility reach production directly, tilting the substitution parameter with ρ in both directions, and the sign reversal and its single interior crossing carry through.

Because the two axes vary independently, the cross-section populates the corner that

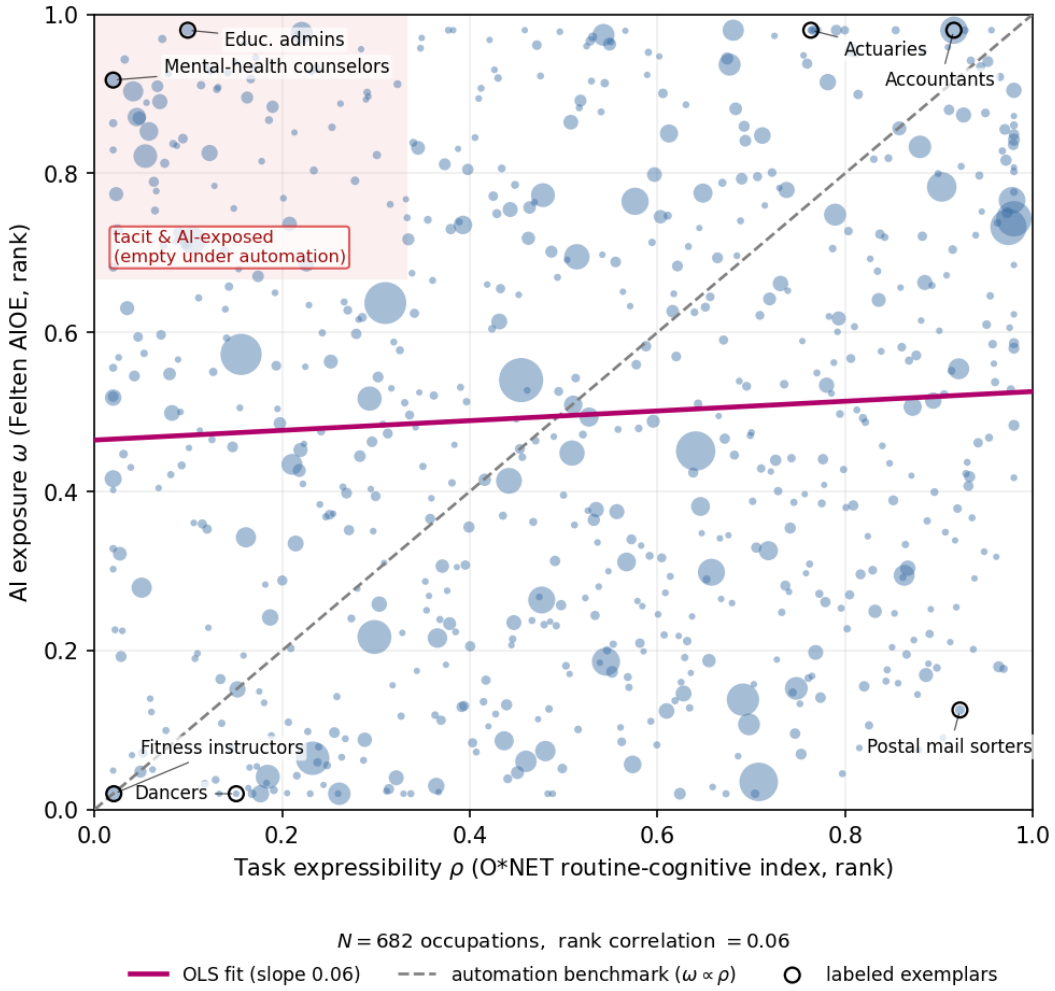


Figure 3: Task Expressibility and AI Exposure across Occupations

Note: Each point is one of the 682 occupations: task expressibility ρ (O*NET routine-cognitive index, rank) against AI exposure ω (Felten AIOE, rank), with point size proportional to employment (BLS OEWS, 2025). The two are nearly independent: their rank correlation is 0.06, and the OLS fit (solid) has slope 0.06 (s.e. 0.04, $R^2 < 0.01$), far from the diagonal (dashed) that exposure would trace if AI reached the same work as routine-biased automation. The decisive region is the top-left corner of tacit yet AI-exposed work, which the automation benchmark would leave empty but which AI fills.

decides the result—tacit yet strongly exposed work, where AI does the job but cannot codify the knowledge back. Automation's diagonal leaves that corner empty; here it is full, and its depths carry the schedule's largest positive values, the heaviest taxes an instrument aligned with the schedule would levy. Cut at terciles—expressibility in the bottom third, exposure in the top third—the corner holds 72 occupations and 9.03 percent of employment, schoolteachers its largest member.

Table 1: Calibrated Parameters

Parameter	Description	Baseline	Provenance	Source
$\gamma(\rho)$	Transmission persistence	$0.85 - 0.40\rho$	Illustrative	Qualitative: Cunha and Heckman, 2007; Cunha et al., 2010; Ericsson et al., 1993
$\eta(\rho)$	Augmentation slope	$\eta(\rho) = \rho$	Anchored	Brynjolfsson et al., 2025
$\beta(\rho)$	Appropriation coefficient	$0.80 - 0.50\rho$	Illustrative	Co-production bundle (§2)
α	Training-time elasticity	0.30	Assumed	—
σ	CES exponent	0.30	Assumed	Sweep spans Krusell et al., 2000; Acemoglu and Restrepo, 2022
(θ_h, θ_A)	Factor shares	$(0.85, 0.15)$	Assumed	—
δ	Intergenerational discount	0.40	Conventional	Drupp et al., 2018
A_0	Current AI level	0.50	Assumed	—
$\underline{A}$	Codified-tooling floor	0.05	Assumed	—

Note: Common parameters; the occupation-varying inputs ρ_j (O*NET), ω_j (Felten AIOE), and n_j (BLS OEWS) are described in the text. The augmentation slope matches the 36 percent novice-productivity gain and the persistence schedule follows the mastery-time proxy, both derived in Online Appendix E; "Qualitative" marks citations that discipline the schedule's qualitative claims rather than its numbers. The exponent σ implies an elasticity of substitution $1/(1-\sigma) \approx 1.4$ at baseline, the discount $\delta = 0.40$ corresponds to roughly 3% annually over a 30-year generation, and every row is probed in Online Appendix H.

The three ρ-functions are common across occupations and linear, each fixed by its two poles. Persistence falls from $\gamma(0) = 0.85$ to $\gamma(1) = 0.45$: expert quality carries across cohorts more faithfully where knowledge is tacit. The tacit pole is held at the stability ceiling, consistent with the skill-persistence evidence; the pole-to-pole gap is matched to the spread in time-to-mastery. The augmentation slope rises from $\eta(0) = 0$, the Polanyi floor—what cannot be told, AI cannot teach—to $\eta(1) = 1$, pinned to the 36 percent novice-productivity gain AI's teaching channel delivers in the field. The appropriation coefficient falls from $\beta(0) = 0.80$ to $\beta(1) = 0.30$: tacit formation happens inside production, so the novice's doing carries a large contemporaneous contribution, while codifiable formation arrives partly apart from the doing and carries less. Online Appendix E gives each endpoint's derivation and

provenance.

A level choice sets the training productivity B_j so that each occupation's decentralized steady-state stock at A_0 equals one, $h^*(A_0) = 1$, making the reported index values and welfare gaps comparable across occupations.[19]

In the calibrated class the level choice turns the single-crossing hypothesis of Propositions 5 and 6 into a theorem. At the normalized decentralized steady state the sign of the stock response reduces to the sign of $\nu(\rho) - \alpha g$, with $g \equiv \partial \ln \Phi / \partial A$ read at the normalized stock and the occupation's effective AI level. That read makes g common to every occupation at a given exposure. The reach rises strictly with expressibility, so the contest crosses zero at most once (Proposition A.1, Online Appendix A). The condition $\nu = \alpha g$ is the force-free cut $\nu = \alpha\mu$ written through the fixed-stock compression $\hat{\mu} = (g - \nu)/(1 - \alpha)$, which coincides with μ at the cut, and reading it exposure by exposure delivers the cut curve $\bar{\rho}(\omega)$. The theorem covers the decentralized read at the normalized baseline; the corrected equilibrium relocates the crossing continuously, and Section 6.3 verifies single crossing there.

The cross-section results below are comparative statics, read at the current AI level $A_0 = 0.5$. An AI trajectory enters only the transitional exercises of Online Appendices J and L, which specify the path and its parameters. Online Appendix F details the solution algorithm.

Each parameter in Table 1 carries its own provenance: anchored, conventional, assumed, or illustrative. That provenance is the whole of the calibration's discipline. The occupational sign reversal, the tax–subsidy split, and the welfare gap that follow are outputs of the calibrated model rather than targets it was fit to.

6.3 The Occupational Damage Index and the Sign Reversal

We evaluate the schedule τ^* of (21) at each occupation's own expressibility and exposure (ρ_j, ω_j), at $A_0 = 0.5$ and the corrected steady state at which the index is read (Section 4.2). Figure 4 plots the 682 index values against expressibility, with AI exposure in color and employment in point size.

The schedule reverses sign across the expressibility axis, as Proposition 6 predicts: τ^* tracks the sign of $\alpha\mu - \nu(\rho_j)$, with the reach read at each occupation's AI level. It is positive at low expressibility, where the stock erodes, and negative at high, where AI augments it.

[19] The choice fixes the units in which τ^* is denominated. It is a calibration object rather than a pure normalization: under the CES aggregator the stock's level sets the production share s_h. Online Appendix E shows that re-targeting it moves the cut and the tax-side share within the ranges of the structural sweep. The persistence cap $\gamma \leq 0.85$, the stability ceiling, keeps the steady state well-behaved: a single stock with γ near one compounds explosively.

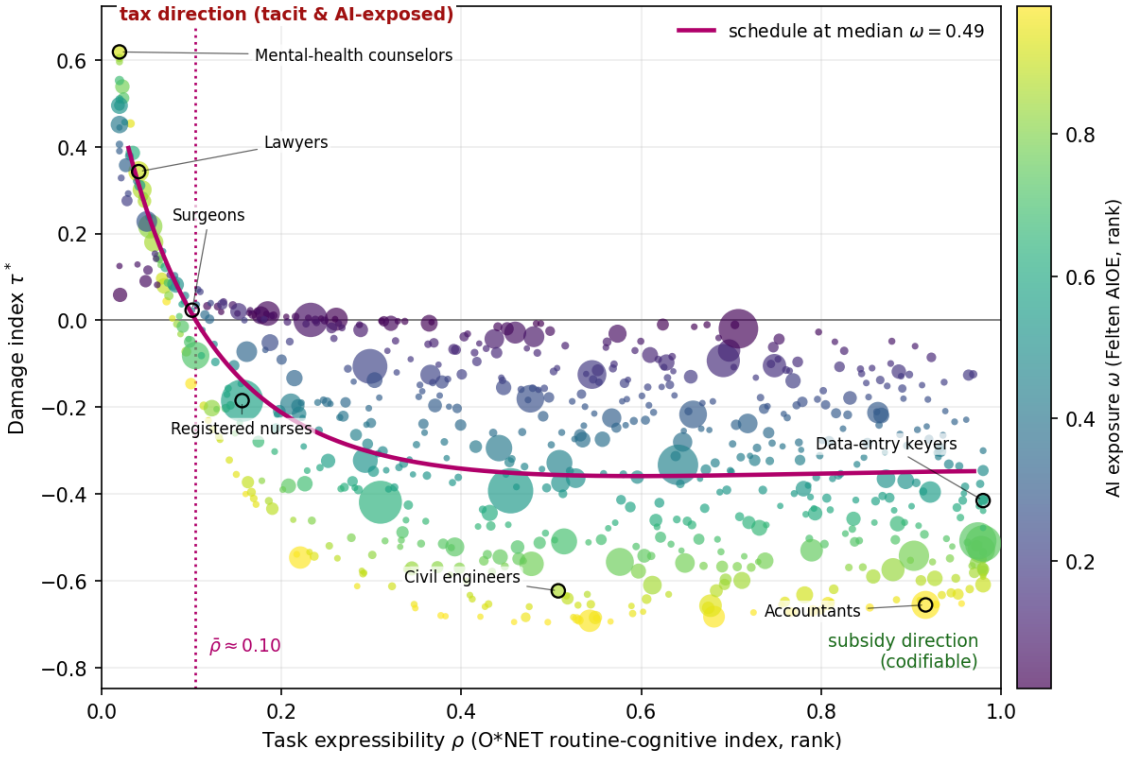


Figure 4: The Pigouvian Damage Index across the Expressibility Spectrum

Note: Each point is one of the 682 occupations: the Pigouvian schedule τ^* at $A_0 = 0.5$ against task expressibility ρ (O*NET routine-cognitive index, rank), colored by AI exposure ω (Felten AIOE, rank), with point size proportional to employment (BLS OEWS, 2025). The index is positive—the tax direction—for tacit occupations and negative—the subsidy direction—for codifiable ones. Because exposure varies independently of expressibility, the values form a cloud rather than a curve; the solid line is the schedule at median exposure, crossing zero at the force-free cut $\bar{\rho} \approx 0.10$ where $\nu(\bar{\rho}) = \alpha\mu$. Labels mark occupations chosen to span the spectrum.

The index is positive for 13.78 percent of occupations, carrying 13.16 percent of employment, and negative for the remainder.[20] The tax-side share is the calibration's least-pinned headline object: it ranges over $[7\%, 18\%]$ across the $\pm 30\%$ structural sweep, over $[0.4\%, 37\%]$ across the augmentation schedule's unidentified curvature, over $[7\%, 21\%]$ as production-side substitutability tilts with expressibility, over $[14\%, 18\%]$ as instruction bypasses the match, and falls to about 4% under cardinal measurement. The sign reversal and the cut survive

[20]The tax-side share is read at the corrected steady state (Section 4.2). The tax-side set is narrower than the decentralized erosion set because correcting the wedge raises co-production and the stock, shifting the contest and shrinking the set slightly. The erosion set—the occupations where the stock erodes under Proposition 1 and the external welfare effect of Lemma A.1 is accordingly negative—covers 17.30 percent of occupations and 15.70 percent of employment (Online Appendix H). The two sign sets are nested and differ for the 24 occupations (2.54 percent of employment) between them, which erode at the decentralized read yet sit on the subsidy side of the corrected index. The occupation and employment shares nearly coincide because occupation size is essentially uncorrelated with expressibility (rank correlation 0.03), so the tax-side set holds employment in proportion to its number.

every case (Online Appendix H).

The occupations where the index runs most positive are not the routine ones. They are the tacit occupations AI nonetheless reaches: mental-health counselors, clergy, arts and humanities professors, education administrators, healthcare social workers, actors. Their knowledge forms by doing rather than by prescription, yet they rank high on AI exposure, so AI compresses the co-production that forms the next cohort while codifying little of that knowledge back. This is the bundle breaking in the data: where AI does the work but cannot teach it, the corrective direction is a tax, protecting formation.

The erosion is silent over most of its range: output rises in over nine-tenths of employment even as stocks wear down beneath it, and the aggregate rises throughout. No economy-wide output statistic therefore flags the damage. Output itself turns down only in the corner's depths: in the most tacit, exposed occupations the stock's fall outweighs AI's direct gain even under gross substitutes (Online Appendix H). Thus where the output statistic finally speaks, the erosion is already deepest.

The index runs most negative in highly exposed analytical occupations of intermediate expressibility—epidemiologists, mathematicians, market analysts—where AI most strongly re-supplies the formation it touches. The naming is read off the task-content measure: acquisition-based alternatives relocate occupations along the axis while the reversal and the cut stand, so the naming, unlike the structure, is measure-specific (Online Appendix H).

At median exposure the schedule crosses zero once, near the tacit pole at $\bar{\rho} \approx 0.10$. Single crossing at the decentralized read is Proposition A.1, and the corrected equilibria preserve it at every exposure level in the calibration. Because each occupation meets the contest at its own effective level $\omega_j A_0$ atop the common floor, the cut is a curve across the exposure range, $\bar{\rho}(\omega)$. It falls from about 0.28 at the lowest exposures to about 0.08 at the highest, since the compression saturates in exposure while the reach keeps pace. Exposure also scales the magnitude: away from the curve, the more exposed the occupation, the more strongly signed its index (Figure 4).

The employment-weighted index is negative, a net subsidy in direction ($\bar{\tau} = -0.27$), yet the occupation values span $[-0.70, +0.62]$; the average is descriptive, and Section 6.5 prices actual uniform instruments on the adoption margin. A uniform AI instrument therefore carries one sign where the correction carries both. Set at the employment-weighted average, it would subsidize AI even as the tacit occupations carry the tax direction: seven-eighths of employment sits on the subsidy side of the cut. The error is one of sign, not magnitude—the central policy message of the paper, and the case for *expressibility-targeted* AI policy. It is a statement about the cross-section, distinct from the aggregate case against a blanket AI tax (Afrouzi et al., 2026; Growiec et al., 2026). Here the same instrument must reverse sign

across occupations in the same period, because AI's effect on formation is itself signed by codifiability.

The co-production subsidy $s^* = \delta V'/\beta$ is positive in every occupation—a 75 to 105 percent range across the middle four-fifths, employment-weighted mean 85 percent—while the AI correction reverses sign. The two correct different distortions. The intergenerational externality of Section 4.1 always depresses training, so its corrective subsidy is everywhere positive; the AI index prices the distinct AI-marginal distortion, which shifts the training incentive down where knowledge is tacit and up where it is codifiable. Whichever correction the planner holds alone, its schedule carries that geography; holding both, the planner corrects the wedge at its source and leaves AI unpriced (Corollary 1).

The correction at stake is large: the discounted planner-versus-decentralized welfare gap is positive in every occupation, an employment-weighted +67.5% (+22% to +183% across the persistence sweep, Online Appendix H). It is largest in the high-persistence tacit occupations, where the intergenerational externality compounds each generation's undertraining most heavily. The gap is measured with the novice paying nothing toward his own formation. Online Appendix B bounds how it falls when he pays all he can credibly commit: it drops to +10.6% and no further, because the generations beyond his lifetime lie outside any contract he can sign. About four-fifths of the gap is the baseline wedge that predates AI, the AI-marginal slice the remaining fifth (Online Appendix I).

6.4 Automation and the Value of Re-Supply

Automation, in the task-substitution tradition, is the polar case $\nu \equiv 0$ of Section 5.3: it displaces the doing with no reach into transmission. Historical automation also stopped at the codifiable boundary—its exposure would trace the diagonal of Figure 3—so the counterfactual is deliberately stronger: each occupation keeps AI's measured exposure, and only the teaching is removed. AI's defining departure is the reach $\nu > 0$, the codified route through which it re-supplies the formation it displaces, and the cross-section lets us price that departure. We hold each occupation's fundamentals fixed—training productivity, persistence, appropriation, exposure, and production—and replace AI's augmentation with automation's by setting the augmentation slope to zero, $\eta_{\max} = 0$. We then recompute the schedule and the stock response for all 682 occupations.

The contrast between the two regimes is stark (Figure 5). Under automation the contest $\alpha\mu - \nu$ collapses to $\alpha\mu > 0$ everywhere: automation raises production, compresses co-production, and erodes the stock in every occupation. The index is therefore positive in all 682 and the schedule never crosses zero. The entire subsidy region, the more codifiable 86.84

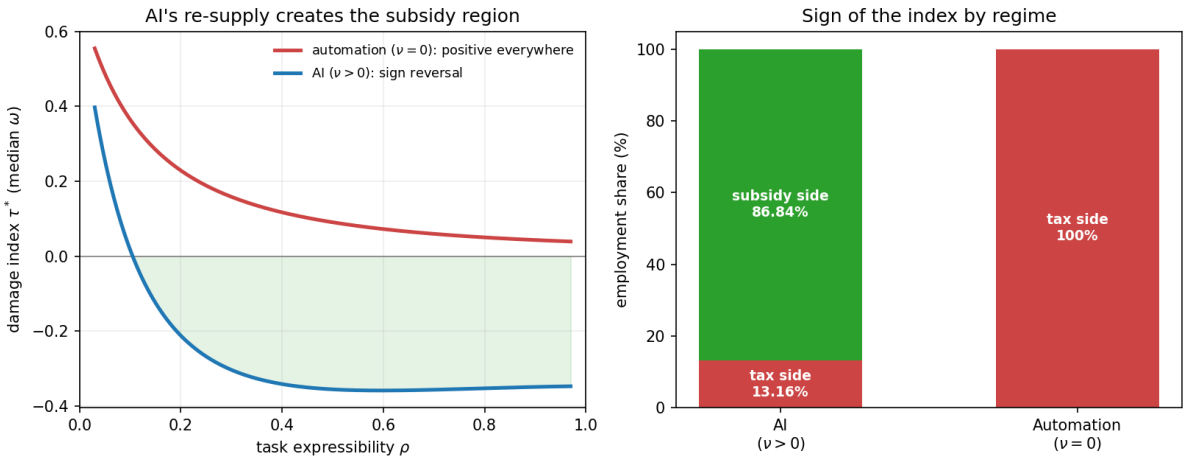


Figure 5: Automation versus AI

Note: Automation ($\nu = 0$) versus AI ($\nu > 0$) over the 682-occupation cross-section, holding each occupation's fundamentals fixed. Left: the damage-index schedule at median exposure; under AI it reverses sign (the subsidy direction in the shaded region), under automation it is positive at every expressibility. Right: the employment shares on the tax and subsidy sides—13.16/86.84 under AI, 100/0 under automation. The subsidy region exists only because AI re-supplies formation.

percent of employment under AI, is the creation of the re-supply channel alone. Re-supply also lifts the decentralized knowledge stock 34 percent above its automation counterpart and turns 84.30 percent of employment from erosion to augmentation, read at the decentralized steady state. The employment-weighted index itself changes sign: positive under automation, $\bar{\tau} = +0.13$, against negative under AI, $\bar{\tau} = -0.27$. Whether the AI correction is, on balance, a tax or a subsidy in direction thus turns entirely on the one feature that sets AI apart from automation—that where knowledge is codifiable, AI can teach.

