

The Valley at the Frontier: Who Gains from Generative AI

Frontier, Leverage and the Distribution of Productivity Gains

Jorge Menéndez-Pidal

Project Frontier, Madrid

ABSTRACT. Does generative AI widen or compress productivity differences between workers? This paper develops a microeconomic model in which an AI tool has two separate margins: a *frontier*, the problems it can solve for anyone, and *leverage*, the extra work it lets a person get through. The gain from adoption is V-shaped in ability, with its minimum at the worker whose knowledge equals the AI's frontier. Frontier advances are equalising and leverage is skill-biased. The V survives leverage that rises with ability as long as its gradient is below a stated bound, and when the frontier is a distribution rather than a point the kink becomes a smooth valley whose floor sits at an observable quantile. The average AI×skill complementarity therefore changes sign as the frontier moves through the ability distribution. This offers an account of conflicting experimental findings that differs from the autonomy-based account of Ide and Talamàs (2025) and can be tested against it: tools whose main effect is speed rather than knowledge should favour the able even when they act only as co-pilots. Relative inequality never rises with ability-neutral leverage, but inequality in levels can. When the frontier is jagged, the ability to judge AI output becomes the skill-biased margin, and over-trusting workers can lose. Adoption is two-tailed, so the sign of selection bias depends on the vintage of the technology. Workers capture a share of the gain that rises with diffusion across employers. The return to ability falls to zero below the frontier, which removes the incentive for juniors to climb rungs the AI already occupies. The paper closes with seven testable predictions, the implied sign of selection in each existing study, and the designs needed to test them.

Preliminary draft. This is a theory note. Sections 3–9 develop the model, with all propositions proved in Appendix A and verified symbolically; Section 10 illustrates them with stylised parameters that are not estimates; Section 11 sets out the predictions and the research design needed to test them. Comments welcome; please do not cite without permission.

Keywords: generative AI, technology adoption, skill complementarity, knowledge hierarchies, wage inequality, human capital, rent sharing. **JEL:** D24, J24, J31, M53, O33.

1 Introduction

Whether generative AI makes workers more productive is largely settled. In the settings studied so far it does, often substantially. Who becomes more productive is less clear. In customer support and mid-level writing, the least experienced workers gained most and differences between workers narrowed (Brynjolfsson et al., 2025; Noy and Zhang, 2023). Among entrepreneurs given an AI adviser, the best performers gained and the weakest lost (Otis et al., 2024). Among consultants, the same tool helped on some tasks and hurt on others (Dell’Acqua et al., 2023). These results are usually read as evidence for one of two competing views: that AI is skill-biased, amplifying the advantage of the able, or that it diffuses knowledge, letting the less able catch up.

This paper argues that, within a job, the two views can describe different regions of a single gain profile, and develops a microeconomic model that says which region a given job is in. The model starts from a simple description of ability. A worker’s ability is the hardest problem they can solve unaided (Garicano, 2000). An AI tool has two separate properties. Its *frontier* is the range of problems it can solve for anyone who asks. Its *leverage* is the additional work it lets a person get through. The frontier sets a floor under everyone’s output. Leverage multiplies output, so it is worth more to those who already produce more.

The central result (Proposition 1) is that the gain from AI is V-shaped in ability, with its minimum at the worker whose own knowledge coincides with the AI’s frontier. Below the frontier, the less able gain more because the floor lifts them further. Above it, the more able gain more because leverage scales their output. The cross-derivative of output with respect to AI and ability, the object that separates the two views, is negative on one side of the frontier and positive on the other. Its population average changes sign as the frontier moves through the ability distribution (Proposition 2). Early tools whose frontier lies below most workers are skill-biased. Later tools whose frontier lies above most workers are equalising. The sign of an estimated AI×skill interaction therefore reflects the technology vintage and the job as much as the technology itself.

Relation to Ide and Talamàs. The closest paper is Ide and Talamàs (2025), who place AI in the knowledge hierarchies of Garicano (2000) and study its equilibrium effects on organisations and incomes. They also argue that their results reconcile the conflicting evidence, but through a different mechanism: autonomous AI, which can do production work itself, mainly benefits the most knowledgeable, while non-autonomous AI, which only advises, mainly benefits the least knowledgeable. This paper is complementary. It sets aside organisational equilibrium and works within a single job, and it separates two properties of a tool that their model does not distinguish: what the tool knows and how much faster it makes a person work. The two accounts overlap where both say that knowledgeable co-pilots help the least able. They come apart for tools whose main effect is speed. In this paper a co-pilot with high leverage and a low frontier is skill-biased, because leverage scales output; in theirs, a co-pilot favours the least knowledgeable. The paper’s distinct contribution is this separation, the within-job predictions it delivers, and the design that tells the two accounts apart (Section 11).

Five further results follow from the same structure.

Dispersion (Proposition 4). With leverage that is the same for every worker, AI never increases relative inequality in output: the Lorenz curve shifts inward. It can, however, increase inequality in levels, because leverage scales absolute gaps. Studies using logs and studies using levels can therefore reach opposite conclusions from the same data.

Judgment (Proposition 5). When the AI also attempts problems it cannot reliably solve, and its errors look plausible, using it beyond one's own competence requires the ability to recognise wrong answers. Judgment becomes the skill-biased margin of AI. Workers who over-trust the tool can be made worse off, and those losses are concentrated at the bottom. Verification can be bought, so firms can partly substitute review processes for worker discernment.

Adoption and selection (Proposition 6). With a cost of learning to use the tool, adoption is two-tailed. The able adopt for leverage, the less able for the floor, and the competent middle adopts last. As the frontier advances, adopters go from being positively to negatively selected on ability, so the bias in comparisons of users with non-users changes sign over time. Randomised access with voluntary take-up turns the valley into a flat trough.

Incidence (Proposition 7). How the gain is split between worker, employer and vendor depends on the share of employers using the tool. Early in diffusion the employer captures most of it, and wages catch up as the tool becomes standard and enters the worker's outside option. In the long run the wage return to ability vanishes below the frontier and rises above it, so the skill premium develops a kink at the AI's frontier.

Learning (Proposition 8). Because ability below the frontier is not paid, a junior worker has no reason to climb rungs the AI already occupies unless they can reach the frontier within the horizon. Learning becomes all-or-nothing, and each frontier advance stops it for a wider band of juniors. AI used as a tutor works against this. Employers that do not capture the returns to general skills have no reason to pay for practice when the tool can do the work.

The paper is a theory paper. All propositions are proved in Appendix A and checked symbolically and numerically in the replication code. Section 10 illustrates them with stylised parameters that are not estimates. Section 11 sets out seven testable predictions, reads the existing experimental evidence against them, and describes the data and designs needed to test them directly. The most informative designs are model upgrades used as shocks to the frontier with leverage held fixed, staggered rollouts with mandated use that estimate effects by decile of baseline ability, and comparisons of speed-oriented with knowledge-oriented tools, which separate this account from that of Ide and Talamàs (2025).

This note belongs to a series. WP 1 studies how local housing markets absorb population shocks. RN 1 studies how inference costs change the economics of software firms when AI becomes their product. This paper studies the human work that interacts with that product. Together they follow a common sequence of shock, adjustment and incidence, applied to markets, firms and workers.

2 Related literature

Tasks, skills and technology. The task-based approach treats technologies as performing tasks previously done by workers and derives wage effects from which tasks are affected (Autor et al., 2003; Acemoglu and Autor, 2011; Acemoglu and Restrepo, 2018). Earlier computerisation automated routine tasks and complemented the problem-solving of educated workers, which

is the canonical skill-biased case (Bresnahan et al., 2002). Autor (2024) argues that generative AI may reverse this pattern by extending expert judgment to non-experts. Acemoglu (2025) bounds the macroeconomic gains from tasks exposed to AI. This paper works inside a single job and asks how the gain depends on the worker, which requires modelling within-task heterogeneity in ability rather than the allocation of tasks between capital and labour.

Knowledge hierarchies. The model of ability builds on Garicano (2000) and Garicano and Rossi-Hansberg (2006), in which knowledge is the range of problems a person can solve and organisations route hard problems to those who know more. The closest paper is Ide and Talamàs (2025), who place AI in this framework in general equilibrium. Their results run along two dimensions. Along the first, the AI’s knowledge: basic autonomous AI, with the knowledge of pre-AI workers, harms the least knowledgeable and benefits the most knowledgeable, while advanced autonomous AI, with the knowledge of pre-AI solvers, benefits both, and each reorganises firms differently. Along the second, autonomy: autonomous AI mainly benefits the most knowledgeable, non-autonomous AI the least knowledgeable, and output is higher with autonomy. They argue that these results reconcile the conflicting experimental evidence. Proposition 2 is the single-worker counterpart of their first dimension. This paper does not model autonomy or the reallocation of problems across people. It adds a margin that their AI agents do not have, leverage, the speed-up the tool gives a person, and works at the level of the individual gain profile. That delivers predictions that can be estimated within a job: the V shape and the location of its minimum, the separation of frontier from leverage, the role of judgment, the selection pattern in adoption, incidence and learning. It also yields a prediction that distinguishes the two accounts, for speed-oriented co-pilots (Section 11).

Prediction and judgment. Agrawal et al. (2018, 2019) argue that AI lowers the cost of prediction and raises the value of judgment. Section 6 gives judgment a precise role: it is the ability to detect wrong answers, and it is the channel through which AI becomes skill-biased once its frontier is jagged. The experiment of Dell’Acqua et al. (2023) documents the jagged frontier directly.

Experimental evidence. A growing body of experiments measures the effect of generative AI on individual productivity (Noy and Zhang, 2023; Peng et al., 2023; Brynjolfsson et al., 2025; Dell’Acqua et al., 2023; Otis et al., 2024; Cui et al., 2025). Their average effects are large and their heterogeneity results differ, with some showing compression and others divergence. Section 11 argues that the model accounts for both and shows which features of those designs would have to change to test it. Exposure measures at the occupation level (Felten et al., 2021; Eloundou et al., 2024) and usage measures from AI providers (Handa et al., 2025) correspond to the two different objects distinguished in Section 3.

Selection, incidence and training. The adoption problem is a Roy model (Roy, 1951) with selection on gains (Heckman and Vytlacil, 2005). The incidence section uses a standard Nash-bargaining split with outside options. The novel element is that the outside option depends on the diffusion of the technology across employers. Evidence of small earnings

effects to date is in Humlum and Vestergaard (2025), and of positively selected early adoption in Bick et al. (2024). The training section draws on the theory of general human capital (Becker, 1964; Acemoglu and Pischke, 1998). The inequality results use the classical conditions under which a transformation of incomes reduces Lorenz inequality (Jakobsson, 1976; Fellman, 1976).

3 The model

Workers and problems. A worker is described by ability $a \in [0, 1]$, distributed in the population with cumulative distribution G and density $g > 0$. Ability is knowledge in the sense of Garicano (2000): a is the hardest problem the worker can solve unaided. Each worker faces a unit flow of problems whose difficulty $z \in [0, 1]$ has distribution H with density $h > 0$. A solved problem of difficulty z is worth $v(z) > 0$, and an unsolved one is worth nothing. Without AI, a worker of ability a produces

$$Y_0(a) = V(a) \equiv \int_0^a v(z) h(z) dz, \quad (1)$$

which is strictly increasing. This is the $Y_i = A_i F(\cdot)$ of the paper’s motivating question, with the production function made explicit: ability matters because it determines which problems get solved.

The AI technology. An AI tool is described by two numbers.

- The *frontier* $k \in [0, 1]$. Knowledge up to difficulty k has been codified. A worker with access to the tool can solve any problem with $z \leq k$, whatever their own ability, by asking for the answer.
- The *leverage* $\lambda \geq 0$. The tool drafts, searches and formats. This frees time, so the worker gets through $(1 + \lambda)$ times as many problems.

A worker who adopts the tool therefore produces

$$Y_1(a) = (1 + \lambda) V(\max\{a, k\}). \quad (2)$$

The gain from adoption is $\Delta(a) \equiv Y_1(a) - Y_0(a)$.