6.5 The Welfare Cost of Mistargeted Policy

To this point the AI correction has been a shadow index; pricing mistargeted policy requires a setting in which AI use is actually taxed. We turn to the adoption margin of Section 4.3: each expert chooses AI use $a \in [0, \omega_j A]$ at resource price p, and the planner levies a per-unit AI rate, rebated lump-sum, chosen from a rate grid extended by outright prohibition. At the baseline price $p = 0.8$, an illustrative value at which the use margin is active, adoption is interior for most employment and both corners are populated.

Welfare is the discounted steady-state social value under the allocation each policy induces: production and appropriated contribution net of the AI resource cost, employment-weighted. Gains are relative to no policy, expressed as shares of the first-best gain. The benchmark is the occupation-targeted schedule. Against it we price, in turn, the policies a planner can set without observing expressibility: the single uniform rate, ten brackets on AI

exposure, and the automation model's occupation-specific rates. Last, we price two brackets on expressibility, the right axis at its coarsest (Figure 6). Every policy is a constant rate per occupation, optimized on the grid; none is the Markov schedule of Proposition 4.

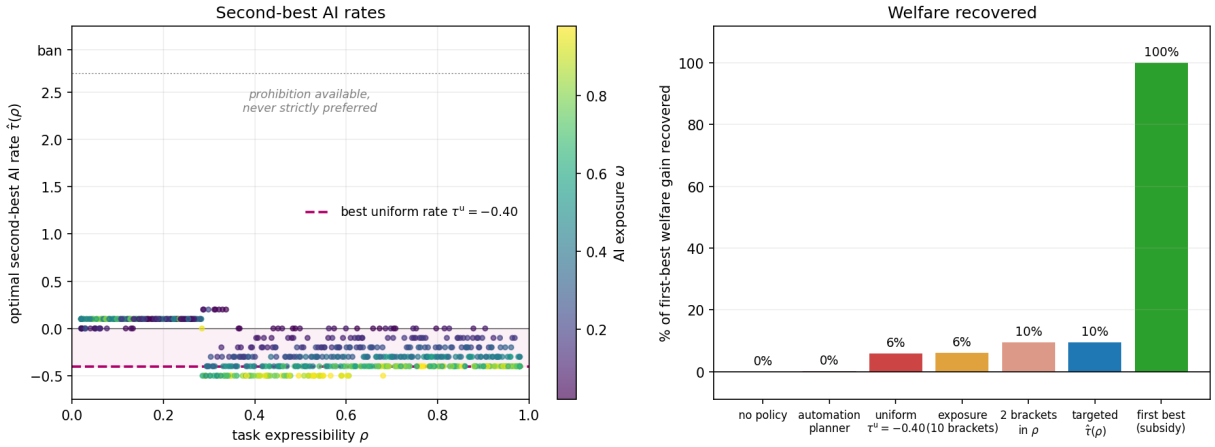


Figure 6: Uniform versus Targeted AI Policy in the Adoption Model

Note: Each point in the left panel is one of the 682 occupations: the occupation-optimal second-best AI rate at the baseline price $p = 0.8$ (Section 4.3), chosen over a rate grid on $[-2.5, 2.5]$ extended by prohibition, colored by exposure. Prohibition is available and never strictly preferred: the tooling floor bounds the first unit's marginal product, so finite prices implement exactly zero use, and each taxed occupation's rate is the smallest on the grid supporting its zero-use corner—welfare-equivalent to prohibition, ties resolved to the finite rate. Subsidies are reported as the smallest rate reaching full adoption, beyond which welfare is flat; zero rates mark occupations already at a corner without intervention. Right: welfare recovered as a share of the first-best (co-production subsidy) gain, shown for the automation model's rates, the best uniform rate, ten exposure-decile brackets, two expressibility brackets, and the occupation-targeted policy. The wrong-axis policies recover from essentially nothing to about three-fifths of what targeting delivers; two brackets on expressibility recover essentially all of it. The AI rate is second-best throughout—the first-best correction is the co-production subsidy. The welfare bars are the steady-state criterion; the transition-inclusive counterparts, which turn the steady-state-sized AI-price policies' gains negative (the automation model's near-zero rates excepted), are reported in Online Appendix K together with the transition-optimal rates.

The targeted optimum is a partition with supporting rates; the sign pattern again reverses across expressibility (Figure 6, left panel). Taxes fall where use erodes the stock, and every taxed occupation is driven to exactly zero use. By the tooling floor of (18), a finite price implements the allocation prohibition would pick (Proposition 4(iv)). Elsewhere either a subsidy drives use to the adoption frontier or no rate is needed where a corner already binds. The taxed side covers 25.95 percent of employment, wider than the shadow schedule's 13.16: the use-margin condition $\nu = \alpha\mu_a$ is read at each occupation's chosen use level rather than at the frontier, and it crosses at higher expressibility. The targeted schedule recovers just under a tenth of the first-best gain. The policy content is the partition—which occupations are priced out of use, which are pushed to the frontier—and the axis on which it is drawn.

The uniform rate comes first. The welfare-maximizing single rate follows the codifiable majority of Section 6.3: it is a subsidy, $\tau^{\mathrm{u}} = -0.40$, half the baseline AI price ad valorem. It recovers 5.8 percent of the first-best gain, three-fifths of what targeting delivers, and it harms the employment it mis-signs: 24.96 percent, whose welfare falls below its no-policy level. A blanket policy forgoes the gain in one definite way: it pushes one side of the force-free cut in the wrong direction.

Finer targeting on the wrong axis does not close the gap. Ten brackets keyed to AI exposure recover 6.1 percent of the first-best gain, scarcely more than the single rate, and still mis-sign 27.20 percent of employment. The reason is the independence documented in Figure 3: expressibility varies nearly as widely within each exposure decile as it does economy-wide, so every bracket inherits the sign conflict the uniform policy faces.

The automation model carries mistargeting to its limit: occupation-specific rates derived from a model in which AI only erodes ($\nu = 0$). Its damage direction is a tax everywhere, but the rates it levies move few margins. Priced in the true economy, the policy recovers essentially none of the attainable gain: 0.1 percent of the first-best gain, less than the single uniform rate. It recovers so little because it mis-signs 88.80 percent of employment, leaving the codifiable majority at zero or small positive rates where the optimum subsidizes. The wrong axis, free to discriminate, is worth less than the right sign applied uniformly.

Two brackets on expressibility suffice. A single split near the use-margin cut, one policy on each side, recovers the targeted gain to within rounding, and a third bracket adds nothing (Figure 6, right panel). The fine shape of the schedule carries little welfare; the sign carries almost all of it.

The AI rate is, throughout, a second-best instrument. The tenth it recovers under full targeting is its ceiling; the co-production subsidy of Section 4.3, which corrects the distortion at its source, recovers the whole first-best gain (Figure 6). Read as that section's two distances, the tenth is the value of internalization and the rest the value of contractibility. This is the targeting principle of Corollary 1: the co-production margin dominates where it can be instrumented, and the AI rate earns its place only where co-production cannot be contracted on.

The steady-state criterion is the AI rate's best case. Re-evaluated through the transition at the calibrated generational discount, the steady-state-sized policies lose welfare (the automation model's near-zero rates excepted). The transition-optimal rates keep the geography and halve the size, and they recover a small positive gain where the full-sized schedule loses. The co-production subsidy keeps a first-order gain under either criterion. Online Appendix K reports the allocation details, the price robustness of the ranking, and the full transition analysis. The AI rate's case is long-run; its irreplaceable content, on either

horizon, is the sign.

7 Conclusion

Artificial intelligence acts on the transmission of skill through two opposing forces: it displaces the doing through which novices learn, and, where knowledge can be put into words, it teaches directly. This paper has studied their balance in an overlapping-generations model in which each occupation holds one body of knowledge, indexed by its expressibility. Because the training contract rewards an expert only over the generation she shares with her novice, transmission is underprovided even before AI arrives. The externality precedes AI: expressibility sets which way AI cuts, and with it the sign of the correction.

Two results organize the analysis. The steady-state stock falls with AI where knowledge is tacit and rises where it is codifiable, the two regimes split by a single interior crossover (Proposition 5). The Pigouvian correction on AI reverses sign at a force-free cut, running in the tax direction where the stock erodes and in the subsidy direction where it is augmented: tax the doing, subsidize the telling. The reversal follows from a contest monotone in expressibility with one interior sign change, its tacit endpoint anchored by the Polanyi floor (Proposition 6).

In the calibration to the cross-section of 682 U.S. occupations, the cut falls near the tacit pole, with about an eighth of employment on the tax side and the codifiable seven-eighths on the subsidy side. Because exposure is nearly independent of expressibility, the heaviest tax-side values spare routine work and fall on tacit yet exposed occupations: counselors and clergy, as the baseline measure names them.

The policy lesson is one of sign rather than size: no uniform AI tax can be right, because the correction takes opposite signs on the two sides of the cut. Policies that cannot see expressibility forfeit most of what targeting delivers; two expressibility brackets recover the gain in full. The first-best instrument is not the AI rate but the co-production subsidy, which corrects the wedge where it arises.

The impulse to restrain AI in defense of skill formation is therefore sound only in the tacit, exposed occupations. For the codifiable majority, AI is on net a formation technology, and the correction points in the subsidy direction. The correction is also front-loaded: in the calibration, AI's erosion is concentrated in the rollout and recedes as AI matures, and the gain from investing in AI's teaching capacity falls with the intervention date, so early action dominates late (Online Appendix L). What targeting buys is the sign, and only the codifiability axis carries it.

The terrain to watch is the corner AI alone could open: work it does but cannot teach,

schoolteaching its largest occupation. In that corner the stock thins beneath output that mostly still improves, and in the corner's depths output itself has begun to fall. There, the telling being impossible to build, what policy protects is the doing through which the knowledge lives.

References

Acemoglu, D. and D. Autor (2011). Skills, tasks and technologies: Implications for employment and earnings. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 4B, pp. 1043–1171. Amsterdam: Elsevier.

Acemoglu, D., D. Kong, and A. Ozdaglar (2026). AI, human cognition and knowledge collapse. NBER Working Paper 34910, National Bureau of Economic Research.

Acemoglu, D., A. Manera, and P. Restrepo (2020). Does the US tax code favor automation? *Brookings Papers on Economic Activity Spring*, 231–285.

Acemoglu, D. and J.-S. Pischke (1998). Why do firms train? Theory and evidence. *Quarterly Journal of Economics 113*(1), 79–119.

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*(6), 1488–1542.

Acemoglu, D. and P. Restrepo (2020). Robots and jobs: Evidence from us labor markets. *Journal of Political Economy 128*(6), 2188–2244.

Acemoglu, D. and P. Restrepo (2022). Tasks, automation, and the rise in u.s. wage inequality. *Econometrica 90*(5), 1973–2016.

Afrouzi, H., A. Blanco, A. Drenik, and E. Hurst (2026). Automation, learning, and career dynamics. NBER Working Paper 35157, National Bureau of Economic Research.

Arrow, K. J. (1962). The economic implications of learning by doing. *Review of Economic Studies 29*(3), 155–173.

Autor, D. and N. Thompson (2025). Expertise. NBER Working Paper 33941, National Bureau of Economic Research.

Autor, D. H. (2014). Polanyi's paradox and the shape of employment growth. NBER Working Paper 20485, National Bureau of Economic Research.

Beane, M. (2019). Shadow learning: Building robotic surgical skill when approved means fail. *Administrative Science Quarterly 64*(1), 87–123.

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy 70*(5, Part 2), 9–49.

Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *Journal of Political Economy 75*(4), 352–365.

Benveniste, L. M. and J. A. Scheinkman (1979). On the differentiability of the value function in dynamic models of economics. *Econometrica 47*(3), 727–732.

Bhagwati, J. and V. K. Ramaswami (1963). Domestic distortions, tariffs and the theory of optimum subsidy. *Journal of Political Economy 71*(1), 44–50.

Brynjolfsson, E., B. Chandar, and R. Chen (2025). Canaries in the coal mine? six facts about the recent employment effects of artificial intelligence. Technical report, Stanford Digital Economy Lab working paper.

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

Brynjolfsson, E., T. Mitchell, and D. Rock (2018). What can machines learn, and what does it mean for occupations and the economy? *AEA Papers and Proceedings 108*, 43–47.

Costinot, A. and I. Werning (2023). Robots, trade, and luddism: A sufficient statistic approach to optimal technology regulation. *Review of Economic Studies 90*(5), 2261–2291.

Cunha, F. and J. J. Heckman (2007). The technology of skill formation. *American Economic Review 97*(2), 31–47.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica 78*(3), 883–931.

de la Croix, D., M. Doepke, and J. Mokyr (2018). Clans, guilds, and markets: Apprenticeship institutions and growth in the preindustrial economy. *Quarterly Journal of Economics 133*(1), 1–70.

Deming, D. J. (2017). The growing importance of social skills in the labor market. *Quarterly Journal of Economics 132*(4), 1593–1640.

Diamond, P. A. (1965). National debt in a neoclassical growth model. *American Economic Review 55*(5), 1126–1150.

Drupp, M. A., M. C. Freeman, B. Groom, and F. Nesje (2018). Discounting disentangled. *American Economic Journal: Economic Policy 10*(4), 109–134.

Eloundou, T., S. Manning, P. Mishkin, and D. Rock (2024). GPTs are GPTs: Labor market impact potential of LLMs. *Science 384*(6702), 1306–1308.

Emanuel, N., E. Harrington, and A. Pallais (2023). The power of proximity to coworkers. NBER Working Paper 31880, National Bureau of Economic Research.

Ericsson, K. A., R. T. Krampe, and C. Tesch-Römer (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review 100*(3), 363–406.

Felten, E., M. Raj, and R. Seamans (2021). Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal 42*(12), 2195–2217.

Garicano, L. (2000). Hierarchies and the organization of knowledge in production. *Journal of Political Economy 108*(5), 874–904.

Garicano, L. and L. Rayo (2026). Training in the age of AI: A theory of career viability. CEPR Discussion Paper 20634.

Growiec, J., K. Prettner, and M. Szkróbka (2026). Workers' incentives and the optimal taxation of AI. *Economics Letters 266*, 113062.

Guerreiro, J. a., S. Rebelo, and P. Teles (2022). Should robots be taxed? *Review of Economic Studies 89*(1), 279–311.

Ide, E. (2026). Automation, AI, and the intergenerational transmission of knowledge. Working paper, IESE Business School. arXiv:2507.16078.

Jarosch, G., E. Oberfield, and E. Rossi-Hansberg (2021). Learning from coworkers. *Econometrica 89*(2), 647–676.

Jovanovic, B. and Y. Nyarko (1996). Learning by doing and the choice of technology. *Econometrica 64*(6), 1299–1310.

Kim, K., N. Mester, and G. Sun (2026). AI and the labor market: A worker's-eye view. Working paper, Washington University in St. Louis. SSRN 6240619.

Korinek, A. and J. E. Stiglitz (2019). Artificial intelligence and its implications for income distribution and unemployment. In A. Agrawal, J. Gans, and A. Goldfarb (Eds.), *The Economics of Artificial Intelligence: An Agenda*, pp. 349–390. Chicago: University of Chicago Press.

Krusell, P., L. E. Ohanian, J.-V. Ríos-Rull, and G. L. Violante (2000). Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica 68*(5), 1029–1053.

Lazear, E. P. (1979). Why is there mandatory retirement? *Journal of Political Economy 87*(6), 1261–1284.

Lipsey, R. G. and K. Lancaster (1956). The general theory of second best. *Review of Economic Studies 24*(1), 11–32.

Ma, X., A. Nakab, and D. Vidart (2026). How do workers learn? theory and evidence on the roots of lifecycle human capital accumulation. NBER Working Paper 35199, National Bureau of Economic Research.

Milgrom, P. and I. Segal (2002). Envelope theorems for arbitrary choice sets. *Econometrica 70*(2), 583–601.

Nelson, R. R. and S. G. Winter (1982). *An Evolutionary Theory of Economic Change*. Cambridge, MA: Belknap Press of Harvard University Press.

Pigou, A. C. (1920). *The Economics of Welfare*. London: Macmillan.

Polanyi, M. (1958). *Personal Knowledge: Towards a Post-Critical Philosophy*. University of Chicago Press.

Polanyi, M. (1966). *The Tacit Dimension*. Garden City, NY: Doubleday.

Sandmo, A. (1975). Optimal taxation in the presence of externalities. *Swedish Journal of Economics 77*(1), 86–98.

Stokey, N. L., R. E. Lucas, Jr., and E. C. Prescott (1989). *Recursive Methods in Economic Dynamics*. Cambridge, MA: Harvard University Press.

Thuemmel, U. (2023). Optimal taxation of robots. *Journal of the European Economic Association 21*(3), 1154–1190.

Webb, M. (2020). The impact of artificial intelligence on the labor market. Technical report, Stanford University working paper.

Wolter, S. C. and P. Ryan (2011). Apprenticeship. In E. A. Hanushek, S. Machin, and L. Woessmann (Eds.), *Handbook of the Economics of Education*, Volume 3, pp. 521–576. Amsterdam: Elsevier.

Online Appendix to "Tax the Doing, Subsidize the Telling: Optimal AI Policy and the Codifiability of Skills"

Zhongchen Fan Ruichi Xiong

July 17, 2026

This Online Appendix contains the proofs of the results stated in the main text, together with the robustness analyses, extensions, data and numerical details, and secondary results of the paper. Equations, figures, tables, and propositions are numbered within sections (A.1, A.2, …); references to numbered objects of the main manuscript—propositions, equations, assumptions, sections—carry the main manuscript's numbering.

A Proofs

Proof of Lemma 1 (expert's optimal co-production time)

The expert's problem is a one-dimensional concave program in the co-production share t_k. The proof has three parts: write down the first-order condition, show the objective is concave, so the condition pins the unique optimum, and then differentiate the condition to sign how the optimum moves with AI.

The program. The expert maximizes

$$Y_k \;=\; (1-t_k)\,\Phi_k \;+\; T\big(H(h_k,t_k,A)\big), \qquad t_k \in [0,1],$$

where $\Phi_k = \Phi(h_k, A)$ is a constant from her point of view: the stock h_k is inherited and the AI level A is taken as given, so output on the AI share moves one-for-one with the time $1-t_k$ spent there but not with the choice itself. Only the second term, the appropriated contribution of the novice, bends with t_k, through the novice's stock $h_{k+1} = H(h_k, t_k, A)$ it builds.

First-order condition. Differentiating and setting the result to zero,

$$\frac{\partial Y_k}{\partial t_k} \;=\; -\Phi_k + T_h(h_{k+1})\, H_t \;=\; 0,$$

which is (3). Reading it as a balance: the marginal hour shifted into co-production forgoes the output Φ_k it would have produced with AI and earns in return the appropriated marginal product $T_h H_t$, and the optimum equates the two.

Concavity and uniqueness. The second derivative is negative throughout:

$$\frac{\partial^2 Y_k}{\partial t_k^2} \;=\; \underbrace{T_{hh}\, H_t^2}_{\leq 0} \;+\; \underbrace{T_h\, H_{tt}}_{<0} \;<\; 0.$$

The marginal return to co-production falls for two compounding reasons. First, each extra hour raises the novice's stock and so slides the expert down a diminishing appropriation schedule ($T_{hh} \leq 0$, T concave). Second, the hours themselves have diminishing product ($H_{tt} < 0$, H strictly concave in t_k, and $T_h > 0$ so this product is valued). These signs are Assumptions 3 and 2, respectively. Strict concavity makes the interior root of the first-order condition the unique maximum.[1] Because the program is twice continuously differentiable and this second derivative is strictly negative, the optimum is regular and the implicit function theorem makes $t^*(h_k; A)$ continuously differentiable—hence continuous, the property the equilibrium map F inherits in the proof of Lemma 2.

Response to AI. To sign $\partial t^*/\partial A$ we differentiate the first-order condition. It is cleanest in logs, because the result is a race between two growth rates. At an interior optimum $T_h, H_t, \Phi > 0$, so we may take logs of $T_h H_t = \Phi$ and differentiate totally in A at fixed h_k, letting $t_k = t^*$ respond and recalling that Φ does not depend on t_k:

$$\underbrace{\frac{\partial \ln(T_h H_t)}{\partial A}}_{\text{direct effect of AI}} \;+\; \underbrace{\frac{\partial \ln(T_h H_t)}{\partial t_k}}_{\text{second-order condition}} \frac{\partial t^*}{\partial A} \;=\; \underbrace{\frac{\partial \ln \Phi}{\partial A}}_{\text{AI's production gain}}.$$

The coefficient on the response is the second-order condition rewritten in logs,

$$\frac{\partial \ln(T_h H_t)}{\partial t_k} \;=\; \frac{T_{hh} H_t^2 + T_h H_{tt}}{T_h H_t} \;<\; 0,$$

the appropriated marginal return falling in co-production. Solving the differentiated condi-

[1] The optimum can sit at a corner. The upper corner $t^* = 1$ binds when co-production pays even at full intensity, $T_h H_t > \Phi_k$ for all $t_k \in [0, 1]$. The lower corner $t^* = 0$ is excluded whenever the very first hour pays, $\lim_{t_k \to 0} T_h H_t > \Phi_k$, which the co-production Inada condition $\lim_{t_k \to 0} H_t = \infty$ guarantees (the marginal appropriation T_h being strictly positive). We take the optimum interior in what follows.

tion for the response gives

$$\frac{\partial t^*}{\partial A} \;=\; \frac{\partial \ln \Phi/\partial A \;-\; \partial \ln(T_h H_t)/\partial A}{\partial \ln(T_h H_t)/\partial t_k},$$

whose denominator is negative. Hence $\partial t^*/\partial A$ carries the sign of its numerator reversed: AI compresses co-production, $\partial t^*/\partial A < 0$, exactly when its production gain outpaces its lift to the appropriated return, $\partial \ln \Phi/\partial A > \partial \ln(T_h H_t)/\partial A$. This is the *substitute regime*, the fixed-h counterpart of the compression (4); the reverse inequality is the *complement regime*. □

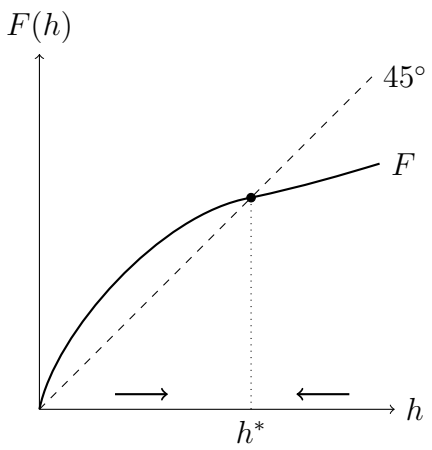
Proof of Lemma 2 (steady state)

The object of study is the one-dimensional dynamical system $h_{k+1} = F(h_k)$ obtained by folding the expert's optimal co-production into the transition—$F(h) = H(h, t^*(h; A), A)$. A steady state is a positive fixed point $h^* = F(h^*)$; its stability and uniqueness are read off the slope and the curvature of F. The three parts of the lemma map to a crossing argument, a slope condition, and a single-crossing argument; Figure A.1 shows the two configurations the argument distinguishes.