The two parameters correspond to different margins of technical progress. The frontier k is what the model knows, and it moves with model releases. Leverage λ is how much faster it lets a person work, and it moves with interfaces, integration into workflows and the worker’s own fluency. The paper’s results come from keeping these two margins apart. A single “AI productivity” parameter would merge two forces that push the distribution in opposite directions.

Equation (2) makes the frontier deterministic. A problem below k is always solved, and the AI never claims to solve one above it. Section 6 drops this assumption: above the floor the AI produces answers that may be wrong, and the worker must judge them.

Exposure is not intensity. Two measures of “how much AI” a worker uses follow from the model. *Exposure* is the share of the job’s problems the AI can solve, $H(k)$. It is a property of the

occupation and the technology and is the same for every worker in the job. Occupation-level indices such as Eloundou et al. (2024) and Felten et al. (2021) measure this object. *Intensity* is the share of a worker's problems that are solved only because of the AI,

$$\iota(a) = \max\{H(k) - H(a), 0\}, \quad (3)$$

which is strictly decreasing in ability below the frontier and zero above it. Intensity is what usage logs and surveys of reliance capture (Handa et al., 2025). In the model the two measures diverge systematically. Within a job, the heaviest users are the least able workers, while the gains from leverage accrue mainly to workers who barely depend on the AI for answers. The correlation between intensity and productivity across workers therefore reveals neither the size nor the incidence of the gains.

Uniform benchmark. Closed forms use the benchmark in which G and H are uniform on $[0, 1]$ and $v \equiv 1$, so that $V(a) = a$ and $Y_1(a) = (1 + \lambda) \max\{a, k\}$. Every qualitative result is stated and proved for general G, H and v . The benchmark is used only to give thresholds a formula.

4 Who becomes more productive: the valley at the frontier

The paper's central object is the cross-derivative of output with respect to AI and ability. With adoption a binary choice, it is the slope of the gain in ability, $\Delta'(a)$. Positive values mean that AI is skill-biased, and negative values that it diffuses knowledge.

Proposition 1 (The valley). *For any G, H and v :*

- (i) *The gain $\Delta(a)$ is strictly decreasing in ability below the frontier and, if $\lambda > 0$, strictly increasing above it. It reaches its minimum at $a = k$, where it equals $\lambda V(k)$.*
- (ii) *An advance in the frontier raises output only for workers below it: $\partial Y_1 / \partial k = (1 + \lambda)v(k)h(k) \mathbf{1}\{a < k\}$, which is weakly decreasing in a . An increase in leverage raises output by $\partial Y_1 / \partial \lambda = V(\max\{a, k\})$, which is weakly increasing in a .*

The cross-derivative in the paper's motivating question therefore has no single sign. It is negative for workers below the frontier and positive for those above it. The two competing hypotheses, skill-biased AI and knowledge-diffusing AI, are not rival descriptions of the technology. They describe different parts of the same gain profile. The worker who gains least is the one whose own knowledge sits exactly at the AI's frontier. That worker no longer needs the AI to supply answers, yet is not expert enough for the leverage margin to be worth much. Part (ii) explains why the two margins should be measured separately. Frontier advances are equalising and leverage is skill-biased. A model release that mainly raises k and a workflow integration that mainly raises λ can have opposite distributional effects with the same average effect.

Which side dominates depends on where the frontier sits in the ability distribution. Define the population slope of gains on ability as the coefficient of a regression of $\Delta(a)$ on a across workers, $S(k, \lambda) = \text{Cov}(\Delta(a), a) / \text{Var}(a)$.

Proposition 2 (The sign flip). *For any G , H and v , $S(k, \lambda)$ is strictly decreasing in k and strictly increasing in λ . It is positive at $k = 0$ whenever $\lambda > 0$ and negative at $k = 1$. There is therefore a unique frontier $\hat{k}(\lambda) \in (0, 1)$ at which AI turns from skill-biased to equalising on average, and $\hat{k}$ is increasing in λ . In the uniform benchmark,*

$$S(k, \lambda) = (1 + \lambda)(1 - k)^2(1 + 2k) - 1. \quad (4)$$

This gives a dynamic prediction. Early versions of a tool, whose frontier lies below most workers in an occupation, favour the able. As the frontier passes through the ability distribution, the average complementarity turns negative even though leverage has not fallen. The sign of the AI×skill interaction estimated in any study is thus specific to a technology vintage and an occupation. In the uniform benchmark with $\lambda = 0.2$, the flip occurs once the AI can solve problems up to the 0.26 quantile of difficulty. Experiments that find knowledge-diffusion patterns, such as those on customer support or mid-level writing (Section 11), are therefore informative about where the frontier sits in those jobs. They are not evidence that AI is equalising in general.

4.1 How robust is the valley?

The V in Proposition 1 comes from two features of equation (2): a floor at a single frontier and leverage that multiplies output equally for everyone. Both are strong. The next proposition relaxes each.

Proposition 3 (Robustness of the valley). (i) *Leverage that rises with ability. Let $Y_1(a) = (1 + \lambda(a))V(\max\{a, k\})$ with $\lambda(a) \geq 0$ differentiable and non-decreasing. Above the frontier the gain is strictly increasing whenever $\lambda(a) > 0$. Below it the gain is strictly decreasing, so the valley keeps its minimum at k , if and only if*

$$\lambda'(a)V(k) < V'(a) \quad \text{for all } a < k. \quad (5)$$

Where the condition fails, the descending branch flattens or turns up and the minimum moves below k .

(ii) *A frontier that is a distribution. Let the frontier vary across the problems in a job, so that the tool can solve a problem only up to a difficulty K drawn from a continuous distribution F , independently of the worker, and $Y_1(a) = (1 + \lambda) \mathbb{E}_K[V(\max\{a, K\})]$. Then*

$$\Delta'(a) = V'(a) [(1 + \lambda)F(a) - 1],$$

so the gain is strictly decreasing below and strictly increasing above

$$a^* = F^{-1}\left(\frac{1}{1 + \lambda}\right). \quad (6)$$

The kink becomes a smooth valley whose floor sits at the $1/(1 + \lambda)$ quantile of the frontier distribution. With a point mass at k , $a^ = k$ and Proposition 1 is recovered.*

Part (i) says when the obvious objection bites. If the able get more out of the tool's speed,

as they plausibly do, the valley survives as long as the gradient of leverage is small relative to the marginal value of one's own knowledge; in the uniform benchmark the condition is $\lambda'(a) < 1/k$. Leverage that rises steeply with ability tilts the whole profile towards skill bias, and the knowledge-diffusion branch shrinks. Part (ii) is the more useful generalisation. In real jobs the tool solves some hard problems and fails some easy ones, so the frontier is a distribution rather than a point. The valley then has no kink, but its floor is still interior and has an observable location: it sits where the share of the worker's problems that the tool can solve, $F(a)$, equals $1/(1 + \lambda)$. More leverage moves the floor down the ability distribution; a better model, which shifts F to the right, moves it up. Both results are verified numerically in the replication code.

5 Dispersion: relative and absolute inequality

Whether AI widens or compresses productivity differences depends on how dispersion is measured. Suppose every worker in a job adopts.

Proposition 4 (Dispersion). (i) Relative dispersion never rises. For any G, H, v, k and λ , the Lorenz curve of Y_1 lies weakly above that of Y_0 . Every Lorenz-consistent index (Gini, Theil, coefficient of variation) is therefore weakly lower with AI. The variance of log output and every interquantile ratio $Q_p(Y)/Q_{p'}(Y)$, $p > p'$, are also weakly lower. The Gini, the coefficient of variation and the variance of log output fall strictly when $0 < G(k) < 1$.

(ii) Absolute dispersion can rise. $\text{Var}(Y_1) = (1 + \lambda)^2 \text{Var}(\max\{Y_0, V(k)\})$. In the uniform benchmark,

$$\frac{\text{Var}(Y_1)}{\text{Var}(Y_0)} = (1 + \lambda)^2(1 - k)^3(1 + 3k), \quad (7)$$

which is strictly decreasing in k . The variance in levels falls if and only if $k > \bar{k}(\lambda)$, and $\bar{k}(\lambda) > \hat{k}(\lambda)$ for every $\lambda > 0$.

Part (i) holds because AI raises each worker's output by a proportion, $Y_1/Y_0 = (1 + \lambda) \max\{1, V(k)/V(a)\}$, that is weakly decreasing in ability. Such a transformation always shifts the Lorenz curve inward (Jakobsson, 1976; Fellman, 1976). Leverage multiplies everyone's output by the same factor, so it leaves relative inequality unchanged. The floor raises the bottom and does nothing to the top. Part (ii) shows the other side. Leverage scales absolute gaps, and a modest λ is enough to widen the variance in levels until the frontier is well into the distribution. In the uniform benchmark with $\lambda = 0.2$, the variance in levels starts to fall only once k exceeds 0.28, later than the slope turns negative ($\hat{k} = 0.26$). There is therefore an intermediate range of frontiers in which a study measuring dispersion in levels finds widening and one measuring it in logs finds compression, using the same data.

Remark (Ability-biased leverage). Part (i) relies on leverage being the same for everyone. If able workers get more out of the tool, so that $\lambda = \lambda(a)$ with $\lambda' > 0$, relative dispersion can rise. The ratio Y_1/Y_0 becomes the product of a rising factor $1 + \lambda(a)$ and a falling factor $\max\{1, V(k)/V(a)\}$. The Lorenz ranking holds whenever the product is non-increasing. Below the frontier this requires $\lambda'(a)/(1 + \lambda(a)) \leq v(a)h(a)/V(a)$: the growth rate of leverage in ability must not exceed the growth rate of unaided output. Above the frontier the second

factor is constant, so any ability bias in leverage widens relative inequality among the workers who already exceed the AI. Section 6 gives a microfoundation: ability-biased leverage arises naturally when the AI's output must be judged.

6 The jagged frontier: judgment as the skill-biased margin

The baseline treats the frontier as a clean line. In practice, AI tools also attempt problems they cannot reliably solve, and their wrong answers look like right ones. Dell'Acqua et al. (2023) call this the “jagged technological frontier”. This section adds it to the model. Above the floor, the AI's reliability cannot be predicted from the problem's difficulty, and using its output requires judgment.

Setup. Set $v \equiv 1$. For a problem above both the worker's own ability and the floor, $z > \max\{a, k\}$, the AI returns a candidate answer. The answer is correct with probability $q \in (0, 1)$, independently across problems. A wrong answer is plausible. The worker detects it with probability $\pi(a)$, where π is strictly increasing because judging an answer requires some of the knowledge needed to produce it. A detected error is discarded, with value zero. An undetected one is submitted and costs $d > 0$, through rework, a customer's loss or a reputational hit. A worker who *reaches*, submitting the AI's answers beyond their own competence, earns the expected value per problem

$$r(a) = q - (1 - q)(1 - \pi(a))d. \quad (8)$$

- Proposition 5 (Judgment).** (i) *The value of reaching, $r(a)$, is strictly increasing in ability. A worker who correctly perceives $\pi(a)$ reaches if and only if $a \geq a_r$, where $r(a_r) = 0$. With $\pi(a) = a$, $a_r = 1 - q/((1 - q)d)$.*
- (ii) *The total gain is $\Delta(a) + (1 + \lambda)r(a)^+ (1 - H(\max\{a, k\}))$. The reach component is zero below a_r . Above it, it is the product of a rising return per problem and a falling number of problems left to reach for. In the uniform benchmark with $\pi(a) = a$ and $a \geq k$, it peaks at $a = (1 + a_r)/2$.*
- (iii) *If workers overestimate their discernment, so that the perceived $\tilde{\pi}(a) > \pi(a)$, some workers with $a < a_r$ reach and lose on every reached problem. Their total gain can be negative. Losses are confined to the bottom of the distribution, since the realised gain is non-negative for every $a \geq a_r$, and a higher floor k reduces them.*

Judgment is the skill-biased margin of AI. In the baseline, the AI supplies answers and the worker's ability matters only through leverage. Once the AI also supplies wrong answers, ability is needed to tell them apart, the complementarity that Agrawal et al. (2019) draw between prediction and judgment. The proposition sorts outcomes reported in the experimental literature (Section 11). On problems inside a clean frontier, the bottom gains most. On problems outside it, users of AI do worse than non-users. Where the frontier is jagged and users over-trust the tool, the bottom can lose outright while the top gains. The model attributes these three patterns to the position of the task relative to k and to the worker's ability relative to a_r , with no difference in the underlying technology.