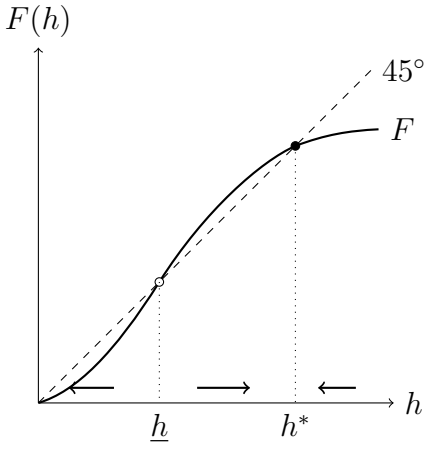
Existence (i). Two anchors bracket the map. At the bottom, $F(0) = 0$: with no inherited stock nothing transmits ($H(0, t, A) = 0$, Assumption 2), so the origin is always a fixed point, and F is continuous because t^* is (proof of Lemma 1). At the top, F eventually falls below the 45° line, in two steps. First, co-production is a time share, $t^* \leq 1$, and H rises in t, so the transition never exceeds its full-intensity value: $F(h) \leq H(h, 1, A)$. Second, $H(h, 1, A)$ grows more slowly than h. Its elasticity in h is the persistence $d \ln H(h, 1, A)/d \ln h$, so the log gap to the diagonal has derivative

$$\frac{d}{d \ln h} \ln \frac{H(h, 1, A)}{h} \;=\; \frac{d \ln H(h, 1, A)}{d \ln h} - 1 \;<\; 0$$

by sub-unit persistence (Assumption 2). The hypothesis of part (i) bounds the persistence away from one at large stocks, so this derivative is eventually below a strictly negative constant; the gap then falls without bound, $H(h, 1, A)/h \to 0$, and $H(h, 1, A) < h$ eventually. The two steps combine to $F(h) \leq H(h, 1, A) < h$ past some $\bar{h}$: diminishing returns to the inherited stock keep a high dynasty from reproducing its own height, and the map drops below the diagonal. So if the map ever rises strictly above the diagonal, $F(h_0) > h_0$ for some $h_0 > 0$, continuity forces it back down across the diagonal between h_0 and $\bar{h}$: the intermediate-value theorem delivers a positive fixed point in $(h_0, \bar{h}]$ (Figure A.1a). The crossing condition is exactly what rules out the trivial economy that collapses to the origin.



(a) A unique stable steady state.



(b) Collapse threshold below a stable steady state.

Figure A.1: Steady States of the Equilibrium Transition Map

Note: The equilibrium transition map $F(h) = H(h, t^*(h; A), A)$ against the 45° line; the origin is always a fixed point ($F(0) = 0$). (a) When F lies above the 45° line just past the origin, sub-unit persistence forces it back across exactly once, at a steady state h^* where it cuts the line from above ($0 < F'(h^*) < 1$). That steady state is the stable, attracting stock (existence (i) and stability (ii)). (b) When F instead starts below the line, the lower crossing $\underline{h}$ is an unstable up-crossing ($F' > 1$) marking the edge of the basin: dynasties starting below $\underline{h}$ collapse toward the origin, those above it converge to the stable h^*. A strictly decreasing transition elasticity (uniqueness (iii)) lets $F(h)/h$ cross one at most twice, so at most these two crossings occur and the stable steady state is unique. Arrows on the horizontal axis show the induced dynamics.

Stability (ii). A fixed point attracts nearby dynasties if and only if $|F'(h^*)| < 1$. The slope decomposes into two channels,

$$F'(h_k) \;=\; \underbrace{\gamma}_{\text{direct inheritance}} \;+\; \underbrace{H_t \,\frac{\partial t^*}{\partial h_k}}_{\text{induced co-production}}, \qquad \gamma = H_h \text{ at the fixed point.}$$

A richer expert transmits more directly, and she also adjusts how much she co-produces. Of the two sides of $|F'| < 1$, the upper one is what can bind. Sub-unit persistence already places the direct term below the unstable ceiling, $\gamma < 1$ (Assumption 2), leaving headroom $1 - \gamma$; the upper bound $F'(h^*) < 1$ then asks only that the induced term not spend more than that headroom—the *non-amplification* condition $H_t\, \partial t^*/\partial h_k < 1 - \gamma$. It is slack whenever co-production does not rise in the inherited stock ($\partial t^*/\partial h_k \leq 0$), and otherwise merely caps how fast co-production may rise; the remaining side $F' > -1$ fails only under an implausibly steep decline, $\partial t^*/\partial h_k < -(1+\gamma)/H_t$, which we set aside. With both met the slope lies in $(-1, 1)$ and the fixed point is locally exponentially stable; in the calibration $F'(h^*) \in (0, 1)$, so convergence is monotone—the map cuts the 45° line from above (Figure A.1).

Uniqueness (iii). It is enough that the map cross the diagonal cleanly. Track the log gap $g(\ln h) \equiv \ln\big(F(h)/h\big)$, whose derivative is $dg/d\ln h = \mathcal{E}_F(h) - 1$, with $\mathcal{E}_F(h) \equiv d\ln F/d\ln h$ the transition elasticity. When $\mathcal{E}_F$ is strictly decreasing, this derivative is strictly decreasing too, so g is strictly concave in $\ln h$; a strictly concave function attains any value, in particular zero, at most twice, so $F(h) = h$ has at most two positive roots. Two cases. If $F > h$ just above the origin, there is a single positive root and the map crosses from above ($F' < 1$)—the stable steady state, unique (Figure A.1a). If instead $F < h$ near the origin, the lower root is an up-crossing ($F' \geq 1$, unstable) and only the upper root satisfies $F'(h^*) < 1$; the stable steady state is again unique, with the unstable fixed point beneath it marking the edge of its basin of attraction (Figure A.1b).

Calibration. The parametric model of Section 6 satisfies all three conditions. Its CES production raises the human-capital elasticity of output as the stock grows, making the transition elasticity $\mathcal{E}_F$ strictly decreasing, so the uniqueness case (iii) applies; the stable slope lies inside $(0, 1)$ throughout the calibrated range. □

Proof of Proposition 1 (AI's impact on the equilibrium)

At the steady state H_h equals the persistence $\gamma < 1$ (Assumption 2), so $1 - H_h > 0$. The implicit function theorem then makes $h^*(A)$ continuously differentiable. Totally differentiating the steady-state identity $h^* = H(h^*, t^*, A)$ in A gives

$$(1 - H_h)\,\frac{dh^*}{dA} \;=\; H_A + H_t\,\frac{dt^*}{dA}.$$

Here dt^*/dA and the compression $\mu \equiv -d\ln t^*/dA$ are the *total* steady-state responses along the fixed point, not the partial $\partial t^*/\partial A$ at fixed h—they already fold in the adjustment of t^* induced by h^* moving with A. Substituting the steady-state values $H_A = \nu h^*$, $dt^*/dA = -\mu t^*$, $t^* H_t = \epsilon h^*$, and $H_h = \gamma$ yields

$$\frac{dh^*}{dA} \;=\; \frac{h^*\,(\nu - \epsilon\mu)}{1 - \gamma},$$

which is (5). □

Proofs of Proposition 2 and Lemma 3 (wedge and shadow value)

Both are immediate, and we record them for completeness. The wedge (8) is the planner's first-order condition (7) minus the expert's (3): the two differ only in the term $\delta V' H_t$, which is strictly positive because V increases in the stock and $H_t > 0$. The closed form (9) evaluates

the envelope recursion of Section 4.1 at a steady state, where $h_{k+1} = h_k$ and $H_h = \gamma$: the recursion becomes $V' = (1 - t^*)\Phi_h + (T_h + \delta V')\,\gamma$, and solving for V' gives (9), positive term by term with $1 - \delta\gamma > 0$. □

The exact externality of AI capability

The damage index of Proposition 3 is anchored to an exact welfare object, derived here. The experiment: at the decentralized steady state, raise A permanently and differentiate discounted social welfare along the decentralized path. That welfare, written W, is the same discounted flow the planner's value V maximizes, read instead along the path the experts steer; $W \leq V$, and the gap is what correction can recover. The welfare derivative splits into two parts. The expert already internalizes the first: her direct production gain and her appropriated share of the reach. The external part is the stream of uninternalized stock displacements, one per period, each valued at the social shadow value—located exactly where Assumption 4 put it, in the stock the expert does not own.

Lemma A.1 (Marginal externality of AI capability)**.** *Let $W(h_k)$ be discounted social welfare along the decentralized path from h_k, and W' its derivative, both read at the decentralized steady state. The derivative of W with respect to a permanent increment in the AI level A decomposes into terms the expert internalizes plus the external component*

$$E(A) \;=\; \frac{\delta\, W'}{1-\delta}\,\frac{\partial h_{k+1}}{\partial A}, \qquad\qquad \frac{\partial h_{k+1}}{\partial A} \;\equiv\; H_A + H_t\,\frac{\partial t^*}{\partial A}, \tag{A.1}$$

the impact response of transmission to AI at a fixed inherited stock, capitalized at the planner's discount and valued at the shadow price. Its sign is the sign of the stock response, $\operatorname{sign} E(A) = \operatorname{sign}\big(\nu(A) - \epsilon\mu\big) = \operatorname{sign} dh^*/dA$*: the external welfare effect is negative where AI erodes the stock and positive where AI augments it.*

Proof. Along the decentralized path from h_k, social welfare obeys the recursion $W(h_k; A) = Y_k + \delta\, W(h_{k+1}; A)$, with $Y_k = (1 - t^*)\Phi(h_k, A) + T(h_{k+1})$ and $h_{k+1} = H(h_k, t^*(h_k; A), A)$. Its stock derivative $W' \equiv \partial W/\partial h_k$ at the steady state is positive. Differentiating the recursion in h_k, the expert's condition (3) removes the flow part of the co-production response, leaving $W' = (1 - t^*)\Phi_h + T_h H_h + \delta\, W' F'$, with $F' = H_h + H_t\,\partial t^*/\partial h_k$ the slope of the equilibrium transition of Lemma 2. At the steady state $H_h = \gamma$, and solving gives $W' = \big[(1 - t^*)\Phi_h + T_h\,\gamma\big]\big/\big(1 - \delta\, F'(h^*)\big) > 0$: the numerator is positive, and the denominator is at least $1 - \delta$ by non-amplifying stability, $|F'(h^*)| < 1$. Now differentiate the recursion in A at the steady

state, where the same state and policy recur:

$$\frac{dW}{dA} \;=\; \underbrace{(1-t^*)\,\Phi_A + T_h H_A}_{\text{internalized}} \;+\; \Big(\underbrace{-\Phi + T_h H_t}_{=0 \text{ by (3)}} + \delta W' H_t\Big)\frac{\partial t^*}{\partial A} \;+\; \delta W' H_A \;+\; \delta\,\frac{dW}{dA}.$$

The expert's first-order condition zeroes the private part of the co-production response, and the first brace, AI's direct production gain plus the appropriated share of the reach, is what her own objective already carries. What remains is external. Collecting,

$$(1-\delta)\,\frac{dW}{dA}\bigg|_{\text{external}} \;=\; \delta W'\left(H_A + H_t\,\frac{\partial t^*}{\partial A}\right) \;=\; \delta W'\,\frac{\partial h_{k+1}}{\partial A},$$

which is (A.1). For the sign: differentiating the fixed point $h^* = H(h^*, t^*(h^*; A), A)$ gives $dh^*/dA = (\partial h_{k+1}/\partial A)\big/\big(1 - dH/dh_k\big)$ with a positive denominator by stability (Lemma 2), so $\partial h_{k+1}/\partial A$ and dh^*/dA share their sign, which is $\operatorname{sign}(\nu - \epsilon\mu)$ by Proposition 1; and $W' > 0$. □

The formula is the discounted transition in closed form. With F' the slope above, the displacement accumulates along the transition as $\Delta_k = (\partial h_{k+1}/\partial A)\big(1 - (F')^k\big)\big/(1 - F')$, climbing to dh^*/dA, and the externality is this accumulation's discounted valuation: $E(A) = \sum_{k\geq 1} \delta^k \big[(1-t^*)\Phi_h + T_h\gamma\big]\,\Delta_k$, each period's displacement valued at the stock's per-period marginal value, the numerator of (9). Summing the geometric series returns (A.1). As $\delta \to 1$, the annuitized externality $(1-\delta)E(A)$ converges to that value times dh^*/dA: the transition's weight vanishes and only the settled displacement is priced.

Proof of Proposition 3 (the Pigouvian damage index)

The index is the discounted marginal social damage of AI on the steady-state stock, $\tau^* = \delta V'\,(-dh^*/dA)$, read at the corrected steady state. There the operative first-order condition is the planner's (7), carrying the effective marginal appropriation $T_h + \delta V'$ rather than the bare T_h of (3). The displacement $-dh^*/dA$ is read at that corrected allocation, the correction $\delta V'$ held fixed. The formula and its sign follow by substitution; the force-free property needs more, and we prove it last.

Formula and sign (i). The comparative static of Proposition 1 applies at this steady state, and substituting $-dh^*/dA$ from (5) yields the rate (10). Its prefactor $\delta V' h^*/(1-\gamma)$ is strictly positive, since $V' > 0$, $h^* > 0$, and $\gamma < 1$, so $\operatorname{sign}\tau^* = \operatorname{sign}(\epsilon\mu - \nu) = -\operatorname{sign}(dh^*/dA)$, every term read at the corrected steady state: the index is positive, a damage, exactly where the stock response at that allocation is negative. At a common allocation this is the opposite orientation to the external welfare effect $E(A)$ of Lemma A.1, with the shared algebraic zero

$\nu = \epsilon\mu$; the lemma reads E at the decentralized steady state, a different equilibrium.

The force-free cut (ii). The rate vanishes where its second factor does, at $\nu = \epsilon\mu$; the force-free property is that this condition contains neither the appropriation T nor the shadow value. At the cut the stock does not respond to AI, $dh^*/dA = 0$, so the total and fixed-h responses of t^* coincide and the cut collapses to

$$H_A + H_t \frac{\partial t^*}{\partial A} = 0,$$

a balance between AI's direct reach into transmission (H_A) and the transmission lost as AI compresses co-production ($H_t\, \partial t^*/\partial A$).

The co-production response. Write $\tilde{T}_h$ for the operative marginal appropriation, T_h in the decentralized economy and $(1+s^*)\,T_h$ in the implemented one with s^* held fixed; one calculation then covers both, the steps below being invariant to this positive scaling. Differentiating the operative condition $\tilde{T}_h(h_{k+1})\, H_t = \Phi$, as in the proof of Lemma 1, gives

$$\frac{\partial t^*}{\partial A} \;=\; \frac{\Phi_A - \tilde{T}_{hh} H_A H_t - \tilde{T}_h H_{tA}}{\tilde{T}_{hh} H_t^2 + \tilde{T}_h\, H_{tt}}.$$

The cancellation. Substitute this response into the cut condition $H_A + H_t\, \partial t^*/\partial A = 0$ and clear the denominator $D \equiv \tilde{T}_{hh} H_t^2 + \tilde{T}_h H_{tt}$, giving $H_A D + H_t\big(\Phi_A - \tilde{T}_{hh} H_A H_t - \tilde{T}_h H_{tA}\big) = 0$. Expanding $H_A D$ and collecting terms,

$$\underbrace{\tilde{T}_{hh} H_A H_t^2}_{\text{from } H_A D} + \tilde{T}_h H_A H_{tt} \;+\; H_t \Phi_A \;\underbrace{-\,\tilde{T}_{hh} H_A H_t^2}_{\text{from } H_t\, \partial t^*/\partial A} - \tilde{T}_h H_t H_{tA} \;=\; 0,$$

so the two $\tilde{T}_{hh} H_A H_t^2$ terms cancel, leaving $\tilde{T}_h\big(H_A H_{tt} - H_t H_{tA}\big) + H_t \Phi_A = 0$. The appropriation now survives only through its slope $\tilde{T}_h$; imposing the optimality condition $\tilde{T}_h = \Phi/H_t$ and dividing by $\Phi H_{tt}/H_t$ removes even that,

$$H_A \;-\; \frac{H_t H_{tA}}{H_{tt}} \;+\; \frac{H_t^2\, \Phi_A}{\Phi\, H_{tt}} \;=\; 0,$$

free of T in level, slope, and curvature, and free of the shadow value, which entered $\tilde{T}_h$ only through the level of s^* and dropped out with the scaling. The elasticity form is read off the cut condition itself. Dividing $H_A + H_t\, \partial t^*/\partial A = 0$ by H turns the first term into the reach ν. Multiplying and dividing the second term by t^* then turns it into $-\epsilon\mu$. The compression here is the fixed-h response, which coincides with the total one at the cut. The cut is therefore $\nu = \epsilon\mu$, which proves (ii).

Why it cancels. The appropriation and the shadow value act on the cut only by relocating the equilibrium (h^*, t^*) at which it is read; they are never forces in the contest. The reason is the bundle: the expert's entire within-period return flows through the very stock h_{k+1} that the externality acts on. Once her optimality is imposed, that return is pinned to the production margin, $\tilde{T}_h = \Phi/H_t$, and leaves no separate handle on the margin where AI's two forces contend. The tax-side cut therefore *coincides* with the erosion cut of Proposition 1, both read at the same steady state. (The reading holds the supporting rate fixed; tracking the re-optimized planner path instead, with the rate adjusting as A varies, would add a valuation term through $\partial V'/\partial A$ that the implemented reading excludes by construction.) □

Proof of Corollary 1 (first-best instrument assignment)

The claim is that a single instrument does two jobs: the co-production subsidy that corrects the co-production margin also corrects the AI-use margin, so the optimal tax on AI use is zero. The reason is that both distortions are the *same* missing term, the continuation value $\delta V'$ the expert ignores, and the subsidy restores it wherever the novice's stock enters.

The subsidized expert. With a proportional subsidy s on the appropriation, the expert maximizes

$$(1 - t_k)\,\Phi(h_k, a) + (1 + s)\,T\big(H(h_k, t_k, a)\big) - p\,a,$$

choosing co-production and AI use, with first-order conditions on the two margins

$$-\Phi + (1 + s)\,T_h H_t = 0, \qquad (1 - t_k)\,\Phi_a + (1 + s)\,T_h H_a = p.$$

The right subsidy reproduces the planner. Set $s = s^* = \delta V'/T_h$, read at the corrected steady state, so the effective marginal appropriation becomes $(1+s^*)\,T_h = T_h + \delta V'$, precisely the planner's marginal valuation of the novice's stock. The planner, maximizing (13), has the two first-order conditions

$$-\Phi + (T_h + \delta V')\,H_t = 0, \qquad (1 - t_k)\,\Phi_a + (T_h + \delta V')\,H_a = p,$$

which become the subsidized expert's two conditions term for term once $(1+s^*)\,T_h$ is written as $T_h + \delta V'$: they coincide, margin for margin. The coincidence extends to corner allocations: the same effective marginal enters each margin's condition, so the inequalities that support a corner compare the same objects for expert and planner, and both select the same corners. The subsidy alone supports the planner's allocation, and nothing is levied on AI use. Off the steady state the ad valorem form is local: a rate $s(h_k) = \delta V'(h_{k+1})/T_h(h_{k+1})$, fixed within the period and evaluated at the planner's intended next state, aligns the expert's condition

at that point. Global implementation uses the payment schedule of Section 4.3: paying $S(h_{k+1})$ alongside the appropriation, with marginal payment $S'(h_{k+1}) = \delta V'(h_{k+1})$, makes the expert's effective marginal appropriation $T_h(h_{k+1}) + \delta V'(h_{k+1})$ at every candidate next state. The schedule is defined by its marginal: integrating, $S(h_{k+1}) = \delta\, V(h_{k+1}, A)$ up to a constant, so her objective is the planner's up to a constant. Her problem therefore coincides with the planner's along the whole corrected path, corners included.

Hence no AI tax. Since the subsidy already attains the planner's optimum, any nonzero price correction on use would move the use level away from it and lower welfare locally: $\tau^{\text{FB}}(A) = 0$. This is the principle of targeting: place the instrument on the distorted margin, co-production, and leave the other margin, use, alone. That margin is undistorted once the source is corrected. □

Proof of Proposition 4 (second-best AI pricing)

When co-production cannot be subsidized, the only lever is the price of AI use. The same friction that creates the wedge also makes co-production non-contractible. We find the welfare-maximizing price correction $\hat{\tau}$, show it prices the per-period damage AI does to the novice's stock, sign it, and locate its cut, which is force-free for the same reason as the shadow index's.