An implication for organisations follows. Because $r(a)$ depends on $\pi(a)$, a firm can raise the value of AI for its weaker workers by raising detection rather than ability: review layers, test suites, checklists, or a second worker who verifies. Verification is a complement of the AI that can be bought. This links the worker-level model to the firm-level question of RN 1, where the cost of verifying an agent's output determines whether software can be sold by outcome.

7 Who adopts, and why comparing users with non-users misleads

Return to the baseline with a clean frontier. Adoption is now a choice. Learning to use the tool well takes time and attention, which costs the worker $c > 0$. A worker adopts if and only if $\Delta(a) \geq c$. Let $D(a) = \mathbf{1}\{\Delta(a) \geq c\}$.

Proposition 6 (Two-tailed selection). *(i) The set of adopters is the union of a lower tail and an upper tail, $[0, a_L] \cup [a_H, 1]$, either of which may be empty. Non-adopters form an interval around the frontier.*

(ii) The lower tail exists if and only if $(1 + \lambda)V(k) \geq c$, with $V(a_L) = (1 + \lambda)V(k) - c$. The upper tail starts at $a_H = \max\{k, V^{-1}(c/\lambda)\}$. Everyone adopts once $\lambda V(k) \geq c$. Provided $\lambda V(1) > c$, as the frontier advances adoption therefore spreads from the top of the ability distribution to the bottom and finally to the middle.

(iii) The raw productivity gap between users and non-users decomposes as

$$\mathbb{E}[Y \mid D = 1] - \mathbb{E}[Y \mid D = 0] = \underbrace{\mathbb{E}[\Delta \mid D = 1]}_{\text{effect on adopters}} + \underbrace{\mathbb{E}[Y_0 \mid D = 1] - \mathbb{E}[Y_0 \mid D = 0]}_{\text{selection}}. \quad (9)$$

The selection term is positive while only the upper tail adopts. It falls, and can change sign, once the lower tail appears. Moreover, $\mathbb{E}[\Delta \mid D = 1] \geq c > \mathbb{E}[\Delta \mid D = 0]$, so the effect on adopters overstates the gain non-adopters would obtain.

Part (ii) predicts a change in who uses AI, not only in how many. In early diffusion, users are positively selected on ability. This is consistent with survey evidence that early adopters of generative AI are younger, more educated and better paid (Bick et al., 2024), though the model attributes it to the location of the frontier rather than to familiarity with technology. Later, the least experienced become heavy users, as the intensity measure (3) already implies. Throughout, the workers near the frontier, the competent middle, are the last to adopt. For them the tool supplies no answers they lack and little leverage.

Part (iii) makes the selection problem that motivates the paper precise, and shows that its sign is not fixed. Whether “good workers adopt AI” or “weak workers adopt AI” depends on the vintage of the technology. A fixed-effects correction for ability removes the selection term but not the second problem in the decomposition. Adopters are selected on gains, so the effect on adopters is not the effect on the population (Roy, 1951; Heckman and Vytlačil, 2005).

Experiments with voluntary take-up. Suppose access is randomised but use is voluntary. The intention-to-treat effect at ability a is $D(a)\Delta(a)$. This is zero in the middle and positive in the tails, so a valley in gains becomes a flat trough in measured effects. Heterogeneity by baseline

performance in such an experiment confounds two objects: who chooses to use the tool and how much it helps those who do. Separating them requires either measuring take-up by ability, so that the Wald ratio can be computed within ability bins, or a design with mandated use (Section 11).

8 Who keeps the gain: incidence and the price of expertise

A productivity gain is not a wage gain. This section splits the surplus created by AI among the worker, the employer and the vendor of the tool.

Setup. A firm employing a worker of ability a deploys the tool at a price p per seat, set by the vendor (the object of RN 1). The net gain from deployment is $\Gamma(a) = \Delta(a) - p$, and the firm deploys whenever $\Gamma(a) \geq 0$. The worker's outside option comes from a competitive fringe of employers that pay workers their output. A share φ of fringe employers use the tool, so an outside offer pays $V(a) + \Gamma(a)$ with probability φ and $V(a)$ otherwise. The outside option is $\omega(a) = V(a) + \varphi \Gamma(a)$. Wages are set by Nash bargaining over the match surplus, with worker bargaining power $\beta \in (0, 1)$:

$$w(a) = \omega(a) + \beta [V(a) + \Gamma(a) - \omega(a)] = V(a) + [\varphi + \beta(1 - \varphi)] \Gamma(a). \quad (10)$$

Proposition 7 (Incidence). (i) *The gross gain $\Delta(a)$ is split as follows: p to the vendor, a share $s = \varphi + \beta(1 - \varphi)$ of $\Gamma(a)$ to the worker and $(1 - \beta)(1 - \varphi)$ of $\Gamma(a)$ to the firm. The shares do not depend on ability.*

(ii) *The worker's share is increasing in diffusion: $\partial s / \partial \varphi = 1 - \beta > 0$. In early diffusion, $\varphi \approx 0$, the worker keeps only a share β and the employer captures the rest. When all employers use the tool, the worker keeps the entire net gain.*

(iii) *With full diffusion, $w(a) = (1 + \lambda)V(\max\{a, k\}) - p$ for every worker whose gain covers the price, $\Delta(a) \geq p$; workers with $\Delta(a) < p$ are not equipped and earn $V(a)$. Among equipped workers, the return to ability, $w'(a)$, is zero below the frontier and $(1 + \lambda)v(a)h(a)$ above it. Relative to the economy without AI, the wage–ability profile flattens below k and steepens above it.*

Part (ii) explains why measured productivity gains can coexist with small wage effects early in diffusion. The gain exists, but the worker's outside option has not yet moved, so the employer keeps most of it. Wages catch up as the tool becomes standard across employers, because the gain then becomes part of what the worker can earn elsewhere. Evidence that generative AI has so far had small effects on earnings despite reported time savings (Humlum and Vestergaard, 2025) fits this pattern. The model reads it as a statement about diffusion and bargaining, not about the size of the productivity effect.

Part (iii) is the long-run wage implication of the floor. Expertise below the frontier stops being paid, because every worker can obtain it from the tool. Expertise above the frontier is paid more, because leverage multiplies it. The skill premium does not rise or fall uniformly. It develops a kink at the AI's frontier, and the kink moves up with each model release. Workers just below the frontier see the price of the knowledge they built fall to zero, while earning more than before in absolute terms.

Remark (The vendor's pricing and the valley). Under a uniform seat price, the firm deploys the tool for worker a only if $\Delta(a) \geq p$. The workers excluded first are those in the valley around the frontier, since their gain is smallest. A vendor that can meter use (RN 1) charges in proportion to consumption. If consumption tracks intensity, $\iota(a)$ in (3), usage-based pricing charges most for the floor gains of the least able. Seat pricing charges the same for the leverage gains of the most able. How the tool is priced therefore affects how its surplus is shared across workers, not only between firm and vendor.

9 Productivity today, human capital tomorrow: the missing rungs

The previous sections take ability as given. This section allows workers to build it. The question is whether a tool that answers a junior's questions also removes the reason to learn the answers.

Setup. There are two periods. A worker starts with ability a_1 and uses the tool in both. In period 1 the worker can *practise*: attempt a block of problems at the edge of their competence unaided rather than delegating them. Practising a share $e \in [0, 1]$ of the block costs κe in period-1 output, the value the tool would have produced on those problems. Practice builds ability through learning by doing,

$$a_2 = a_1 + (1 + \tau) \gamma e, \quad (11)$$

where $\gamma > 0$ is the learning rate and $\tau \geq 0$ the extent to which the tool acts as a tutor, for instance by explaining its answers, making practice more productive. In period 2 all employers use the tool and the labour market is competitive, so by Proposition 7(iii) the worker earns $w_2 = (1 + \lambda)V(\max\{a_2, k\}) - p$. The worker chooses e to maximise $-\kappa e + \delta w_2$, with discount factor $\delta \in (0, 1)$. To isolate the mechanism, use the uniform benchmark, $V(a) = a$, and compare with a world without AI ($k = \lambda = \tau = 0$).

Proposition 8 (The missing rungs). *(i) Without AI, the worker practises fully if $\delta\gamma \geq \kappa$ and not at all otherwise, whatever their initial ability.*

(ii) With AI, the optimal e is also all or nothing. A worker who starts at or above the frontier practises if and only if $\delta(1 + \lambda)(1 + \tau)\gamma \geq \kappa$. A worker who starts below the frontier practises if and only if

$$a_1 \geq a_1^* \equiv k - (1 + \tau)\gamma + \frac{\kappa}{\delta(1 + \lambda)}. \quad (12)$$

(iii) The threshold a_1^ rises one-for-one with the frontier, falls with tutoring τ and falls with leverage λ . A frontier advance therefore stops learning for a widening band of junior workers, who are more productive today and no more capable tomorrow.*

Proposition 8 is the human-capital counterpart of the valley. Below the frontier, ability is not paid (Proposition 7). Climbing from one rung to the next below k therefore earns nothing, and practice pays only if it can carry the worker past the frontier within the horizon. Learning becomes all-or-nothing: either the worker can jump the flat stretch of width $k - a_1$, or they do

not start. With many periods the logic is the same, and the relevant comparison is between the length of the flat stretch and the present value of the premium beyond it.

Two forces work against the effect. Tutoring lowers the threshold, and so does leverage, because it raises the reward for expertise above k . Whether AI acts mainly as a crutch or mainly as a tutor in a given occupation is therefore the race between $\tau\gamma$ and the speed at which k moves.

Who pays for the rungs? The proposition assumes that the worker bears the cost of practice and collects its return. In firms, the employer often decides how work is allocated. When human capital is general, a competitive employer gains nothing from a worker's future ability (Becker, 1964), so it assigns all problems to the tool, $e = 0$, whatever the worker's own incentive. Labour-market frictions that let employers capture part of the return to general training restore some investment (Acemoglu and Pischke, 1998). AI makes the problem harder, because it raises the forgone output κ of letting a junior struggle. Entry-level jobs have historically financed training by bundling it with low-value work. Proposition 8 suggests that this bundle comes apart as the frontier rises, and that the stock of workers who eventually exceed the frontier, the people whose judgment Section 6 shows to be the complement of AI, may shrink.

A frontier that limits itself. The last point has a dynamic consequence that the two-period model can only sketch. Suppose the frontier advances by learning from expert work, so that its next step is proportional to the share of the workforce whose ability exceeds it: $k_{t+1} = k_t + \theta [1 - G_t(k_t)]$ with $\theta > 0$. Suppose also that each new cohort of juniors climbs past the frontier only if it starts above the threshold $a_1^*(k_t)$ of equation (12), which rises one for one with k_t . Then the share of experts above the frontier in each new cohort is $1 - G(a_1^*(k_t))$ and shrinks as the frontier advances. Under these assumptions, and once earlier cohorts of experts retire, the frontier rises at a decreasing rate and cannot advance beyond the point where the threshold reaches the top of the junior ability distribution, $k_\infty = 1 + (1 + \tau)\gamma - \kappa/[\delta(1 + \lambda)]$ when that is below one. Tutoring and leverage raise the limit; a technology that only answers questions lowers it. The sketch is not a result about any actual model, whose training data need not come from the same occupation. But it shows how the missing rungs of Proposition 8 could feed back on the frontier itself, and it is the dynamic question this note leaves open.

10 Numerical illustration

This section illustrates the propositions in the uniform benchmark. All parameter values are illustrative and are not estimates. Unless stated otherwise, leverage is $\lambda = 0.2$, a 20% increase in the problems a worker can handle.

The valley moves. Figure 1 plots the gain from adoption against ability for three frontiers. Each profile is V-shaped with its minimum at $a = k$. When the frontier is low ($k = 0.2$), most workers lie above it and gains rise with ability. When it is high ($k = 0.8$), most lie below it and gains fall with ability. The same leverage parameter produces both patterns.