The expert's response. Facing the price $p + \tau$, the expert sets

$$G^t \equiv T_h(h_{k+1})H_t - \Phi = 0, \qquad G^a \equiv (1 - t_k)\Phi_a + T_h(h_{k+1})H_a - (p + \tau) = 0,$$

which define the induced choices $t^*(h_k; a)$ and $a^*(\tau)$ (with $da^*/d\tau \neq 0$ by regularity). The tax enters only the use condition G^a, so co-production feels the price solely through use: $dt^*/d\tau = (\partial t^*/\partial a)\, da^*/d\tau$.

Welfare. The planner values the path the price induces, rebating the revenue lump-sum:

$$\widehat{W}(h_k; \tau) \;=\; (1 - t^*)\,\Phi + T(h_{k+1}) - p\,a^* + \delta\,\widehat{W}(h_{k+1}; \tau),$$

whose shadow value $\widehat{W}' \equiv \partial\widehat{W}/\partial h$ is the V' that appears in (15); its sign is established in step (ii) below.

The optimal correction. A marginal change in τ shifts the expert's choices (t^*, a^*), which move the period's flow and, through h_{k+1}, every later period. Take one period's contribution, holding the inherited stock fixed; its derivative in τ is

$$\big(s_t + \delta\widehat{W}'H_t\big)\frac{dt^*}{d\tau} \;+\; \big(s_a + \delta\widehat{W}'H_a\big)\frac{da^*}{d\tau},$$

where $s_t \equiv -\Phi + T_h H_t$ and $s_a \equiv (1 - t_k)\Phi_a + T_h H_a - p$ are the social per-period returns to the two margins. The expert's own first-order conditions evaluate them: her co-production condition is $s_t = 0$, and her use condition, private return equated to the *full* price $p + \tau$, gives $s_a = \tau$. These flow terms thus drop out by the expert's own optimality (the envelope theorem), and with $dt^*/d\tau = (\partial t^*/\partial a)\, da^*/d\tau$ the contribution collapses to

$$\left[\delta \widehat{W}' \left(H_a + H_t \frac{\partial t^*}{\partial a}\right) + \tau\right] \frac{da^*}{d\tau} \;=\; \left[\delta \widehat{W}' \frac{\partial h_{k+1}}{\partial a} + \tau\right] \frac{da^*}{d\tau}.$$

The bracket is the whole result: the planner wants the price to equal the discounted damage a marginal unit of AI use inflicts on the novice's stock, $\delta \widehat{W}' (-\partial h_{k+1}/\partial a)$, read at the current state. Setting the bracket to zero state by state gives the Markov schedule of part (i). With the bracket zero at every state along the induced path, no marginal perturbation of the use path raises $\widehat{W}$. The schedule therefore makes the expert's use condition coincide with the constrained planner's throughout.

For a *constant* rate, the derivative of $\widehat{W}$ away from a steady state is a discounted sum of state-dependent brackets, and no single rate zeroes them all. At the steady state that the constant rate induces, the same state, policy, and shadow value recur every period. The sum is then geometric, and its root is the fixed point (15) with every object read at that allocation. This is the constant-rate characterization of part (i).

Sign (ii). The shadow value is positive where the schedule reads it. At the steady state the schedule induces, the value recursion closes as in the proof of Lemma A.1: the co-production response loses its flow part to the expert's condition, and the use response drops against the zero bracket. Solving gives $\widehat{W}' = \big[(1 - t^*)\Phi_h + T_h\, \gamma\big] \big/ \big(1 - \delta\, F'\big) > 0$, with F' the transition slope at the chosen use level: the numerator is positive, and the denominator is at least $1 - \delta$ by non-amplifying stability (Lemma 2). Away from a steady state, positivity is read as part of the regularity of Section 4.1. In the calibrated class the transition slope lies in $(0, 1)$ throughout (proof of Lemma 2), so the flow, an expert envelope rising in the stock, and the transition both rise with the stock, and $\widehat{W}$ increases outright. With $\widehat{W}' > 0$, the sign of $\hat{\tau}$ is the opposite of $\partial h_{k+1}/\partial a = H_a + H_t\, \partial t^*/\partial a$: a tax where AI use erodes the novice's stock, a subsidy where it builds it.

Force-free cut (iii). The cut is $\partial h_{k+1}/\partial a = 0$, that is $H_a + H_t\, \partial t^*/\partial a = 0$—the same balance as the shadow index's cut, now in the use margin a rather than the level A. The fixed-h co-production response is obtained by differentiating the co-production condition, exactly as in the proof of Lemma 1, with a in place of A:

$$\frac{\partial t^*}{\partial a} \;=\; \frac{\Phi_a - T_{hh} H_a H_t - T_h H_{ta}}{T_{hh} H_t^2 + T_h\, H_{tt}}.$$

Substituting and running the cancellation from the proof of Proposition 3 verbatim (the two $T_{hh}H_aH_t^2$ terms cancel, then $T_h = \Phi/H_t$ removes the marginal appropriation) leaves

$$H_a \;-\; \frac{H_t H_{ta}}{H_{tt}} \;+\; \frac{H_t^2\,\Phi_a}{\Phi\,H_{tt}} \;=\; 0,$$

free of T in level, slope, and curvature, exactly as in Proposition 3. In elasticity form this is $\nu = \epsilon\mu_a$, with $\mu_a = -\partial \ln t^*/\partial a$ the use-compression.

Corners (iv). At a zero-use corner the expert's use condition is the inequality $(1-t^*)\Phi_a + T_h H_a \le p + \tau$, read at $a = 0$: exactly the rates $\tau \ge \tau_0(h_k) \equiv (1-t^*)\Phi_a + T_h H_a - p$ support $a = 0$, an allocation identical to prohibition's, so their welfare coincides with the ban's. At the frontier corner the condition is the mirror inequality at $a = A$, supported by every rate no larger than $\tau_A(h_k)$. The interior formula prices the band between. □

Proof of Proposition 5 (cross-occupation erosion and augmentation)

By Proposition 1, $\operatorname{sign} dh^*(A;\rho)/dA = \operatorname{sign}\big(\nu(A;\rho) - \epsilon\mu\big)$: each occupation's stock moves with its own contest. At the tacit pole the reach vanishes (Assumption 5), so the contest there equals $-\epsilon\mu$, negative wherever AI compresses co-production ($\mu > 0$). By hypothesis the contest is continuous and monotone in ρ and changes sign in the interior. Starting negative, it therefore rises through zero exactly once: continuity and the sign change deliver a crossing (intermediate value theorem), and monotonicity makes it unique. Below the crossing the stock erodes, above it the stock is augmented—the single interior crossover, with the stock falling pointwise at every ρ below it and rising pointwise at every ρ above it. □

Proof of Proposition 6 (sign reversal and the force-free cut)

Parts (i) and (ii). By Proposition 3, $\operatorname{sign}\tau^*(A;\rho) = \operatorname{sign}\big(\epsilon\mu - \nu(A;\rho)\big) = -\operatorname{sign}\big(dh^*/dA\big)$. At the tacit pole the reach vanishes (Assumption 5), so the contest there equals $\epsilon\mu$, positive wherever AI compresses co-production: the index points in the tax direction at that pole. Continuous and monotone in ρ, changing sign in the interior, the contest falls through zero exactly once: continuity and the sign change deliver the zero (intermediate value theorem), and monotonicity makes it unique. This is the single reversal of (i), and the crossing is the unique cut-point $\bar\rho(A)$ of (ii), characterized by $\nu(A;\bar\rho) = \epsilon\mu$. The cut is force-free by Proposition 3(ii): in its proof above, the curvature terms of T cancel and the expert's optimality substitutes out the marginal appropriation, so neither T nor V' appears in the condition. The strictly positive prefactor of (10) fixes only $|\tau^*|$. That magnitude inherits the

occupation-specific variation of $V'(A;\rho)$, $h^*(A;\rho)$, and the persistence, and is not in general monotone in ρ.

Part (iii). The same argument runs on the use margin. By Proposition 4, $\operatorname{sign}\hat{\tau}(A;\rho) = -\operatorname{sign}\partial h_{k+1}/\partial a$, and the use-margin contest $\epsilon\mu_a - \nu$ settles that sign, with its force-free cut $\nu = \epsilon\mu_a$ established there. Under the continuity, monotonicity, and interior sign change hypothesized for this contest, it too falls through zero exactly once: the levied price is a tax on the tacit side of its own cut and a subsidy on the explicit side. □

A sufficient primitive class for the cross-occupation contest

The monotone-contest hypothesis of Propositions 5 and 6 is delivered here from primitives. The class leaves the persistence profile, the appropriation scale, and the augmentation profile free functions of expressibility. Production is arbitrary in part (i) and log-separable in part (ii); the parts differ in where the contest is read.

Lemma A.2 (Sufficient primitive class for the contest)**.** *Let transmission be isoelastic in co-production time with a multiplicative augmentation, $H(h_k, t_k, A;\rho) = c(h_k;\rho)\, t_k^a\, \psi(A;\rho)$ with $a \in (0,1)$, satisfying Assumptions 2 and 5, and let appropriation be the power $T(h) = \beta(\rho)\, h^b$ with $\beta(\rho) > 0$ and $b \in (0,1]$. The profiles $c(\cdot\,;\rho)$, $\beta(\rho)$, and $\psi(\cdot\,;\rho)$ are otherwise unrestricted. Write $g \equiv \partial \ln \Phi/\partial A$, and read each part at interior regular optima.*

(i) ***Any production, common-stock read.*** *For any aggregator Φ satisfying Assumption 1, the fixed-stock contest is*

$$\nu(A;\rho) - \epsilon\,\hat{\mu} \;=\; \frac{\nu(A;\rho) - a\, g(h_k, A)}{1 - ab}, \qquad \hat{\mu} \equiv -\left.\frac{\partial \ln t^*}{\partial A}\right|_{h_k}. \tag{A.2}$$

Read at a common stock, it is strictly increasing in ρ, negative at the tacit pole wherever $g > 0$, and zero at most once—in the interior of the spectrum exactly when $\nu(A;1) > a\, g$.

(ii) ***Log-separable production, equilibrium read.*** *If in addition $\Phi = \phi_1(h_k)\,\phi_2(A)$, then $g = \phi_2'(A)/\phi_2(A)$ carries no stock argument, and at every stable steady state*

$$\operatorname{sign}\frac{dh^*(A;\rho)}{dA} \;=\; \operatorname{sign}\big(\nu(A;\rho) - a\, g(A)\big). \tag{A.3}$$

The sorting of Proposition 5 then holds at each occupation's own steady state: erosion below, augmentation above a single crossover at $\nu(A;\bar{\rho}) = a\, g(A)$, interior exactly when $\nu(A;1) > a\, g(A)$.

Proof. The optimum in closed form. With $T_h = b\beta\,(h_{k+1})^{b-1}$ and $H_t = aH/t_k$, the expert's condition (3) reads $ab\,\beta\,H^b = t_k\,\Phi$. Substituting $H = c\,t_k^a\,\psi$ and collecting,

$$(1-ab)\,\ln t^* \;=\; \ln(ab\,\beta) \;+\; b\ln c(h_k;\rho) \;+\; b\ln\psi(A;\rho) \;-\; \ln\Phi(h_k,A),$$

with $1-ab>0$, so the interior optimum is unique, and the learning-by-doing elasticity is the constant $\epsilon = a$.

Part (i). Differentiate the closed form in A at fixed h_k: $(1-ab)\,\partial\ln t^*/\partial A = b\,\nu - g$, so $\hat{\mu} = (g - b\,\nu)/(1-ab)$. Substituting,

$$\nu - a\,\hat{\mu} \;=\; \nu - \frac{a\,(g-b\,\nu)}{1-ab} \;=\; \frac{\nu - a\,g}{1-ab},$$

which is (A.2). At a common stock the term $a\,g$ is common across occupations, so the contest moves in ρ through the reach alone. It is strictly increasing by Assumption 5, equals $-a\,g/(1-ab)$ at the tacit pole by the floor $\nu(A;0)=0$, and therefore crosses zero at most once; the crossing is interior exactly when the reach at the explicit pole clears $a\,g$.

Part (ii). At a steady state, $h^* = c(h^*;\rho)\,(t^*)^a\,\psi(A;\rho)$. Write $\gamma_c \equiv \partial\ln c/\partial\ln h$ and $s_h \equiv \partial\ln\Phi/\partial\ln h$; the closed form gives $\partial\ln t^*/\partial\ln h_k = (b\,\gamma_c - s_h)/(1-ab)$. Total differentiation of the fixed point in A collects to

$$\frac{d\ln h^*}{dA}\Big[(1-ab) - (\gamma_c - a\,s_h)\Big] \;=\; \nu - a\,g.$$

The bracket is the stability margin. The equilibrium map $F(h) = H(h, t^*(h;A), A)$ has elasticity $\gamma_c + a\,(b\,\gamma_c - s_h)/(1-ab) = (\gamma_c - a\,s_h)/(1-ab)$ at the steady state, and it lies below one exactly when the bracket is positive, which holds at the stable steady state of Lemma 2. Hence sign $dh^*/dA = \text{sign}(\nu - a\,g)$, which is (A.3). With g free of the stock, the sign-carrier $\nu(A;\rho) - a\,g(A)$ is strictly increasing in ρ, negative at the tacit pole wherever $g>0$, and zero at most once, exactly as in part (i); erosion below the crossing and augmentation above it are the sorting of Proposition 5. □

The calibrated class meets part (i) with $b=1$, and its level normalization reads the cross-section at the common stock $h^*=1$: the read the proof below exploits. Its CES aggregator lies outside part (ii), and the crossing at the corrected equilibria is the numerical verification of Section 6.3.

Single crossing in the calibrated class

Proposition A.1 (Single crossing in the calibrated class)**.** *At the decentralized steady state under the level choice $h^*(A_0) = 1$, the stock response reduces to*

$$\operatorname{sign}\Big(\frac{dh^*}{dA}\Big) \;=\; \operatorname{sign}\big(\nu(\rho) - \alpha\, g\big), \qquad g \equiv \frac{\partial \ln \Phi}{\partial A}\Big|_{h=1,\, A=A_j},$$

where g is common to every occupation at a given exposure. Since the reach at the occupation's AI level, $\nu(\rho) = \eta_{\max}\rho\,\omega_j/(1 + A_j)$, is strictly increasing in ρ while αg is constant in ρ, the contest is strictly monotone in ρ and crosses zero at most once; the crossing $\bar\rho(\omega)$ solves $\nu(\bar\rho) = \alpha g$ and is interior exactly when $\alpha g < \eta_{\max}\,\omega_j/(1 + A_j)$.

Proof. Under the parametric technologies of Section 6, the expert's interior optimum is $t^* = (\alpha\mathcal{W}/\Phi)^{1/(1-\alpha)}$ with $\mathcal{W} = \beta B h^\gamma \psi(A)$, so at fixed stock

$$\frac{\partial \ln t^*}{\partial A}\Big|_h = \frac{\nu - g}{1 - \alpha}, \qquad \frac{\partial \ln t^*}{\partial \ln h} = \frac{\gamma - s_h}{1 - \alpha},$$

where $\nu = \partial \ln \psi/\partial A$, $g = \partial \ln \Phi/\partial A$, and $s_h = \partial \ln \Phi/\partial \ln h$; the appropriation is linear, so no T-curvature term enters. Total differentiation of the steady state $\ln h^* = [\ln B + \alpha \ln t^* + \ln \psi]/(1 - \gamma)$ in A, folding in the policy response, collects to

$$\frac{d \ln h^*}{dA}\Big[1 - \alpha - \gamma + \alpha s_h\Big] \;=\; \nu - \alpha g.$$

The bracket is the stability margin: the equilibrium map's elasticity is $(\gamma - \alpha s_h)/(1 - \alpha)$, and it lies below one exactly when the bracket is positive, which holds at the stable steady state of Lemma 2. Hence $\operatorname{sign}(dh^*/dA) = \operatorname{sign}(\nu - \alpha g)$.

Under the level normalization the decentralized stock is $h^*(A_0) = 1$ in every occupation, so at a given exposure ω_j both g and s_h are evaluated at the same point $(1, A_j)$ and are common across ρ: the occupation's expressibility enters the contest only through the reach. With $\nu(\rho) = \eta_{\max}\rho\,\omega_j/(1 + A_j)$ strictly increasing in ρ and $\nu(0) = 0 < \alpha g$, the contest $\nu(\rho) - \alpha g$ is strictly increasing, negative at the tacit pole, and crosses zero at most once—at $\bar\rho(\omega_j)$ solving $\nu(\bar\rho) = \alpha g$, interior exactly when $\alpha g < \eta_{\max}\omega_j/(1 + A_j)$. Finally, $\nu = \alpha g$ is the force-free cut: at fixed stock $\hat\mu = (g - \nu)/(1 - \alpha)$, so $\alpha\hat\mu - \nu = (\alpha g - \nu)/(1 - \alpha)$, and the two conditions share their zero. □

Comparative statics of the cut in the AI level

The cut $\nu(\bar{\rho}) = \alpha\mu$ moves as A grows because both sides fall. The reach declines,

$$\frac{\partial \nu(\rho, A)}{\partial A} \;<\; 0, \qquad \nu(\rho, A) \to 0 \;\text{ as } A \to \infty$$

(for the calibrated $\psi = (1+A)^{\eta(\rho)}$, $\nu = \eta(\rho)/(1+A)$, so $\partial\nu/\partial A = -\eta(\rho)/(1+A)^2 < 0$). The compression $\alpha\mu$ also falls. Under gross-substitutes production ($\sigma \in (0,1)$) the opportunity cost of co-production grows only linearly in A at high levels, so its proportional growth $\partial \ln \Phi/\partial A$ vanishes; in the parametric model the fixed-h compression is the gap $\big(\partial \ln \Phi/\partial A - \nu\big)/(1-\alpha)$, and both terms go to zero, so $\mu = -\,d\ln t^*/dA \to 0$. Which side falls faster is *not* determined by the qualitative structure—it depends on the production technology—so the direction in which $\bar{\rho}$ drifts is a quantitative, calibration-specific question rather than a theorem.

Calibrated direction. In the CES calibration of Section 6 the compression falls faster than the reach along the logistic path, so the tax-side set $\{\rho : \alpha\mu > \nu(\rho)\}$ *shrinks* as A rises. The cut $\bar{\rho}$ drifts *toward* the tacit pole, and AI's net erosion is concentrated at low-to-moderate AI, receding as AI matures. Direct evaluation gives $\bar{\rho}$ falling monotonically from ≈ 0.10 at $A = 0.5$ toward 0 at high A. The erosion the tacit pole suffers during the transition is what early intervention addresses.

Front-loaded corrective need. The discounted-welfare gain of a complementary policy that raises the augmentation reach, introduced at intervention generation k_0, is

$$\Delta V(k_0) \;=\; \sum_{k \geq k_0} \delta^k \cdot \big[Y_k^{\eta_{\text{new}}} - Y_k^{\text{baseline}}\big].$$

Discounted to the present, $\Delta V(k_0)$ is decreasing in k_0: a later start forgoes the early, lightly discounted periods, and the augmentation reach ν_{new} the policy delivers declines along the AI path, so a later policy also acts on a smaller reach. Early intervention dominates. Numerical verification is in Appendix L: $\Delta V(k_0)$ falls monotonically in k_0 for every occupation in the cross-section, with the gain largest in the codifiable occupations where AI can teach. The drift *direction* is a property of the calibrated production side, not of the Polanyi structure; the structural result is the static sign reversal.

B Robustness to Altruism and Novice Financing

In the main text, the expert is myopic with respect to post-retirement productivity (Assumption 4). This appendix establishes that the central wedge result is robust to Beckerian altruism on the expert's side and to novice-side financing of formation.

Setup.— We replace the expert's objective with

$$U_k = Y_k + \alpha_b \cdot V(h_{k+1}; A), \tag{B.1}$$

where $\alpha_b \in [0, 1]$ is the weight the expert places on the social value her novice's stock carries forward, V the value function of (6).

Wedge under altruism.— Differentiating the augmented objective with respect to t_k gives the expert's modified first-order condition under altruistic preferences:

$$-\Phi_k + \left(T_h + \alpha_b V'\right) H_t = 0.$$

Against the planner's first-order condition (7), the wedge between the two is now

$$\text{WEDGE}_k^{\alpha_b} = (\delta - \alpha_b)\, V'\, H_t.$$

Three observations follow: (i) $\alpha_b = 0$ recovers the main-text wedge in (8); (ii) $\alpha_b = \delta$ makes the wedge vanish (full alignment of decentralized and planner allocations); (iii) for $\alpha_b \in (0, \delta)$, the wedge is strictly positive and, at a fixed allocation, scales linearly: $\text{WEDGE}^{\alpha_b}/\text{WEDGE}^0 = (\delta - \alpha_b)/\delta = 1 - \alpha_b/\delta$.

Proposition B.1 (Wedge robustness to altruism)**.** *For any $\alpha_b \in [0, \delta)$, the intergenerational wedge remains strictly positive, and at a fixed allocation its magnitude scales by $1 - \alpha_b/\delta$ relative to the $\alpha_b = 0$ baseline of (8).*

Proof. The altruistic expert's first-order condition carries the effective marginal appropriation $T_h + \alpha_b V'$ where the planner's (7) carries $T_h + \delta V'$, so their gap on the marginal co-production hour is $(\delta - \alpha_b)\, V'\, H_t$. It is strictly positive for $\alpha_b < \delta$, since $V' > 0$ (Section 4.1) and $H_t > 0$. Holding the allocation—and with it V' and H_t—fixed, dividing by the baseline wedge $\delta V' H_t$ of (8) gives the ratio $1 - \alpha_b/\delta$. □

Calibrated magnitude.— The relevant range for α_b sits well below the parent–child benchmark on which Becker and Murphy (1988) build the theory of family investment: the expert–novice relationship is non-kin and contractual. We therefore take an illustrative upper bound of

$\alpha_b = 0.15$. At that bound, with $\delta = 0.40$, the wedge retains $1 - 0.15/0.40 = 62.5\%$ of its baseline magnitude, and every α_b below the bound retains more.