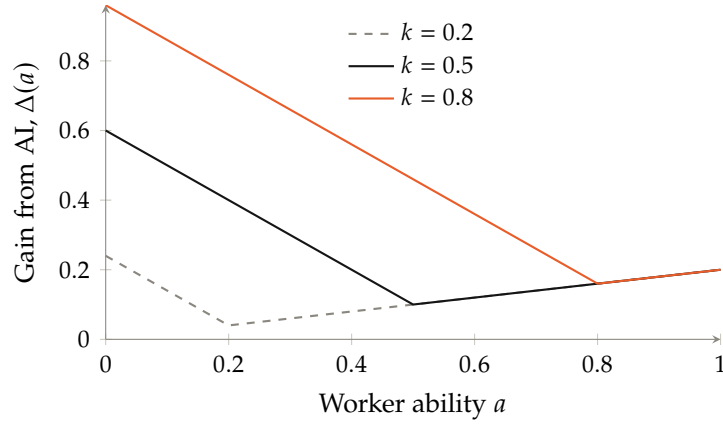


FIGURE 1. Gain from adoption by ability for three AI frontiers ($\lambda = 0.2$, uniform benchmark). The minimum is always at the frontier.

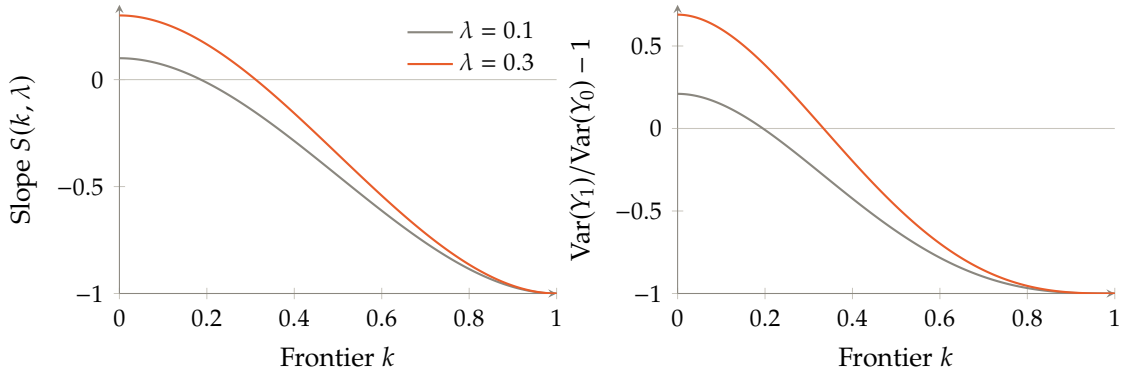


FIGURE 2. Left: slope of the gain on ability. Right: proportional change in the variance of output in levels. Uniform benchmark.

Slope and variance. Figure 2 plots the population slope of gains on ability, equation (4), and the proportional change in the variance of output, equation (7), against the frontier. The slope turns negative at $\hat{k} = 0.19$ when $\lambda = 0.1$ and at $\hat{k} = 0.31$ when $\lambda = 0.3$. The variance in levels starts to fall slightly later, at $\bar{k} = 0.20$ and $\bar{k} = 0.33$ respectively. Between the two thresholds, the average AI×skill interaction is already negative but absolute dispersion is still rising. Relative dispersion (Proposition 4(i)) falls throughout.

Judgment. Figure 3 adds the jagged frontier of Section 6 with a low floor ($k = 0.15$), AI answers beyond the floor correct with probability $q = 0.4$, an error cost $d = 1.5$ and discernment $\pi(a) = a$. Rational workers reach beyond their competence only above $a_r = 0.56$, and their gain then rises above the floor-only gain. Workers who trust the tool as if they could catch every error reach regardless. Those below $a = 0.42$ end up worse off than without AI, with a loss of 0.33 for the least able. The same tool produces gains at the top and losses at the bottom.

Adoption. Figure 4 shows who adopts as the frontier advances, with adoption cost $c = 0.12$. Below $k = 0.10$ only the upper tail adopts. A lower tail then appears and grows, while non-adopters are confined to a shrinking band around the frontier, until everyone adopts at $k = 0.60$. The raw ability gap between adopters and non-adopters, the selection term in (9),



FIGURE 3. Gains with a jagged frontier ($k = 0.15$, $q = 0.4$, $d = 1.5$, $\pi(a) = a$, $\lambda = 0.2$).

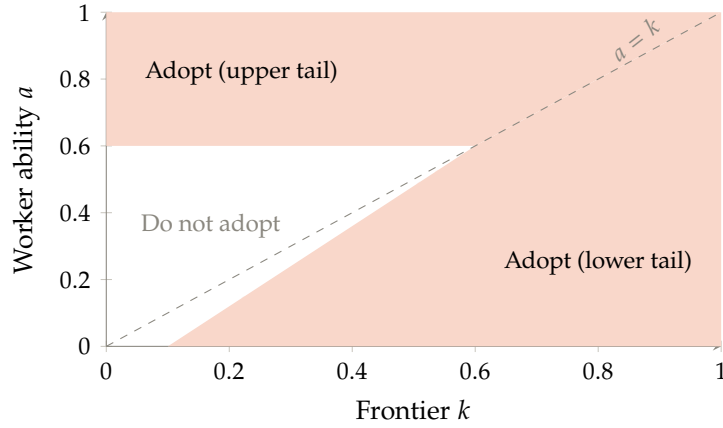


FIGURE 4. Adoption regions by ability as the frontier advances ($c = 0.12$, $\lambda = 0.2$). Shaded: adopters.