Novice-side financing.— The altruism above works on the expert's side and leaves the novice passive; the mirror-image robustness lets him pay for his formation, as the below-product training wages of Section 2 suggest he does. Let $\lambda \in [0, 1]$ be the share of his own-lifetime return the novice can credibly commit to the expert, so her operative marginal return to co-production becomes $T_h + \lambda\,\delta\,m$, where $m(h) \equiv (1 - t^*)\Phi_h + T_h\,\gamma$ is the stock's per-period marginal value, the novice's own-lifetime stake in what the co-production forms. Since the steady-state shadow value of Lemma 3 is $V' = m/(1 - \delta\gamma)$, the residual wedge is

$$\text{WEDGE}^{\lambda} \;=\; \delta\left(V' - \lambda\, m\right) H_t \;=\; \left[1 - \lambda\,(1 - \delta\gamma)\right] \delta\, V'\, H_t,$$

scaling by $1 - \lambda(1 - \delta\gamma)$ at a fixed allocation, and the corrective subsidy scales with the same factor. Full one-link financing, $\lambda = 1$, still leaves fraction $\delta\gamma$ of the wedge: 0.34 at the tacit pole, 0.18 at the codifiable one. The novice finances the link he owns. The recursive tail beyond his own lifetime, his novice's formation and every later one, remains outside any contract he can sign. The wedge closes at no $\lambda \leq 1$, and it is largest exactly where the persistence is—the tacit occupations keep the most of their correction.

The internalization enters the operative appropriation as the planner's correction does, so the force-free structure carries over: λ relocates the equilibrium at which the contest is read and never enters the cut condition. Quantitatively, re-solving the cross-section with the λ-equilibrium as the decentralized benchmark lowers the employment-weighted welfare gap from +67.5% at $\lambda = 0$ to +30.5% at $\lambda = 0.5$ and +10.6% at full one-link financing. The headline gap of Section 6 is thus a zero-internalization value; what survives every λ is its sign, its cross-occupation shape, and the sign-reversing AI schedule read at the corrected allocation.

The novice's opportunity cost of co-production time is likewise absent from the baseline—his hours cost him nothing—which flatters the level of the subsidy s^*. A positive outside option would lower the decentralized allocation and the corrective subsidy together, touching neither the wedge's sign nor the cut.

Implication for the damage index.— At a fixed allocation the damage index scales with the wedge, $\tau^*(\alpha_b) = (1 - \alpha_b/\delta)\,\tau^*(0)$: at the bound $\alpha_b = 0.15$ it retains 62.5% of its baseline magnitude. Occupation-specific values inherit the same scaling, preserving the sign reversal across occupations documented in Figure 4.

C Robustness to Appropriation in the Expert's Own Stock

The main text takes the appropriation technology to depend on the novice's stock alone, $T(h_{k+1})$ (Section 3.2). A natural enrichment lets it depend also on the expert's own stock, $T(h_k, h_{k+1})$—the expert co-producing rather than only forming the novice. This appendix shows the enrichment leaves the expert's optimality condition and the force-free threshold untouched, altering only the shadow value; that a level wage paid to the novice inside the match alters nothing at all; and that an appropriation carrying AI directly is the one enrichment that would place a term in the threshold itself. Output becomes $Y_k = (1 - t_k)\Phi(h_k, A) + T(h_k, h_{k+1})$ with $h_{k+1} = H(h_k, t_k, A)$; write $T_1 \equiv \partial T/\partial h_k$ and $T_2 \equiv \partial T/\partial h_{k+1}$ for the two marginal appropriations, T_2 being the analogue of the baseline T_h, and keep T increasing and concave in the novice's stock (Assumption 3).

Proposition C.1 (Own-stock appropriation)**.** *Enrich the appropriation to $T(h_k, h_{k+1})$. Then:*

(i) the expert's first-order condition is $T_2(h_k, h_{k+1})\, H_t = \Phi(h_k, A)$—the form of (3) with T_2 in the role of T_h, the direct term T_1 absent;

(ii) the force-free threshold is unchanged: the cut is the T-free condition of Proposition 3, $\nu = \epsilon\mu$, involving neither T_1 nor T_2;

(iii) the shadow value alone acquires the new dependence, gaining T_1,

$$V'(h_k) = (1 - t_k)\,\Phi_h + T_1 + \big(T_2 + \delta\, V'(h_{k+1})\big)\, H_h,$$

so the damage index's sign and force-free cut are unchanged and only its magnitude, carried by the shadow value, is affected.

Proof. (i) First-order condition. When the expert chooses t_k, her own stock h_k is predetermined, so $\partial h_k/\partial t_k = 0$ and the direct argument contributes nothing: $\partial T(h_k, h_{k+1})/\partial t_k = T_2\, \partial h_{k+1}/\partial t_k = T_2 H_t$. Differentiating Y_k gives $-\Phi + T_2 H_t = 0$, which is (3) with T_h replaced by T_2; T_1 does not enter.

(ii) Force-free threshold. At the cut, $dh^*/dA = 0$, so the total and fixed-h_k responses of t^* coincide and the cut collapses to $H_A + H_t\, \partial t^*/\partial A = 0$ (proof of Proposition 3). With h_k held fixed, $\partial t^*/\partial A$ solves the same optimality condition as in the baseline, with T_2 for T_h and T_{22} for T_{hh}; the cross-derivative T_{12} never appears, since it would multiply $\partial h_k/\partial A = 0$. The cancellation of Proposition 3 then runs verbatim—the two curvature terms cancel and

the optimality condition $T_2 = \Phi/H_t$ removes the marginal appropriation—leaving the same T-free cut, free of T_1 and T_2 alike.

(iii) Shadow value. Applying the envelope theorem to the planner's value (6) with the enriched return $(1 - t_k)\Phi(h_k, A) + T(h_k, h_{k+1}) + \delta V(h_{k+1}, A)$ and differentiating in h_k gives the displayed shadow value. The stock h_k is now the state, so its direct appearance in T contributes. The result is the baseline envelope, with T_2 for T_h, plus the direct term T_1 (at the steady state the closed form (9) gains T_1 in its numerator). Because V' enters the rate $\tau^* = \delta V'(-dh^*/dA)$ only as the strictly positive prefactor (Proposition 3), the added term rescales the rate's magnitude but leaves its sign and its zero, $dh^*/dA = 0$, unchanged. □

A level wage paid to the novice inside the match is likewise immaterial. Let the expert pay $w \geq 0$, a payment independent of co-production time and of the formed stock. She then maximizes $Y_k - w$, and a constant enters no first-order condition: the decentralized allocation is unchanged. The planner's period return is the period's total output, the two parties' payoffs summed, so the transfer cancels and w never enters the planner's problem. Both allocations, and with them every object built on them—the wedge, the shadow value, the damage index—are exactly as in the baseline.

An appropriation with a direct AI argument, $T(h_{k+1}, A)$, marks the boundary of these invariances. Differentiating the expert's condition $T_h(h_{k+1}, A)\, H_t = \Phi$ in A adds the term $-T_{hA}\, H_t$ to the numerator of the co-production response in the proof of Proposition 3. The cancellation there removes the appropriation's curvature and slope but leaves this cross-term standing: the threshold gains a T_{hA} term, exactly as the exclusion of Section 3.2 asserts. The force-free property thus rests on AI entering the appropriated return only through the formed stock.

D Robustness to Occupational Exposure Intensity

The calibration lets occupations differ in how much of their work AI can perform, scaling the effective AI level an occupation faces to $A\,\omega_j$ with an exposure intensity $\omega_j \in [0, 1]$ that varies across occupations independently of expressibility (Section 6). This appendix shows the scaling is a calibration refinement, not a structural one: it relocates an occupation along the A-axis and rescales the magnitude of its response, leaving the reach-versus-compression contest, its sign, and the force-free cut unchanged. Hold an occupation fixed, so $\omega \equiv \omega_j$ is a positive constant; the case $\omega = 0$ is the trivial one in which AI never enters that occupation. Let H and Φ depend on A only through the effective level $\tilde{A} \equiv A\,\omega$. Tildes denote objects read in $\tilde{A}$: $\tilde{\nu} \equiv \partial \ln H/\partial \tilde{A}$ and $\tilde{\mu} \equiv -\,d \ln t^*/d\tilde{A}$.

Proposition D.1 (Robustness to exposure intensity)**.** *With effective level* $\tilde{A} = A\,\omega$*,* $\omega \in (0, 1]$*, the reach and compression in the nominal level scale by* ω*,* $\nu = \omega\tilde{\nu}$ *and* $\mu = \omega\tilde{\mu}$*, while the persistence* γ *and the learning-by-doing elasticity* ϵ *are unchanged. Hence*

$$\frac{dh^*}{dA} = \omega\,\frac{h^*\,(\tilde{\nu} - \epsilon\tilde{\mu})}{1-\gamma}, \qquad \tau^* = \omega\,\frac{\delta\,V'\,h^*}{1-\gamma}\,(\epsilon\tilde{\mu} - \tilde{\nu}).$$

The positive factor ω *rescales both magnitudes but cancels from the sign of the stock response, the sign of the index, and the force-free cut, which remains* $\tilde{\nu} = \epsilon\tilde{\mu}$*. Exposure intensity therefore relocates the occupation on the A-axis: a given effective level, and the cut with it, is reached at nominal level* $A = \tilde{A}/\omega$*. It also scales the occupation's response, leaving the contest, the sign reversal, and the cut condition intact.*

Proof. Production and transmission depend on A only through $\tilde{A} = A\omega$, so the expert's optimality condition $T_h H_t = \Phi$ does too, making t^* a function of $\tilde{A}$; the chain rule then sends $\partial/\partial A \mapsto \omega\,\partial/\partial\tilde{A}$ on every derivative in the nominal level. The reach and compression therefore scale by ω, $\nu = \partial \ln H/\partial A = \omega\,\partial \ln H/\partial\tilde{A} = \omega\tilde{\nu}$ and $\mu = -\,d \ln t^*/dA = \omega\tilde{\mu}$, whereas $\gamma = \partial \ln H/\partial \ln h$ and $\epsilon = \partial \ln H/\partial \ln t$, read at the same steady state, carry no A-derivative and are unchanged. Substituting into Proposition 1 and Proposition 3 gives the displayed expressions, with contest $\nu - \epsilon\mu = \omega(\tilde{\nu} - \epsilon\tilde{\mu})$ and prefactor $\delta V' h^*/(1-\gamma)$—a level object, free of any A-derivative and so of ω. Since $\omega > 0$, $\operatorname{sign}(dh^*/dA) = \operatorname{sign}(\tilde{\nu} - \epsilon\tilde{\mu})$ and $\operatorname{sign}\tau^* = \operatorname{sign}(\epsilon\tilde{\mu} - \tilde{\nu})$, both independent of ω; and $\tau^* = 0 \Leftrightarrow \tilde{\nu} = \epsilon\tilde{\mu}$, the cut in the effective level, free of ω. The exposure intensity enters only the magnitudes and the map $A \mapsto \tilde{A} = A\omega$. □

E Calibration Data Details

Task expressibility ρ_j from O*NET

We construct ρ_j for each occupation from the O*NET 30.3 task-content database, adapting the task-composite approach of Acemoglu and Autor (2011). For every occupation we form a routine-cognitive index: the standardized average of items capturing codifiable, rule-bound work, minus the standardized average of items capturing tacit work. The routine-side items are "importance of repeating the same tasks," "degree of automation," "documenting and recording information," "processing information," "performing administrative activities," and "working with computers." The tacit-side items are "thinking creatively," "assisting and caring for others," "establishing and maintaining interpersonal relationships," "resolving conflicts and negotiating," "providing consultation and advice," and "freedom to make deci-

sions." Each item is z-scored across the 682 occupations, and the index is rank-normalized to the unit interval. It orders occupations as expressibility should: court reporters, data-entry keyers, and payroll clerks at the top, and choreographers, clergy, counselors, surgeons, and the performing arts at the bottom.

Two of the routine-side items, "degree of automation" and "working with computers," are automation-adjacent, so we re-derive the index without them. The leave-out index has rank correlation 0.96 with the baseline, and its rank correlation with exposure falls to 0.04. Recomputing the headline gives a tax-side share of 18.48% of occupations (21.53% of employment) at an employment-weighted index of -0.23.

AI Occupational Exposure ω_j from Felten et al.

We take the AI Occupational Exposure (AIOE) index of Felten et al. (2021) at the Standard Occupational Classification (SOC) 6-digit level—it links AI capabilities to occupational abilities—merge it to the O*NET occupations, and rank-normalize to $[0, 1]$.

Expressibility and exposure are nearly independent across occupations, and the association is estimated precisely enough for the near-independence to carry weight. The rank correlation is 0.06 ($p = 0.09$), the Pearson correlation is 0.06 with a 95% confidence interval of $[-0.02, 0.14]$, and an OLS fit of exposure on expressibility gives slope 0.06 (s.e. 0.04) with $R^2 < 0.01$. Expressibility thus explains less than half a percent of the variance in exposure. Under cardinal rather than rank measurement the slope is 0.10 (s.e. 0.05, $R^2 < 0.01$).

With both axes cut at terciles, the tacit-yet-exposed corner—expressibility in the bottom third, exposure in the top third—holds 72 of the 682 occupations and 9.03% of employment. Its largest members by employment are elementary, middle, and secondary school teachers, lawyers, and management analysts. AI exposure tracks the cognitive content of work rather than its routineness: the most exposed occupations are actuaries, accountants, judges, and financial analysts, the least exposed dancers, fitness trainers, and construction helpers. It does not line up with routine task content, as the first AI-exposure measures already showed (Brynjolfsson et al., 2018; Webb, 2020; Eloundou et al., 2024).

In the model this near-independence separates the two measures' roles, exposure scaling the magnitude of an occupation's response and expressibility setting its sign. It places the tacit, highly exposed occupations, where AI does the work but cannot codify it back, at the locus of erosion (Section 2).

Employment weights n_j from BLS OEWS

The employment weight n_j is national employment by detailed SOC occupation from the BLS Occupational Employment and Wage Statistics (OEWS, May 2025), retrieved through the BLS public API and merged to the cross-section. Fourteen O*NET detailed occupations are reported by OEWS only under coarser, SOC-vintage-consolidated codes (for example the three buyer codes under "Buyers and Purchasing Agents," or substance-abuse and mental-health counselors under their merged code). These are recovered through a crosswalk that distributes each consolidated total across its constituents, net of any already-matched member. All 682 occupations thus carry an employment weight, covering 125 million workers; the weights enter the aggregate index $\bar{\tau}$ and the welfare gap. Employment is essentially uncorrelated with expressibility (rank correlation 0.03), so the tax-side set holds employment in proportion to its number, and occupation and employment shares nearly coincide throughout.

Transmission persistence $\gamma(\rho)$

We set the persistence schedule $\gamma(\rho)$ to reproduce a stylized regularity rather than estimate it: tacit, hard-to-codify skills persist across cohorts more than codifiable ones and take correspondingly longer to master. We stipulate the working ranges: time-to-mastery of 5–15 years in tacit-intensive crafts and 1–3 years in codifiable skills. The ranges contain the institutional durations Section 2 documents: the seven-year surgical track and the five-to-six-year doctoral training on the tacit side, the three-year analyst track at the codifiable end. Mastery here means full command of the occupation's expertise, so the tacit range runs past the entry requirements that occupational data record. That range brackets the ten or more years of preparation the expertise literature documents for expert performance across complex domains, medicine among them (Ericsson et al., 1993). The schedule uses the three- to fivefold spread between the poles, not the levels.

Time-to-mastery is not the persistence itself, and the two should not be conflated. The persistence γ is a cross-*generation* object—the elasticity of the next generation's stock in the expert's, how faithfully quality carries forward—while time-to-mastery is within-*generation*, the years a novice needs to acquire the skill. We use the latter as a proxy for the former: knowledge that takes longer to master is deeper and more tacit, and deeper knowledge both takes longer to master and decays more slowly down the chain of later experts. The proxy disciplines two features of the schedule and no more. It fixes the *ordering*, γ decreasing in ρ; and it fixes the *gap* between the poles. Persistence is read through the horizon $1/(1-\gamma)$, the cumulative weight a cohort's stock carries down the chain. Thus $\gamma = 0.85$ and $\gamma = 0.45$

give weights of 6.7 and 1.8, a ratio of about 3.7 that sits within the three- to fivefold spread in time-to-mastery. The raw ratio of the γ values, 1.9, is not the object being matched; the horizon is.

The *level* is anchored separately, not by time-to-mastery. The tacit pole is held at $\gamma = 0.85$ because that is the stability ceiling—a single stock with γ near one compounds explosively (Section 6.2), so the level is a modeling bound rather than an estimate. The later-stage self-productivity estimates of Cunha et al. (2010) are 0.90 for cognitive and 0.87 for noncognitive skills (their Table I, corrected for measurement error). These are measured across two-year periods, not 30-year generations, so they discipline the qualitative claim that skill stocks reproduce with high persistence, not the level itself. Given that pole and the target horizon ratio, the codifiable pole falls at 0.45, and we interpolate linearly: $\gamma(\rho) = 0.85 - 0.40\,\rho$ across $[0.45, 0.85]$.

This is the loosest of the calibration's anchors—a proxy mapping rather than a direct estimate, so the endpoints are illustrative—but the results do not rest on it. Reading the mastery spread at its upper end, for instance, would put the codifiable pole near $\gamma = 0.25$ rather than 0.45. Across the $\pm 30\%$ persistence perturbations and the constant-persistence counterfactual $\gamma(\rho) = 0.7$ of Appendix H (Table H.2), the sign reversal and the force-free cut are preserved. Persistence reaches the cut only through the compression μ, so $\bar{\rho}$ barely moves; its main role is to scale the dispersion of $|\tau^*|$ and, through the multiplier $1/(1-\delta\gamma(\rho))$, the level of the welfare gap.

Becker appropriation $\beta(\rho)$

We set the appropriation schedule $\beta(\rho)$ from the structure of co-production itself. In the model, $T(h) = \beta(\rho)\,h$ is the novice route's entire contemporaneous output: the contribution the novice's doing delivers within the shared period, all of which accrues to the expert. So β is the current output a unit of formed stock carries with it: a productivity coefficient of the route, not a division of its proceeds. The period flow Y_k thus counts the novice's contribution once, in full.

The bundle fixes the ordering and the poles. Where knowledge is tacit, formation happens inside production: the novice's doing is the occupation's real work—the resident operates, the apprentice fabricates. Each unit of stock formed thus delivers a large contemporaneous contribution, and we set $\beta \approx 0.80$ at the tacit pole. Where knowledge is codifiable, part of formation arrives apart from the doing, through instruction, manuals, and worked examples. A unit of formed stock thus embodies less contemporaneous doing and delivers less current output: $\beta \approx 0.30$ at the codifiable pole. The direction is the same gradient that separates

on-the-job from classroom acquisition across occupations. Between the poles we interpolate linearly: $\beta(\rho) = 0.80 - 0.50\,\rho$. The levels are illustrative; what the structure pins is the ordering and the sign of the gap.

Of the three schedules, β is the one whose calibration matters least. It enters the index only through the strictly positive prefactor (Proposition 6), so the sign reversal and the force-free cut are *structurally* free of it, not merely numerically robust: the cut condition $\nu(\bar{\rho}) = \alpha\mu$ carries no appropriation term. Varying β rescales the magnitude $|\tau^*|$ and the co-production subsidy $s^* = \delta V'/\beta$. Its only route to the sign runs through the equilibrium at which the contest is read. The ±30% sweep of Appendix H (Table H.2) bounds that channel at a cut movement of ±0.01, with the tax-side employment share moving by about two percentage points (11.15% to 13.41%).

AI training augmentation $\psi(A;\rho) = (1+A)^{\eta(\rho)}$ from Brynjolfsson–Li–Raymond

We calibrate the AI training augmentation to match the 36% productivity gain that Brynjolfsson et al. (2025) report for agents in the lowest skill quintile in a staggered deployment of a generative-AI assistant to customer-support work. In the cross-section, customer-service representatives sit near the codifiable pole with high exposure ($\rho_{\text{cs}} = 0.97$, $\omega = 0.73$); the training-side augmentation $\psi(A_j;\rho) = (1+A_j)^{\eta(\rho)}$ at the baseline $A_0 = 0.5$ should reproduce the gain. With $\eta(\rho) = \rho$ (so $\eta_{\max} = 1$), the occupation-effective AI level is $A_j = A_0\,\omega \approx 0.37$, and the augmentation is $(1.37)^{0.97} \approx 1.36$, reproducing the published estimate to within a point. Mapping the moment into ψ takes two readings, and they carry different weight. The first is clean: over the study's horizon the experts' co-production allocation is fixed, so the treated–untreated gap is free of the reallocation of t^*, a generational-frequency response that enters the model elsewhere. The second is a calibration choice the evidence supports but does not fix: reading the gap as formation rather than contemporaneous tool assistance. Treated agents retain part of the gain during system outages, when the tool is unavailable—the signature of acquired skill rather than assisted output—so part of the gain is formation, with the share unquantified. The baseline takes the full gain as formation. The augmentation-slope sweep of Appendix H reads directly as the share: $\eta_{\max} = 0.7$ is the calibration in which 70% of the observed gain is formation. The sweep, not the point, thus carries the moment into the model. A parallel asymmetry appears in the field (Dell'Acqua et al., 2026): consultants assigned AI scored roughly 30% higher than peers on tasks within the technology's frontier but performed worse on a task selected to lie beyond it, where they were 19 percentage points less likely to produce a correct solution. The one-parameter form

thus reproduces the target without a separate scale parameter; the augmentation slope $\eta_{\max}$ is raised to 1.25 in the closing-window counterfactual of Appendix L. The full transitional dynamics are reported in Appendix J.

The augmentation slope is the one anchor the results are most sensitive to—the counterpoint to the appropriation coefficient. It sits directly in the force-free cut: with $\nu(\rho) = \eta_{\max}\rho/(1+A)$, the condition $\nu(\bar{\rho}) = \alpha\mu$ makes $\bar{\rho}$ scale as $\alpha\mu/\eta_{\max}$, so the cut-point is as sensitive to $\eta_{\max}$ as to any parameter in the model, and $\eta_{\max}$ alone rests on a single field study. Appendix H therefore probes it hardest: a slope sweep over $[0.5, 2.0]$, the $\pm 30\%$ perturbation, and the curvature bend. The reversal survives every case. What $\eta_{\max}$ moves is where the tax region ends, not whether it exists.

Level normalization B_j

Each occupation's training productivity B_j is set so that its decentralized steady-state stock at the current AI level equals one, $h^*(A_0) = 1$, which makes index values and welfare gaps comparable across occupations and fixes the units in which τ^* is denominated. The choice is a calibration object rather than a pure normalization. Under the constant-elasticity-of-substitution (CES) aggregator the level of the stock sets the human-capital production share $s_h = \partial \ln \Phi / \partial \ln h$, which enters the compression μ, so re-targeting the normalized stock moves the contest the way the factor shares do. Recomputing the cross-section at $h^*(A_0) = 0.5$ and 2 moves the cut to 0.13 and 0.08 and the tax-side employment share to 15% and 9%, within the ranges of the structural sweep (Table H.2), with the sign reversal preserved throughout.

The remaining scalars

The scalar parameters carry conventional or assumed values, each probed in Appendix H. The learning-by-doing elasticity $\alpha = 0.30$ is assumed. The Ben-Porath (1967) tradition supplies the diminishing-returns form, and its estimated curvature attaches to the composite time–stock input, a different object from the time-only elasticity α measures. The substitution parameter $\sigma = 0.30$ is hand-set; it is the CES exponent, so the implied elasticity of substitution is $1/(1-\sigma) \approx 1.4$, at the gross-substitutes end of the capital–skill range of Krusell et al. (2000); the σ-sweep spans the complements region and the capital–labor substitution values used in Acemoglu and Restrepo (2022). The factor shares $(\theta_h, \theta_A) = (0.85, 0.15)$ are assumed, with θ_A swept. The planner discount $\delta = 0.40$ per generation corresponds to roughly 3% annually at a 30-year generation length. In the expert survey of the social discount rate by Drupp et al. (2018), that rate sits above the median recommendation of 2%

and within the range experts on average deem acceptable, whose mean bounds are 1.12% and 4.14%. The ±30% sweep spans 2.2% to 4.3% annually, from about the survey's mean recommendation to just above the mean upper bound of the acceptable range. The current AI level $A_0 = 0.5$ is assumed and swept; re-matching the novice-productivity moment at $A_0 = 0.35$ or 0.65 returns $\eta_{\max} = 1.38$ or 0.81, inside the swept range $[0.5, 2.0]$. The codified-tooling floor $\underline{A} = 0.05$, one-tenth of the current AI level, is likewise assumed and swept. It enters production as the installed base that AI capability adds to, bounding the marginal product of the first unit of AI capability. Appendix H reports the cross-section at $\underline{A} \in \{0.02, 0.10\}$.

F Numerical Algorithm

The numerical solution proceeds in six steps:

1. *Expert first-order condition* (closed form). Solve the expert's co-production decision analytically: the first-order condition (Lemma 1) under the multiplicative form gives $t^* = (\alpha\mathcal{W}/\Phi)^{1/(1-\alpha)}$ (Section 6.1), with the corner $t^* = 1$ when $\alpha\mathcal{W} \geq \Phi$. No iteration is required at this step.

2. *Steady state and level normalization.* Solve each occupation's decentralized steady state by iterating the transition at the optimal t^* to its fixed point (tolerance 10^{-13}), and set the training productivity B_j by bisection so that $h^*(A_0) = 1$ (Section 6.2).

3. *State transition.* Iterate the closed-form scalar transition $h_{kj} \to h_{k+1,j}$ via (17) starting from the initial condition h_j^* at the $k = 0$ decentralized steady state with $A_0 = 0.5$, simulating $k = 0, 1, \ldots, 10$ generations. The logistic AI path (J.1), specified in Appendix J, feeds into each generation's transition.

4. *Planner steady state.* Solve the planner's shadow value V_j' by damped fixed-point iteration. Starting from $V_j' = 0$, alternate the corrected steady state at the effective appropriation $\beta(\rho) + \delta V_j'$ with the closed-form envelope update of Lemma 3 in its parametric form (20), at damping one-half. Convergence is to tolerance 10^{-11}, reached in about thirty iterations across the cross-section.

5. *Occupation aggregation.* Compute each occupation's trajectory on its own—the occupations evolve independently—and aggregate via the employment-weighted sum $Y_k = \sum_j n_j Y_{kj}$.

6. *Sensitivity to logistic parameters.* Verify the robustness of the closing-window pattern to the $(A_{\max}, r)$ parameters of (J.1) by repeating the simulation over $A_{\max} \in \{2.0, 3.0, 4.0\}$ and $r \in \{0.30, 0.38, 0.50\}$. In every cell of the grid the employment-weighted welfare gain $\Delta V(k_0)$ declines monotonically in k_0. The occupation-level monotonicity, 100% of occupations and of employment, is verified on the baseline path (Appendix L).

Implementation is in Python (NumPy + pandas). A data module assembles the cross-section: the O*NET expressibility index, the Felten exposure measure, and BLS employment via the public API, with a SOC-vintage crosswalk for the consolidated codes. A single solver module implements the primitives and the routines above. Each figure, table, and robustness exercise, including the numerical verification of Proposition G.1, is produced by a short script calling those solvers. The A_0, θ_A, and $\underline{A}$ perturbations of Table H.2, the curvature bend $\eta(\rho) = \eta_{\max}\rho^{\kappa}$, the substitution tilt $\sigma_j = \sigma + s(\rho_j - \rho_{\text{med}})$, and the level-normalization re-target of Appendix E run as parameter overrides on the same solvers. The replication package documents which script produces each exhibit and the order in which they run. Code and data are available upon request from the corresponding author and will be released publicly upon publication.

G Alternative Production Functional Forms

The CES form (18) adopted in the calibration is a parsimonious specialization of the general production aggregator of Assumption 1. This appendix records how the model's central qualitative results fare under three alternative specializations.

Additive ($\sigma \to 1$)

The unweighted-additive specification $Y_k = (1 - t_k)(h_k + A) + \beta(\rho)h_{k+1}$ corresponds formally to $\sigma = 1$ in the CES limit (with the relaxed normalization $\theta_h = \theta_A = 1$). Production is then unbounded in A even at a vanished stock, so the essentiality that drives Proposition G.1 fails and output rises with AI throughout. In the weighted $\sigma \to 1$ limit of the calibrated aggregator the erosion channel all but shuts down as well: the eroding share falls to 6% of occupations (Table H.1). The additive end of the range is thus close to degenerate on both channels. Our baseline $\sigma = 0.30 \in (0, 1)$ gives intermediate behavior: the stock still erodes at the tacit pole (Proposition 5 applies) but output continues to grow.

Cobb–Douglas ($\sigma \to 0$)

The Cobb–Douglas specification $\Phi = h^{\theta_h} A^{\theta_A}$ is the unit-elasticity special case. Inada conditions hold automatically: $h \to 0 \Rightarrow \Phi \to 0$, so output tracks the stock down at the tacit pole. The qualitative architecture survives under Cobb–Douglas: erosion at the tacit pole, augmentation at the explicit pole, the sign-reversing schedule. But flexibility is reduced (the substitution elasticity is fixed at 1 rather than calibrated), and essentiality now bites. In the $\sigma = 0$ row of Table H.1 erosion deepens and spreads (to 25% of occupations versus 17% at the CES baseline), and output is already declining at the tacit pole at the bottom of the scanned range. This is the output reversal of Proposition G.1(i), operative there because transmission at the tacit pole is intense enough, $\gamma + \alpha = 1.15 > 1$.

Leontief ($\sigma \to -\infty$)

The Leontief production form $\Phi = \min\{h/a_h,\, A/a_A\}$ bottlenecks output at the smaller input scaled by its required intensity. It is the sharp end of the complementarity range: AI is a strict complement of human capital, any erosion of the stock caps output immediately through the min operator, and the output-reversal onset $\bar{A}$ collapses to the point where erosion begins. The Leontief structure is an extreme limit; in the calibrated CES family the onset is set by the interaction of essentiality with the training corner. At $\sigma = 0$ output already declines at the bottom of the scanned range at the tacit pole. At $\sigma = -0.7$ the cornered co-production ($t^* = 1$) props the stock up to interior $\bar{A} \approx 0.2$–0.35 before the collapse (Table H.1).

Summary

Table G.1 compares the four specifications across the model's central qualitative results.

Table G.1: Comparison across Production-Function Specifications

	Additive ($\sigma = 1$)	Cobb–Douglas ($\sigma = 0$)	CES baseline ($\sigma = 0.3$)	Leontief ($\sigma \to -\infty$)
Output reversal $\bar{A}$	no	at scan floor	no	yes (any A)
Erosion at the tacit pole	yes	yes	yes	yes
Closed-form τ^*	yes	yes	yes	no (kinked FOC)
Occupational sign reversal	yes	yes	yes	yes

Note: Qualitative comparison at the tacit pole, where the reversal concentrates. "Output reversal $\bar{A}$" reports whether steady-state output turns down in A: under Cobb–Douglas it is already declining at the bottom of the scanned range $A \in [0.1, 5.0]$ ("at scan floor," the case Table H.1 reports as $\bar{A} \approx 0.1$), while the Leontief bound caps output at any erosion. "Closed-form τ^*" fails under Leontief because the kinked technology has no smooth first-order condition.

The cross-occupation erosion pattern holds across the substitution-elasticity range spanning the calibration ($\sigma \in [-0.3, 0.5]$), degrading only at the substitution extremes (Appendix H): the contest of Proposition 1, not the production-function choice, drives it. Output reversal is conditional on $\sigma \leq 0$; under our baseline $\sigma = 0.30$, the reversal does not occur, and the headline policy results, the sign-reversing schedule and its occupational targeting, survive. The CES form $\sigma \in (0, 1)$ is the appropriate balance: it allows non-trivial substitution between human capital and AI while indexing essentiality by the sign of σ. The swept range $\sigma \in [-0.7, 1]$ spans the capital–skill elasticity estimates of Krusell et al. (2000) and the capital–labor substitution values used in Acemoglu and Restrepo (2022). At the baseline $\sigma = 0.30$, $h \to 0 \Rightarrow \Phi$ stays finite and positive: human capital is non-essential, so output grows even as the tacit stock erodes. Strict essentiality ($\Phi \to 0$ as $h \to 0$), together with the conditional output-reversal, arises only at $\sigma \leq 0$.

Output reversal under gross complements

The sign reversal of the main text concerns the *composition* of human capital: which occupations' stocks AI erodes and which it augments. The *level* of aggregate output, $Y^*(A) = (1 - t^*)\Phi^* + \beta(\rho)h^*(A; \rho)$, is a separate question. In the gross-substitutes baseline ($\sigma \in (0, 1)$) human capital is productive but not essential, so aggregate output rises in A even where the stock erodes, because the production aggregator gains directly from AI. Essential production ($\sigma \leq 0$) overturns this.

Proposition G.1 (Output reversal under gross complements)**.** *Read at the tacit pole, where the augmentation is absent ($\psi \equiv 1$):*

(i) Under Cobb–Douglas production ($\sigma = 0$), if transmission is intense enough that $\gamma + \alpha > 1$, steady-state output vanishes at high AI: $Y^*(A) \to 0$ as $A \to \infty$ while $Y^*(A) > 0$ at every finite A. Output therefore strictly declines over a range of AI levels.

(ii) Under gross complements ($\sigma < 0$), production is bounded by the stock alone, $\Phi(h, A) \leq \theta_h^{1/\sigma} h$ for every A, so $Y^* \leq \left(\theta_h^{1/\sigma} + \beta\right)h^*$ and, the stock being bounded there, output is bounded uniformly in the AI level.

Part (i) is the collapse case: with the augmentation absent, AI raises only the opportunity cost of co-production, so co-production and the stock decay as powers of A. When transmission is intense enough that the stock's erosion outweighs the growing AI input, production decays with them and output falls to zero. The condition holds at the calibrated tacit pole, where $\gamma + \alpha = 1.15$.

Part (ii) is the saturation case: under gross complements AI cannot push output past a fixed multiple of the stock, so output is capped by what the stock supports, however capable the AI. The decline itself under $\sigma < 0$, and its onset $\bar{A}$, are quantitative properties of the calibrated economy, reported across the cross-section in Table H.1. Output already declines at the bottom of the scanned range at $\sigma \in \{0, -0.3\}$ and turns down at interior $\bar{A} \approx 0.2$–0.35 at $\sigma = -0.7$, where the co-production corner ($t^* = 1$) props the stock and delays the collapse.

Output reversal is the strongest form of the Polanyi concern: AI worsens not just the composition but the level of output, hollowing out the human capital its own production complements. It arises only in the essential region and at the most tacit occupations, where $\nu(A; \rho)$ is smallest and erosion fastest. Under the baseline gross-substitutes calibration ($\sigma \in (0, 1)$), output rises throughout; the corrective case there rests on the intergenerational externality, not on an output loss.

Proof of Proposition G.1 (output reversal)

Both parts are read at the tacit pole, where the augmentation is absent ($\psi \equiv 1$, Assumption 5), under the parametric model of Section 6 with the linear appropriation $T(h) = \beta h$; steady-state output is $Y^*(A) = (1 - t^*)\,\Phi^* + \beta h^*$. Production carries the tooling floor, $\Phi = h^{\theta_h}(\underline{A} + A)^{\theta_A}$ in the Cobb–Douglas case; for the large-A asymptotics we write A for $\underline{A} + A$, the floor being asymptotically negligible.

Part (i): the steady-state system. Under Cobb–Douglas production the steady state solves the pair

$$h^* \;=\; \left[B\,(t^*)^{\alpha}\right]^{1/(1-\gamma)}, \qquad t^* \;=\; \min\left\{1,\; \left(\alpha\beta B\,(h^*)^{\gamma}/\Phi^*\right)^{1/(1-\alpha)}\right\},$$

the first the transition (17) at its fixed point with $\psi \equiv 1$, the second the expert's condition (3) with its upper corner. Because $t^* \leq 1$, the stock is bounded along the steady states: $h^* \leq B^{1/(1-\gamma)}$.

Output is positive at every finite A. The stock is positive along the steady states, so $Y^* \geq \beta h^* > 0$.

Output vanishes at high A. For large A the corner fails—its condition $\alpha\beta B(h^*)^{\gamma} \geq \Phi^*$ has a bounded left side and a right side growing in A, so the optimum is interior. Take logs of the interior pair. The transition gives $(1-\gamma)\ln h^* = \ln B + \alpha \ln t^*$; the expert's condition gives $(1-\alpha)\ln t^* = \ln(\alpha\beta B) + (\gamma - \theta_h)\ln h^* - \theta_A \ln A$. Substituting the first into the second collects the co-production terms into $\ln t^*\left[(1-\alpha) - \alpha(\gamma - \theta_h)/(1-\gamma)\right]$, and the bracket equals $D/(1-\gamma)$ with $D \equiv 1 - \alpha - \gamma + \alpha\theta_h$. Solving the two linear equations gives power

laws in A:

$$t^* \propto A^{-\theta_A(1-\gamma)/D}, \qquad h^* \propto A^{-\alpha\theta_A/D}, \qquad \Phi^* \propto A^{\theta_A(1-\alpha-\gamma)/D}.$$

Here $D > 0$ is the stability condition of Lemma 2 under Cobb–Douglas: the equilibrium map's elasticity is $(\gamma - \alpha\theta_h)/(1-\alpha)$, constant, and it lies below one exactly when $D > 0$. The exponents tell the story. AI raises the opportunity cost of co-production without bound, so co-production and with it the stock decay as powers of A. Production $\Phi^* = (h^*)^{\theta_h} A^{\theta_A}$ then pits the growing AI input against the eroding stock, and the stock wins exactly when $\gamma + \alpha > 1$: that inequality makes the exponent $\theta_A(1-\alpha-\gamma)/D$ negative, so $\Phi^* \to 0$. With $t^* \to 0$, $h^* \to 0$, and $\Phi^* \to 0$, every term of Y^* vanishes: $Y^*(A) \to 0$ as $A \to \infty$.

The decline. $Y^*(A)$ is continuous in A, strictly positive at every finite A, and vanishing in the limit. Were it non-decreasing on some ray $[A_1, \infty)$, its limit would be at least $Y^*(A_1) > 0$, a contradiction. So on every such ray output strictly declines somewhere: there is a range of AI levels on which Y^* falls, which is part (i).

Part (ii): the saturation bound. For $\sigma < 0$ the map $x \mapsto x^{1/\sigma}$ is decreasing, so enlarging the CES bracket lowers the aggregate:

$$\Phi(h, A) \;=\; \big(\theta_h h^\sigma + \theta_A(\underline{A} + A)^\sigma\big)^{1/\sigma} \;\le\; \big(\theta_h h^\sigma\big)^{1/\sigma} \;=\; \theta_h^{1/\sigma}\, h \qquad \text{for every } A \ge 0.$$

Hence $Y^* = (1-t^*)\Phi^* + \beta h^* \le \big(\theta_h^{1/\sigma} + \beta\big)h^*$: output is capped by a fixed multiple of the stock, whatever the AI level. At the tacit pole the stock is itself bounded, $h^* \le B^{1/(1-\gamma)}$, since $\psi \equiv 1$ and $t^* \le 1$; output is therefore bounded uniformly in A, which is part (ii). □

H Sensitivity Analyses

This appendix reports the robustness exercises behind Section 6: a sweep over the production substitution parameter σ, $\pm 30\%$ perturbations of the remaining structural parameters $(\delta, \beta, A_{\max}, \gamma, \alpha, \eta_{\max}, A_0, \theta_A)$ together with the tooling floor $\underline{A}$, and a curvature check on the augmentation schedule. It also reports a substitution tilt that lets σ vary with expressibility, a match-free variant that lets instruction bypass the co-production match, and measurement alternatives for the two occupational axes—rank versus cardinal, and acquisition-based indices of expressibility. The reported magnitudes are illustrative outputs of the calibrated quantitative model; the robustness conclusions are about the sign and approximate size of the structural objects, not their exact values. Across the substitution range spanning the calibration, under every structural perturbation, every bend of the schedule, and both measurement conventions, the sign reversal and the single interior crossing survive; the pattern degrades only at the substitution extremes reported below. The reversal and its geography

are what these exercises defend, while the cut's exact location is the calibration-tier object, set by the augmentation schedule's height in the tacit region.

Substitution parameter σ

The substitution parameter σ governs both the qualitative structure of the production aggregator Φ and the conditional output-reversal proposition (Proposition G.1). Figure H.1 shows how the erosion share moves with σ and the cross-occupation pattern of the marginal stock response at the baseline; Table H.1 reports the full sweep.

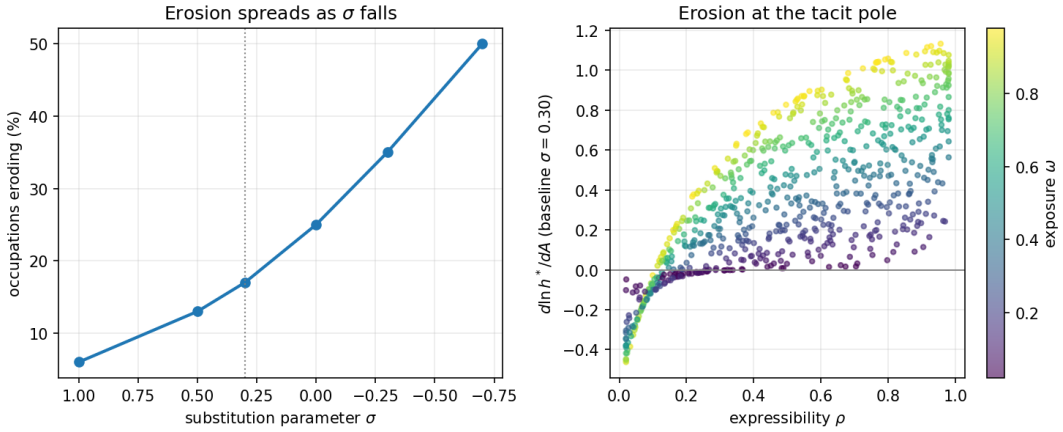


Figure H.1: Sensitivity to the Substitution Parameter σ

Note: Left: the share of occupations whose stock erodes ($dh^*/dA < 0$) as a function of the substitution parameter σ; erosion all but vanishes at the additive limit (6% at $\sigma \to 1$) and spreads as σ falls (50% at $\sigma = -0.7$), with the baseline $\sigma = 0.30$ marked. Right: the marginal stock response $d \ln h^*/dA$ at the baseline $\sigma = 0.30$ across the 682 occupations, colored by AI exposure ω—negative at the tacit pole, positive elsewhere, the single interior crossover of Proposition 5.

Across the cross-section the qualitative pattern is stable: erosion concentrated at the tacit pole, augmentation elsewhere, with a single interior crossover. As σ falls the erosion share rises monotonically: from 6% of occupations at the additive limit $\sigma \to 1$, where erosion all but vanishes, to 50% at strong complements $\sigma = -0.7$. The cut $\bar{\rho}$ moves out from the tacit pole accordingly (Table H.1). The pattern degrades only at the extremes: at $\sigma \to 1$ AI and human capital become near-perfect substitutes, while at $\sigma = -0.7$ co-production corners at $t^* = 1$ over part of the range. The corner props the tacit-pole stock and relocates the cut non-monotonically.

Output reversal (the level of aggregate output falling in A) materializes throughout the essential region $\sigma \leq 0$ and only at the tacit pole, exactly where Proposition G.1 locates it. Output is already declining at the bottom of the scanned range at $\sigma \in \{0, -0.3\}$, and turns down at interior thresholds $\bar{A} \approx 0.2$–0.35 at $\sigma = -0.7$, where the corner delays the

Table H.1: σ-Sensitivity over the Cross-Section

σ	Erosion (occ.)	Erosion (emp.)	Cut $\bar{\rho}$	Output reversal (tacit pole)
+1.0	6%	7%	none	none
+0.5	13%	11%	0.06	none
+0.3 (baseline)	17%	16%	0.10	none
0.0	25%	27%	0.18	$\bar{A} \approx 0.1$
−0.3	35%	37%	0.31	$\bar{A} \approx 0.1$
−0.7	50%	48%	0.11	$\bar{A} = 0.2$–0.35

Note: All quantities at $A_0 = 0.5$. "Erosion (occ.)" and "Erosion (emp.)" are the fractions of occupations and of employment with $dh^*/dA < 0$ under the decentralized comparative static of Proposition 1; "cut $\bar{\rho}$" is the median-exposure force-free cut ($\nu(\bar{\rho}) = \alpha\mu$); "output reversal" is the onset $\bar{A}$ at the tacit pole, where $\bar{A} \approx 0.1$ denotes output already declining at the bottom of the scanned range $A \in [0.1, 5.0]$. As σ falls, erosion deepens and spreads and the cut moves out from the tacit pole; at $\sigma = -0.7$ co-production corners ($t^* = 1$) over part of the range, propping the tacit-pole stock and relocating the cut non-monotonically.

collapse. Across the gross-substitutes range $\sigma \in (0, 1)$, Y^* rises with A in the aggregate and in over nine-tenths of employment. In the deepest tacit-and-exposed occupations (8.9% of employment at the baseline) it declines locally, the stock's fall outweighing the direct gain, while remaining bounded below by the growing AI term. This is decline, never collapse; Proposition G.1 reserves collapse for the essential region.

The aggregate direction follows the tax-side share. The employment-weighted index stays negative, a net subsidy in direction, down to Cobb–Douglas (-0.14 at $\sigma = 0$). It turns positive under gross complements ($+0.16$ at $\sigma = -0.3$; $+2.01$ at -0.7), where erosion reaches half of employment.

Other architecture choices.— Three further architectural perturbations have bounded effects:

- *Common persistence* $\gamma(\rho) = 0.7$. Holding the persistence constant across occupations removes the $1/(1 - \gamma(\rho))$ amplification that steepens the schedule. By the force-free cut $\nu(\bar{\rho}) = \alpha\mu$ of Proposition 6, it leaves the cut-point and the sign reversal unchanged; only the dispersion of $|\tau^*(\rho)|$ across occupations attenuates.

- *Cobb–Douglas vs. CES.* The $\sigma = 0$ row of Table H.1 is the Cobb–Douglas case; the sign pattern of the stock response is identical to the CES baseline.

- *Expert altruism* $\alpha_b \in [0, 0.15]$. Appendix B shows the intergenerational wedge retains $\geq 62.5\%$ of its baseline magnitude even at the illustrative bound $\alpha_b = 0.15$ for non-kin professional altruism; the damage index scales by $1 - \alpha_b/\delta$.

Other structural parameters

We test the robustness of the headline results: the force-free cut $\bar{\rho}$, the sign reversal of the schedule, and the planner-versus-decentralized welfare gap. We perturb the structural parameters $(\delta, \beta, A_{\max}, \gamma, \alpha, \eta_{\max}, A_0, \theta_A)$ by $\pm 30\%$, holding the remaining parameters at baseline. We include the augmentation slope $\eta_{\max}$ deliberately: it enters the reach $\nu = \eta_{\max}\rho/(1+A)$ directly, so the cut is roughly proportional to $\alpha\mu/\eta_{\max}$, yet $\eta_{\max}$ is pinned by a single field study (Appendix E). We include the AI level A_0 and the AI share θ_A for the parallel reason: both are assumed rather than measured, and both enter the sign condition—A_0 through the reach and the compression jointly, θ_A through the production share s_h in the compression μ. The learning-by-doing elasticity α, likewise assumed, multiplies the compression in the cut condition itself. Table H.2 summarizes the responses; the baseline row carries the headline values of Section 6 (Figure 4).

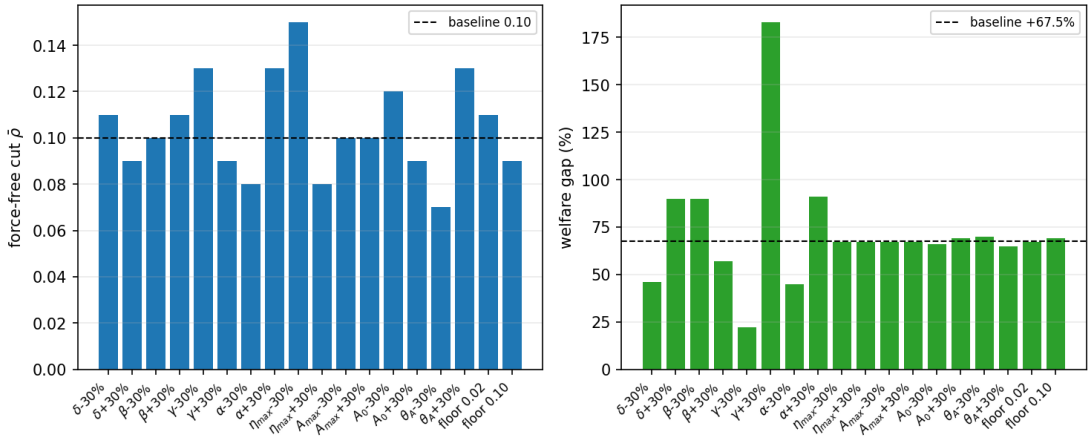


Figure H.2: Headline Objects under $\pm 30\%$ Structural Perturbations

Note: The force-free cut $\bar{\rho}$ (left) and the employment-weighted welfare gap (right) under $\pm 30\%$ perturbations of each structural parameter; dashed lines mark the baseline ($\bar{\rho} = 0.10$, gap $+67.5\%$). The cut stays in $[0.07, 0.15]$, moving most with the contest parameters $\eta_{\max}$, θ_A, α, and A_0, less with γ, and barely with β, δ, or $A_{\max}$. The sign reversal is preserved throughout, while the welfare gap rises with persistence and patience.

Three patterns emerge, each tracing directly to the force-free structure of Proposition 6. First, the cut-point $\bar{\rho}$ is governed by $\nu(\bar{\rho}) = \alpha\mu$ and stays interior under every perturbation, moving within $[0.07, 0.15]$ across the $\pm 30\%$ sweep. The parameters that sit in the contest move it most—$\eta_{\max}$ and α directly, θ_A through the production share in μ, A_0 through both sides. The persistence γ shifts it through the compression. Perturbing β, δ, or $A_{\max}$ leaves it nearly fixed. The parameters β and δ enter only the index's positive prefactor and touch the cut solely by relocating the equilibrium at which the contest is read, a channel the sweeps bound at ± 0.01. The ceiling $A_{\max}$ shapes only the transition path and leaves the comparative

Table H.2: Extended Sensitivity: Structural Perturbations and Specification Variants

Perturbation	Value	Cut $\bar{\rho}$	Tax-side (emp.)	Welfare gap
Baseline	(calibration)	**0.10**	**13.16**%	**+67.5**%
Planner discount factor δ				
$\delta - 30\%$	$\delta = 0.28$	0.11	13.44%	+46%
$\delta + 30\%$	$\delta = 0.52$	0.09	11.10%	+90%
Appropriation coefficient $\beta(\rho)$ (level shift)				
$\beta \times 0.70$	lower appropriation	0.10	11.15%	+90%
$\beta \times 1.30$	higher appropriation	0.11	13.41%	+57%
AI saturation ceiling $A_{\max}$				
$A_{\max} - 30\%$	$A_{\max} = 2.10$	0.10	13.16%	+67%
$A_{\max} + 30\%$	$A_{\max} = 3.90$	0.10	13.16%	+67%
Transmission persistence $\gamma(\rho)$ (multiplicative shift)				
$\gamma \times 0.70$	persistences reduced	0.13	15.03%	+22%
$\gamma \times 1.30$ (capped at 0.85)	persistences raised	0.09	10.30%	+183%
Learning-by-doing elasticity α				
$\alpha - 30\%$	$\alpha = 0.21$	0.08	8.57%	+45%
$\alpha + 30\%$	$\alpha = 0.39$	0.13	15.52%	+91%
Augmentation slope $\eta_{\max}$ (co-determinant of the cut; pinned by one field study)				
$\eta_{\max} - 30\%$	$\eta_{\max} = 0.70$	0.15	17.60%	+67%
$\eta_{\max} + 30\%$	$\eta_{\max} = 1.30$	0.08	8.57%	+67%
AI level A_0 (assumed)				
$A_0 - 30\%$	$A_0 = 0.35$	0.12	13.76%	+66%
$A_0 + 30\%$	$A_0 = 0.65$	0.09	11.09%	+69%
AI production share θ_A (assumed; $\theta_h = 1 - \theta_A$)				
$\theta_A - 30\%$	$\theta_A = 0.105$	0.07	6.79%	+70%
$\theta_A + 30\%$	$\theta_A = 0.195$	0.13	17.40%	+65%
Codified-tooling floor $\underline{A}$ (assumed)				
$\underline{A} = 0.02$	lighter floor	0.11	14.79%	+67%
$\underline{A} = 0.10$	heavier floor	0.09	8.86%	+69%
Augmentation curvature $\eta(\rho) = \eta_{\max}\rho^{\kappa}$, re-anchored to the empirical moment				
$\kappa = 0.5$	concave; $\eta_{\max} = 0.99$	0.01	0.44%	+67%
$\kappa = 2$	convex; $\eta_{\max} = 1.03$	0.34	36.86%	+67%
Substitution tilt $\sigma_j = \sigma + s(\rho_j - \rho_{\text{med}})$: expressibility reaches production				
$s = +0.4$	σ_j rising in ρ, 0.11 to 0.49	0.13	16.74%	+68%
$s = -0.4$	σ_j falling in ρ, 0.49 to 0.11	0.07	7.26%	+67%
Match-free instruction: share λ of the augmentation bypasses the match				
$\lambda = 0.25$	quarter match-free	0.12	13.80%	+67%
$\lambda = 0.50$	half match-free	0.13	15.72%	+66%

Note: Headline objects recomputed over the cross-section at $A_0 = 0.5$. $\bar{\rho}$ is the median-exposure force-free cut $\nu(\bar{\rho}) = \alpha\mu$; "tax-side emp." is the employment share with $\tau^* > 0$; the welfare gap is the employment-weighted discounted planner-versus-decentralized improvement. The sign reversal (the tax direction at the tacit pole, the subsidy direction at the explicit pole) is preserved in every cell.

statics at A_0 untouched.

Second, because $\gamma(\rho)$, $\beta(\rho)$, σ, and δ enter only the strictly positive magnitude prefactor, the *occupational* sign change (the tax direction at the tacit pole, the subsidy direction at the explicit pole) persists across all perturbations. What moves is the dispersion of $|\tau^*(\rho)|$.

Third, the discounted planner-versus-decentralized welfare gap remains positive throughout and rises steeply with persistence and patience. The largest sensitivities come from δ and γ, both of which scale the recursive-transmission multiplier $1/(1-\delta\gamma(\rho))$. The gap is largely insensitive to $A_{\max}$, which shapes the path of A but not the within-period contractual wedge.

The employment-weighted baseline welfare gap of +67.5% is therefore robust in sign and order of magnitude: it ranges from +22% (lower persistence) to +183% (higher persistence) across the sweep, a more patient planner or higher persistences raising it. Every cell supports the conclusion that occupation-targeted intervention substantially outperforms both no-policy and uniform-policy alternatives.

Measurement, the augmentation schedule, and specification variants

The checks that follow probe the cut from every direction the calibration leaves open: how the two occupational axes are measured, how the augmentation schedule is drawn between its anchors, how production substitution tilts with expressibility, and how tightly the technology ties instruction to the match.

Measurement. Expressibility and exposure enter the calibration as rank-normalized indices; we re-derive the headline using their cardinal (min–max-scaled) values instead. The force-free cut is unchanged: $\bar{\rho} = 0.10$ under rank, cardinal-ρ, and cardinal-ρ-and-ω measurement alike. The employment-weighted index stays negative ($\bar{\tau}$ from -0.27 to -0.34), expressibility and exposure stay nearly independent (rank correlation 0.06, cardinal 0.08), and the sign reversal holds throughout (Table H.3). Cardinalization lowers the tax-side employment share substantially (13.16% to about 4%), because the raw indices cluster occupations more tightly toward the middle of the spectrum than ranks do, so fewer occupations sit below the cut. The cut and the reversal are untouched.

Acquisition-based measurement. The task-content index reads expressibility off the work itself; O*NET's education and training files let us read it off how the occupation's skill is acquired. From the reported distributions of required schooling, related work experience, and on-site and on-the-job training we build two alternatives. The *classroom share* of total preparation is schooling years over schooling plus experience plus training: knowledge acquired

Table H.3: Measurement Robustness

Measurement	corr(ρ, ω)	Tax-side (occ.)	Tax-side (emp.)	$\bar{\tau}$	Cut $\bar{\rho}$
Rank ρ, rank ω (baseline)	0.06	13.78%	13.16%	−0.27	0.10
Cardinal ρ, rank ω	0.07	6.2%	4.4%	−0.31	0.10
Cardinal ρ & ω	0.08	5.9%	4.3%	−0.34	0.10
Classroom share (acquisition)	0.69	16.1%	8.1%	−0.32	0.10
Formation length (acquisition)	−0.29	11.4%	11.3%	−0.26	0.10

Note: Headline objects under rank versus cardinal measurement of expressibility and exposure, and under the two acquisition-based indices described in the text, at $A_0 = 0.5$. The schedule and its cut are invariant to the measure by construction—a measure only relocates occupations along the ρ axis—and the sign reversal and the net subsidy persist in every row.

by telling over knowledge acquired by doing. The *formation length* is the experience-and-training years themselves, longest at the tacit end.

Both deliver the structural results (Table H.3): the schedule reverses sign with a single interior crossing at the same cut. The tax-side employment share stays within the structural-sweep range (8.1% and 11.3% against 13.16% at baseline), and the employment-weighted index remains negative (−0.32 and −0.26). Neither reproduces the baseline's occupational geography: the tax-side sets overlap the baseline's by 4% and 14% of its tax-side employment, because each index reads a different facet of the concept.

The classroom share measures the entry channel and is strongly exposure-correlated (+0.69, against +0.06 for the task-content index): schooling intensity and measured AI exposure both track cognitive content. The tacit-and-exposed corner is therefore nearly empty under it. The classroom share also places the credentialed professions at the codifiable pole on the strength of their schooled entry. Yet their reported post-schooling formation is among the longest, six to eight years for lawyers and surgeons. The formation length is the acquisition-side counterpart of the time-to-mastery proxy that disciplines γ, and it restores those professions to the tacit side. But it conflates how much there is to learn with how far it can be told, relocating the deepest tax-direction values to the longest-experience occupations (judges, arbitrators, executives).

The task-content index observes the daily work the novice must learn to perform—the object transmission carries in the model—and is the only one of the three nearly orthogonal to exposure. The occupational naming of Section 6—which occupations sit deepest in the tax-side region and in the corner—is accordingly a property of the task-content measure. The reversal, the cut, the tax-side share, and the net-subsidy stance carry across all three.

Augmentation slope. Because $\eta_{\max}$ enters the reach directly, $\nu = \eta_{\max}\rho/(1+A)$, the cut is roughly proportional to $\alpha\mu/\eta_{\max}$; $\eta_{\max}$ is therefore the parameter the cut moves with most.

Figure H.3 traces the cut and the tax-side share as $\eta_{\max}$ ranges over $[0.5, 2.0]$: the cut falls from 0.22 to 0.05, yet stays interior and the reversal is preserved throughout. We include $\eta_{\max}$ in the $\pm 30\%$ structural sweep above for this reason.

Augmentation curvature. The field estimate disciplines only the schedule's height near the codifiable pole, and the Polanyi floor pins $\eta(0) = 0$; between the two anchors the linear interpolation is a choice. We therefore bend the schedule, $\eta(\rho) = \eta_{\max}\rho^{\kappa}$ with $\eta_{\max}$ re-set so the experimental moment is re-hit exactly. The cut moves with the tacit-region height: $\kappa = 0.5$ (concave, more reach at low ρ) gives $\bar{\rho} = 0.01$ and a 0.44% tax-side employment share; $\kappa = 2$ (convex, less) gives 0.34 and 36.86%. The welfare gap holds at +67% and the employment-weighted index remains negative (-0.54 and -0.08) in both (Table H.2). The cut responds to the schedule only through its value near the cut itself. The slope and curvature sweeps therefore bound the same object from two directions: the reach available to the most tacit occupations. The reversal, the single interior crossing, and the tacit corner's positive index survive every case.

Substitution tilt. Expressibility drives the baseline through transmission alone, the reach $\eta(\rho)$ and the appropriation $\beta(\rho)$, while production carries a single substitution parameter. The conjecture that codifiability should also govern how elastically AI substitutes for human capital in the doing itself has two versions. The level version is measured: how much of an occupation's production AI reaches is the exposure ω_j, and it is nearly independent of expressibility in the data (Figure 3). No moment we observe pins the elasticity version, so we sweep it. The tilt $\sigma_j = \sigma + s(\rho_j - \rho_{\text{med}})$, clipped to $[0.05, 0.95]$, is anchored at the cross-section median so the median occupation keeps the baseline curvature. At $s = +0.4$ AI is a closer production substitute for codifiable knowledge, σ_j rising from 0.11 at the tacit pole to 0.49 at the codifiable one; $s = -0.4$ is the reverse.

The tilt exposes a feature of the mechanism. At the calibrated equilibrium AI is the scarce production input, so the compression falls, rather than rises, with σ: complementarity curvature is what makes the marginal AI unit raise the price of co-production. Tilting substitutability up with expressibility therefore strengthens the schedule at both poles. The tacit pole meets a lower σ_j, and its index roughly doubles, +0.26 to +0.54 at $\rho = 0.05$ and median exposure; the codifiable pole meets a higher one, and its index deepens, -0.35 to -0.37 at $\rho = 0.90$.

The tax-side employment share rises to 16.74% with the cut at 0.13, and the reverse tilt compresses them to 7.26% and 0.07 (Table H.2). A stress tilt $s = +0.8$, spanning $\sigma_j \in [0.05, 0.68]$, pushes the pair to 20.94% and 0.17. The single interior crossing, the negative employment-weighted index ($\bar{\tau}$ between -0.29 and -0.21), and the welfare gap (+67% to +68%) survive every tilt. A direct ρ-channel into production relocates the cut within the

ranges the structural sweep already spans—exactly the latitude the contest hypothesis of Propositions 5 and 6 grants the compression.

Match-free instruction. In the calibrated class the instruction operates through the match: ψ multiplies co-production, so formation vanishes without it. We let a share of the instruction bypass the match, replacing the transmission bracket $t^{\alpha}\psi$ with $t^{\alpha}\psi_m + \lambda\, t_0^{\alpha}(\psi - 1)$. Here $\psi_m = (1+A)^{(1-\lambda)\eta(\rho)}$ carries the match's part, and the second term is AI-borne instruction available without co-production, scaled by the occupation's baseline co-production intensity t_0 and vanishing at $A = 0$ and at the Polanyi floor. The anchor moment is rehit to within a point. The appropriation keeps its Becker reading—the match remains the employment relationship through which the novice contributes—so the exercise isolates the technology.

Instruction that substitutes for the match widens the tax-side region rather than narrowing it. The match's own augmentation weakens to $(1-\lambda)\eta$, so the reach defending the co-production margin falls, while the bypass builds the stock without defending that margin. At $\lambda = 0.25$ the cut moves to 0.12 and the tax-side employment share to 13.80%; at $\lambda = 0.50$, to 0.13 and 15.72% (Table H.2); at a stress $\lambda = 0.75$, to 0.16 and 17.99%. All lie within the ranges above. The single interior crossing, the negative employment-weighted index ($\bar{\tau}$ from -0.22 to -0.14), and the welfare gap (+65% to +67%) survive every case.

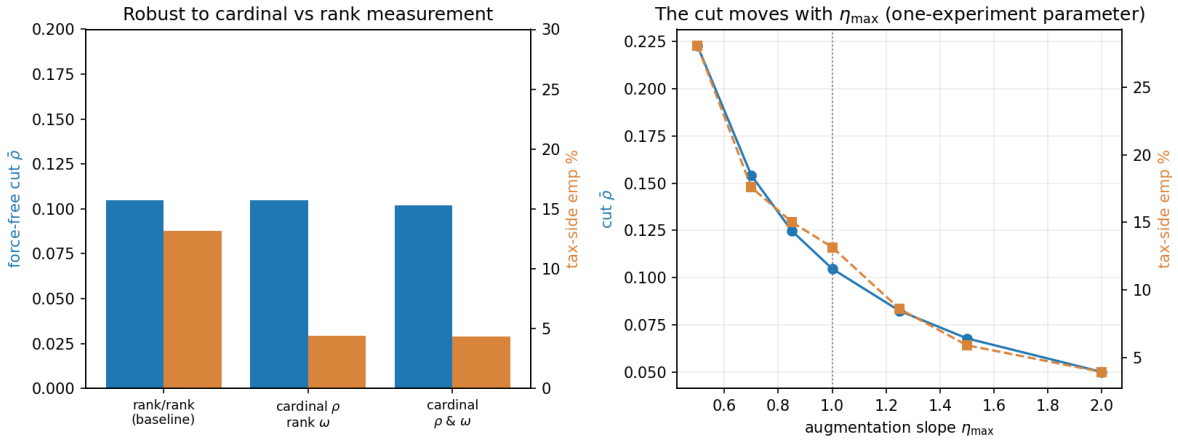


Figure H.3: Measurement and Augmentation-Slope Robustness

Note: Left: the force-free cut $\bar{\rho}$ and tax-side employment share under rank versus cardinal measurement of expressibility and exposure—the cut is invariant. Right: the cut and tax-side share as the augmentation slope $\eta_{\max}$ varies over $[0.5, 2.0]$; the cut moves substantially (it scales as $\alpha\mu/\eta_{\max}$) but stays interior, and the sign reversal holds.

I Decomposition of the Welfare Gap

The intergenerational wedge predates AI. It is positive already at $A = 0$ (Section 4.1): an expert does not appropriate the value her training creates beyond the next generation, whatever the technology. AI does not create the wedge; it raises its quantitative bite. We decompose the employment-weighted planner-versus-decentralized welfare gap at the calibrated AI level into two parts (Figure I.1). The *baseline* component is the gap that would remain at $A = 0$; the *AI-marginal* increment is the part added as A rises from 0 to its current level.

Most of the gap predates AI: the baseline wedge is 79% of the current gap, the AI-marginal slice 21%. The split turns over across the spectrum. In the tacit occupations the gap is largest in level—high persistence compounds each generation's undertraining through the multiplier $1/(1-\delta\gamma(\rho))$—yet there it is almost entirely the pre-existing wedge, AI adding only 12%. In the codifiable occupations the level gap is smaller, but AI accounts for about two-fifths of it, 39%, because there AI augments the stock and so raises the shadow value of the capital the externality already leaves underprovided. The reading for policy is that the corrective burden divides. The instruments that predate AI—training subsidies, licensure, credentialing, residencies—carry most of the load in the tacit occupations, where the baseline wedge dominates. AI's marginal contribution to the distortion falls disproportionately on the codifiable occupations, the same ones on the subsidy side of the schedule.

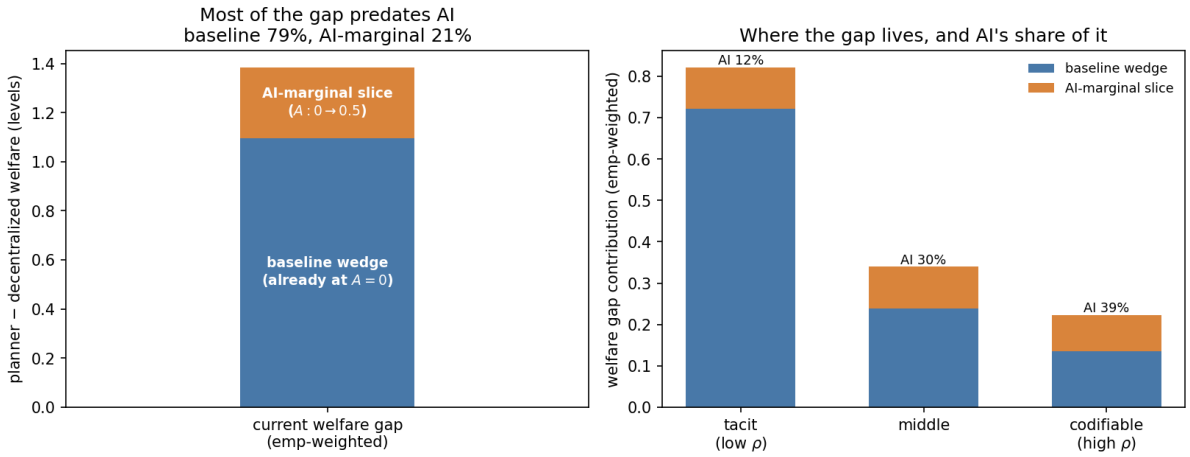


Figure I.1: Baseline and AI-Marginal Components of the Welfare Gap

Note: Decomposition of the planner-versus-decentralized welfare gap. Left: the employment-weighted current gap splits 79% baseline (the wedge already present at $A = 0$) and 21% AI-marginal ($A : 0 \to 0.5$). Right: by expressibility tercile—the gap lives in the tacit occupations, and there it is overwhelmingly the baseline wedge (AI's share 12%), while AI's share rises to 39% in the codifiable occupations.

J Baseline Path, Counterfactual Trajectories, and Welfare Comparison

This appendix reports the transitional dynamics that complement the comparative statics of Section 6. Under the calibration of Table 1, we simulate the no-policy and planner-allocation trajectories of (h_{kj}, Y_{kj}) for $k = 0, 1, \ldots, 10$ across the cross-section, each occupation starting from its decentralized steady state at $A_0 = 0.5$ and the exogenous AI level of Section 3.1 growing along the logistic path

$$A_k = \frac{A_{\max}}{1 + \left[(A_{\max} - A_0)/A_0\right] e^{-rk}}, \tag{J.1}$$

which bounds A_k in the empirically relevant range while preserving monotonicity, and hence the direction of every comparative-static result along it. Its parameters—the ceiling $A_{\max} = 3.0$, the adoption rate $r = 0.38$ per generation, and a generation length of 30 years—are assumed, chosen to bound the path rather than identified from data. The cross-section comparative statics of Section 6 are read at the fixed level $A_0 = 0.5$ and are invariant to the path: the $\pm 30\%$ perturbation of $A_{\max}$ in Table H.2 moves neither the force-free cut nor the welfare gap.

No-policy baseline path

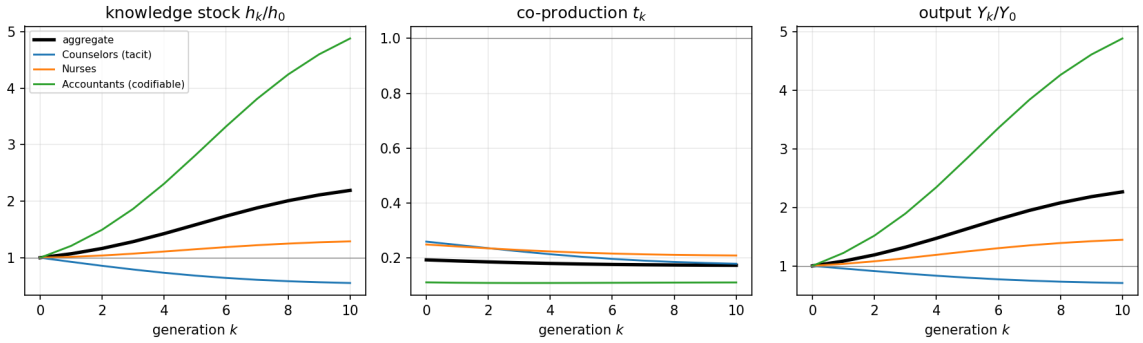


Figure J.1: No-Policy Trajectories along the AI Path

Note: Ten generations along the logistic AI path. The black line is the employment-weighted aggregate; colored lines are illustrative occupations. Left: the knowledge stock h_k/h_0 rises in aggregate but falls at the tacit pole (counselors)—the single interior crossover of Proposition 5. Center: co-production t_k compresses as AI raises its opportunity cost. Right: output Y_k/Y_0 rises in aggregate; at the tacit pole it declines with the stock beneath it.

Two patterns confirm the predictions of Sections 3–5. *(i) The single interior crossover*: over the transition the stock erodes in the 13% of occupations at the tacit pole and rises in

Table J.1: Baseline Transition over Ten Generations

	Knowledge stock Δh	Output ΔY
Employment-weighted aggregate	+119%	+156%
Occupations with eroding stock	13% (the tacit pole)	

Note: Employment-weighted percentage change from the $k = 0$ decentralized steady state along the logistic AI path ($A_k : 0.50 \to 2.70$). The stock rises in aggregate because most employment sits in the augmenting, codifiable occupations; it erodes in the 13% of occupations at the tacit pole—the single interior crossover of Proposition 5.

the rest, exactly the sign pattern of Proposition 5, the crossover sitting near the calibrated cut $\bar{\rho} \approx 0.10$. *(ii) Output continues to grow*: the employment-weighted stock rises +119% and output +156% from $k = 0$ to $k = 10$ under the gross-substitutes calibration ($\sigma = 0.30$), because the production aggregator gains directly from AI and most employment sits in the augmenting occupations, even as the tacit pole erodes. The erosion is therefore a compositional cross-occupation phenomenon, not an output collapse—the latter arises only under gross complements (Proposition G.1).

Counterfactual trajectories under the Pigouvian planner allocation

Figure J.2 overlays the planner allocation on the no-policy path. Internalizing the transmission chain raises co-production in every occupation, and the restored formation compounds down the chain: the aggregate stock and output pull away from the no-policy path generation by generation.

Welfare comparison

The planner allocation, the corrected steady state the occupation-specific rates support, improves on no policy by an employment-weighted +67.5%. The gap rises steeply with persistence—largest in the high-persistence tacit occupations, where the intergenerational externality is most severe, and smallest in the low-persistence codifiable ones, tracing the multiplier $1/(1 - \delta\gamma(\rho))$ of Lemma 3. The employment-weighted average damage index is a net subsidy, $\bar{\tau} = -0.27$, so a uniform instrument set to the employment-weighted average would subsidize AI on net, yet the occupation-specific index values span $[-0.70, +0.62]$ and cross zero. The dispersion of τ^* around its mean is what a uniform-rate alternative misallocates, and the sign change makes the misspecification qualitative, not merely a matter of magnitude.

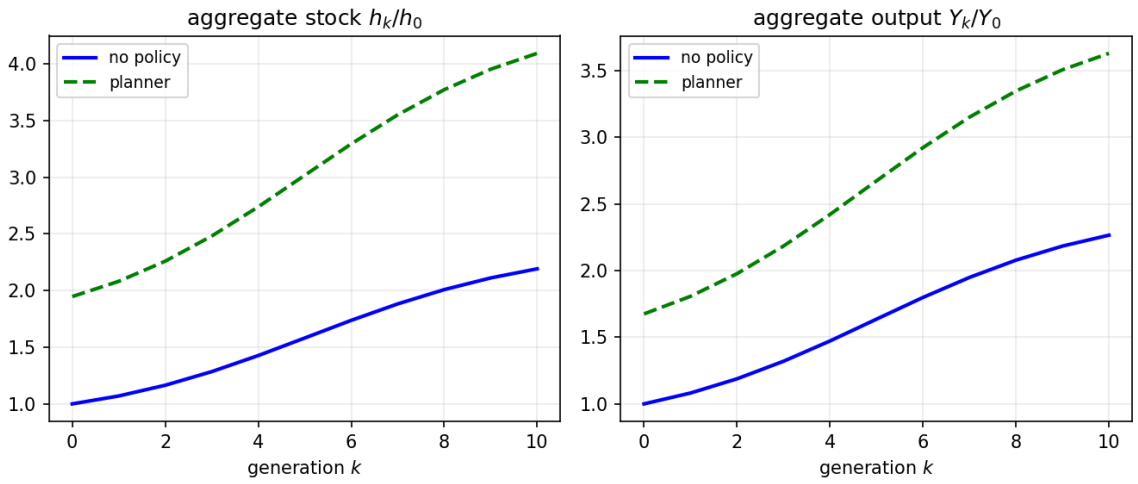


Figure J.2: No-Policy versus Planner Trajectories

Note: Aggregate trajectories under no policy (solid) versus the Pigouvian planner allocation (dashed) over $k = 0, \ldots, 10$, employment-weighted. The planner raises co-production, restoring the knowledge stock (left) and raising output (right), with the largest restoration in the high-persistence tacit occupations where the intergenerational wedge is largest.

K Mistargeted Policy: Allocation Details, Price Robustness, and the Transition

Section 6.5 prices AI-rate policies on the adoption margin and reports the welfare ranking. This appendix records the allocations behind those optima, checks the ranking across AI prices, and carries the exercise through the transition: the transition-inclusive welfare of the steady-state-sized policies, and the rates re-solved for that criterion.

Allocation details

At the baseline price $p = 0.8$, adoption is interior for 82.66 percent of employment, at exactly zero use for 10.75, and at the exposure frontier for 6.59. The targeted optimum's tax rates, up to +0.20, are the smallest on the grid supporting each zero-use corner; their welfare coincides with the ban's, and ties resolve to the finite rate. Subsidies are reported as the smallest rate reaching full adoption, beyond which a larger rebated subsidy moves no margin. Use stays interior at the optimum for 0.01 percent of employment, so the interior formula of Proposition 4 prices only the band between the corners. Under the co-production subsidy of Section 4.3, the first-best benchmark all shares are measured against, the subsidized expert's conditions coincide with the planner's, and the coincidence extends to the full-adoption corner most codifiable occupations reach.

The automation model's positive rates, up to +0.10, cover 3.95 percent of employment,

with a subsidy nowhere. The remaining occupations select a zero rate: use is already at a private corner, or the gain the automation model perceives is too small for the rate grid to distinguish from zero.

Robustness to the AI price

The targeted policy's lead holds at prices $p \in \{0.6, 0.8, 1.0\}$: it recovers 6.9 to 11.3 percent of the first-best gain, one and a half to four times the best uniform rate. The automation model's rates recover essentially nothing at any price (-1.2% to $+0.1\%$).

The transition criterion

The transition starts from the current state: each occupation begins at its no-policy steady state, the policy switches on, and the allocation re-optimizes each period at the baseline AI level. The benchmark is the co-production subsidy held at its steady-state rate, which corrects the co-production margin in the period the distortion arises and bounds the transition first best from below. On the transition-inclusive criterion, every steady-state-sized AI-price policy loses welfare at the calibrated generational discount, except the automation model's rates, which move almost no use and return essentially zero ($+0.1\%$). Against the benchmark's gain, the steady-state-optimal targeted schedule returns -18%, two expressibility brackets -26%, the uniform subsidy -27%, exposure brackets -29%, the ban -3%.

The transition losses have one cause: adjustment costs arrive a generation before formation gains. The subsidized side pays the resource cost of expanded use immediately and collects the re-supplied stock over the following generations; the taxed side forgoes surplus immediately and collects the protected stock as slowly. At $\delta = 0.40$ per generation the near term wins. Rates sized for the steady state therefore overshoot the transition.

Solving the transition problem over the same policy space confirms it. Prohibition again never strictly improves on the finite schedule. The signs agree with the steady-state schedule over 87.68 percent of employment; the taxed region widens slightly, since impatience discounts the subsidies' delayed formation gains. The rates are no larger for 96.50 percent of employment, and the employment-weighted mean magnitude falls from 0.26 to 0.12. The re-sized schedule recovers $+3\%$ of the benchmark gain where the steady-state-sized one loses 18%. At the model's generational frequency, a two-phase schedule—one first-generation rate, one rate thereafter—adds negligible welfare over the best constant rate.

The disagreements sit at the boundary, on a nearly flat stretch of the welfare surface, and run one way. No occupation crosses from tax to subsidy; half of the disagreeing employment settles at a zero rate compatible with either sign; for the rest, subsidies near the cut shrink

to small taxes. Imposing the steady-state signs on every disagreeing occupation forfeits a hundredth of a percent of the benchmark gain. What relocates is the cut, not the poles, and the welfare content of the sign concentrates where the rates are large.

The result is the corrective mirror of the transitional robot taxes of Guerreiro et al. (2022): their distributional motive fades with the transition and the tax vanishes, while our formation externality persists and the rate shrinks instead. As the planner grows patient, the transition's weight in discounted welfare vanishes and the two criteria merge.

What survives both criteria is the geography: the sign pattern carries to the transition-optimal schedule itself. Among the steady-state-sized rate schedules, the targeted one loses least and the exposure brackets most; the ban and the automation model's rates move welfare little either way, reflecting only how little use the baseline price leaves to suppress.

L Closing-Window: Extended Numerical Pattern

This appendix reports the intervention-date pattern underlying the closing-window result, simulated over the full 682-occupation cross-section. We forward-simulate the decentralized transition along the logistic AI path of (J.1) and overlay an η-investment policy that raises the augmentation slope $\eta_{\max}$ from its baseline value of 1.0 to 1.25 at intervention generation $k_0 \in \{0, 2, 4, 6, 8, 10\}$. For every occupation we trace the discounted welfare gain $\Delta V(k_0)$ and the terminal knowledge stock $h_{10}(k_0)$ relative to no policy (Figure L.1). The magnitudes are illustrative—raising $\eta_{\max}$ by a quarter lifts the employment-weighted terminal stock by 43% at $k_0 = 0$, and the qualitative pattern is the content.

Both metrics decline monotonically in k_0 in every occupation of the cross-section, 100% of occupations and of employment, so $\partial V/\partial k_0 < 0$ throughout. The corrective need is front-loaded, and earlier intervention strictly dominates. The two metrics close at different speeds. The discounted-welfare gain is steeply front-loaded: only about 9% of the $k_0 = 0$ gain survives to $k_0 = 4$. The reach the policy amplifies, $\nu(\rho) = \eta(\rho)/(1 + A_k)$, declines as AI matures, so a late investment acts on an already-saturated reach. The terminal-stock gain instead holds nearly in full through $k_0 \approx 6$ and falls away only later, a late start leaving too few generations for the higher augmentation to compound. The $k_0 = 10$ endpoint is the no-policy boundary, an intervention at the horizon acting on no generation within it.

Across occupations the gain concentrates where AI can teach. The terminal-stock gain rises with expressibility (rank correlation +0.64 with ρ at $k_0 = 0$). The welfare gain is largest in the codifiable occupations (credit authorizers, budget analysts, accountants) and negligible at the tacit pole (choreographers, fitness trainers), where the Polanyi floor holds $\eta(\rho) = \eta_{\max}\rho$ near zero whatever the investment. The persistence multiplier $1/(1 - \gamma)$ is

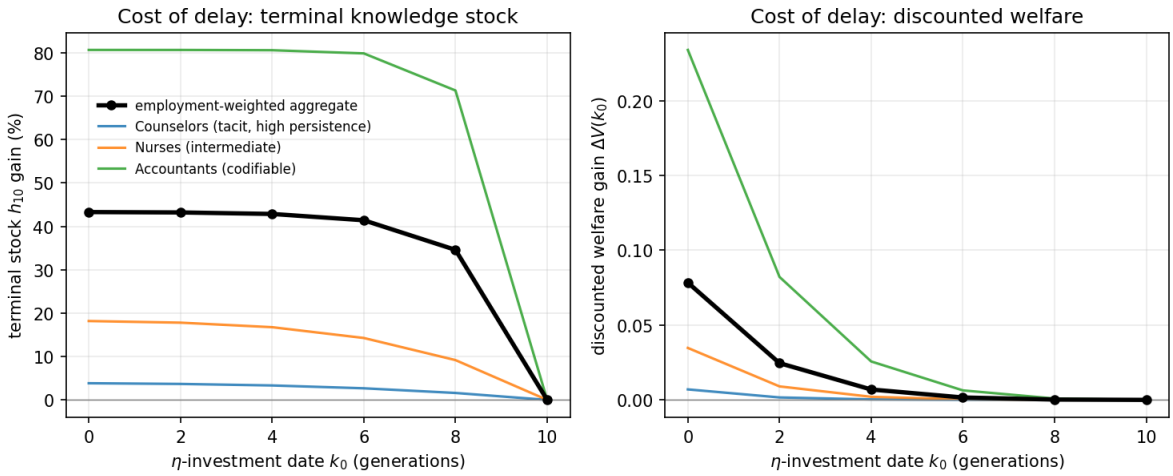


Figure L.1: Closing-Window Dynamics

Note: The η-investment counterfactual across the 682-occupation cross-section. The black line is the employment-weighted aggregate; colored lines are illustrative occupations. Left: terminal knowledge stock h_{10} against the η-investment date k_0 (percentage gain over no policy). Right: discounted welfare gain $\Delta V(k_0)$, valued at the present. Both decline monotonically in k_0 in every occupation, so earlier intervention strictly dominates; the welfare gain is steeply front-loaded, while the terminal-stock gain holds nearly in full through $k_0 \approx 6$ and falls away only later. The gain is largest in the codifiable occupations, where AI can teach, and negligible at the tacit pole (the Polanyi floor).

largest at the tacit pole, but it amplifies only what the reach delivers, and at the Polanyi floor there is nothing to amplify: the ordering stands. The η-investment is thus the mirror of the schedule's tax side: the tax side marks the tacit pole AI erodes, the investment builds the codifiable pole AI augments, and both are front-loaded.

References

Acemoglu, D. and D. Autor (2011). Skills, tasks and technologies: Implications for employment and earnings. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 4B, pp. 1043–1171. Amsterdam: Elsevier.

Acemoglu, D. and P. Restrepo (2022). Tasks, automation, and the rise in u.s. wage inequality. *Econometrica 90*(5), 1973–2016.

Becker, G. S. and K. M. Murphy (1988). The family and the state. *Journal of Law and Economics 31*(1), 1–18.

Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *Journal of Political Economy 75*(4), 352–365.

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

Brynjolfsson, E., T. Mitchell, and D. Rock (2018). What can machines learn, and what does it mean for occupations and the economy? *AEA Papers and Proceedings 108*, 43–47.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica 78*(3), 883–931.

Dell'Acqua, F., E. McFowland III, E. R. Mollick, H. Lifshitz, K. C. Kellogg, S. Rajendran, L. Krayer, F. Candelon, and K. R. Lakhani (2026). Navigating the jagged technological frontier: Field experimental evidence of the effects of artificial intelligence on knowledge worker productivity and quality. *Organization Science 37*(2), 403–423.

Drupp, M. A., M. C. Freeman, B. Groom, and F. Nesje (2018). Discounting disentangled. *American Economic Journal: Economic Policy 10*(4), 109–134.

Eloundou, T., S. Manning, P. Mishkin, and D. Rock (2024). GPTs are GPTs: Labor market impact potential of LLMs. *Science 384*(6702), 1306–1308.

Ericsson, K. A., R. T. Krampe, and C. Tesch-Römer (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review 100*(3), 363–406.

Felten, E., M. Raj, and R. Seamans (2021). Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal 42*(12), 2195–2217.

Guerreiro, J. a., S. Rebelo, and P. Teles (2022). Should robots be taxed? *Review of Economic Studies 89*(1), 279–311.

Krusell, P., L. E. Ohanian, J.-V. Ríos-Rull, and G. L. Violante (2000). Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica 68*(5), 1029–1053.

Webb, M. (2020). The impact of artificial intelligence on the labor market. Technical report, Stanford University working paper.