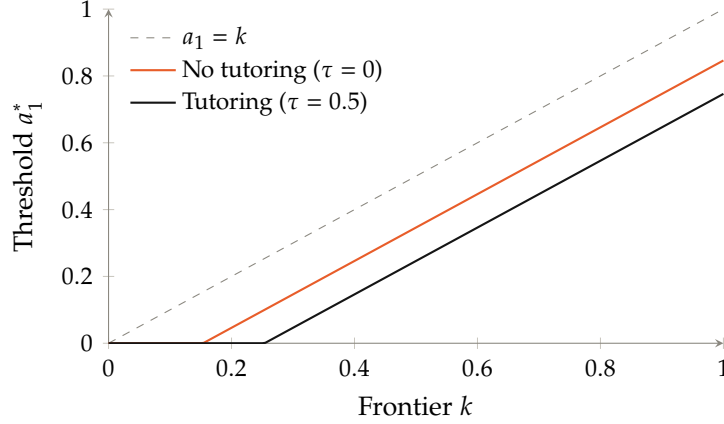
falls from 0.49 at $k = 0.05$ and turns negative at $k \approx 0.44$. A study comparing users with non-users would find that AI users are the able workers early on and the less able ones later, with the same technology and a different frontier.

Incidence. Table 1 reports the worker's share of the net gain, $\varphi + \beta(1 - \varphi)$, for two values of bargaining power. With weak bargaining power and little diffusion, the worker keeps a fifth of the gain. The share rises linearly to one as the tool spreads across employers.

The ladder. Figure 5 plots the lowest initial ability at which a worker below the frontier still practises, equation (12), with $\gamma = 0.2$, $\kappa = 0.05$ and $\delta = 0.9$. Without AI, all workers practise under these values. With AI and no tutoring, a frontier at $k = 0.6$ stops practice for every worker below $a_1 = 0.45$. Tutoring with $\tau = 0.5$ lowers the threshold to 0.35 but does not remove it.

TABLE 1. Worker’s share of the net AI gain, by bargaining power and diffusion.

Bargaining power	Share of employers using AI, φ				
	0.00	0.25	0.50	0.75	1.00
$\beta = 0.2$	0.20	0.40	0.60	0.80	1.00
$\beta = 0.5$	0.50	0.62	0.75	0.88	1.00

FIGURE 5. Lowest initial ability at which a worker below the frontier invests in learning ($\gamma = 0.2$, $\kappa = 0.05$, $\delta = 0.9$, $\lambda = 0.2$). Workers between zero and the curve stop learning.

11 Empirical implications and research design

This paper does not estimate the model. This section does three things: it lists the model’s testable predictions, reads the existing experimental evidence against them, and sets out the data and designs that would test them directly.

11.1 Predictions

Table 2 collects the predictions. Most concern the *shape* of effects across workers, not their average size. A test of the model therefore needs individual-level outcomes, a measure of baseline ability that is not affected by AI, and enough workers to estimate effects by ability bin.

11.2 The existing evidence through the lens of the model

The first wave of field and laboratory experiments was not designed to test this model, but several of its results map onto it. In customer support, access to an AI assistant raised resolutions per hour by about 14% on average and by 34% for novice and low-skilled agents, with little effect on the most experienced (Brynjolfsson et al., 2025). In mid-level professional writing, ChatGPT reduced time by 40%, raised quality and compressed differences between participants (Noy and Zhang, 2023). Among software developers, AI coding assistants raised output, with larger gains for less experienced developers (Peng et al., 2023; Cui et al., 2025). In the model, these are jobs in which the tool’s frontier lies above much of the workforce. They are the descending branch of the valley, consistent with Propositions 1 and 2 when k is high relative to G . They are equally consistent with Ide and Talamàs (2025), since the tools studied acted as co-pilots; these experiments cannot tell the two accounts apart.

The experiment of Dell’Acqua et al. (2023) with management consultants is the closest to

TABLE 2. Testable predictions.

	Prediction	What would refute it
P1	Gains are V-shaped in baseline ability, with the minimum near the AI's frontier.	Gains monotone across the whole ability range in a job where the tool solves some but not all problems.
P2	The AI $\times$ skill interaction falls with each frontier advance at fixed leverage.	The interaction rising after a model upgrade that improves accuracy rather than speed.
P3	Log or ratio dispersion falls; dispersion in levels can rise.	Rising Gini or 90/10 ratio with ability-neutral use of the tool.
P4	On tasks outside the frontier, losses are concentrated among the least able, and reduced by verification.	Losses that are larger for experts, or unaffected by review.
P5	Early adopters are the able. Later the least able become the heaviest users, and the middle adopts last.	Adoption that is monotone in ability at all stages.
P6	Pass-through of productivity to wages rises with the share of local employers using the tool.	Pass-through unrelated to employer-side diffusion.
P7	Practice and skill growth of juniors below the frontier fall when the frontier rises; tutoring designs attenuate this.	Faster skill growth among AI-assisted juniors whose tasks lie entirely below the frontier, with no tutoring.

Proposition 5. On tasks inside the tool's frontier, below-average performers gained more than above-average ones. On a task chosen to lie outside it, consultants using AI were less likely to reach the correct answer than those without it. Otis et al. (2024) find that an AI business mentor raised the performance of high-performing Kenyan entrepreneurs and lowered that of low performers. Advice given by the tool had to be judged, and the ability to judge it differed. This is the over-trust case of Proposition 5(iii).

On adoption and wages, survey evidence shows positive selection of early users on education and income (Bick et al., 2024), the early phase of Proposition 6. Linked administrative and survey data for Denmark show average time savings of about 3% and precisely estimated near-zero effects on earnings two years after adoption (Humlum and Vestergaard, 2025), consistent with low pass-through in early diffusion (Proposition 7). None of these studies tests the full shape of the predictions. Most report effects for two or three groups, which cannot distinguish a V from a monotone profile.

The sign of selection, study by study. Proposition 6(iii) implies that comparisons of users with non-users are biased in a direction that depends on where the frontier sits in each sample. Table 3 reads each study's reported ability gradient as a signal of that location and states the implied sign of the selection term. It is a set of falsifiable predictions for anyone with take-up data from these settings, not a re-analysis.

11.3 What a direct test requires

Data. The unit is worker $\times$ task $\times$ time. The minimum requirements are an output measure at the task level, a pre-period ability measure (performance before access, tenure-adjusted), the task's difficulty or category, and a record of AI use. Usage logs give intensity, $\iota(a)$. Combined with a task-level indicator of whether the tool can solve the problem, they also identify the frontier k within the job, which is the model's key latent variable.

TABLE 3. Implied sign of selection in user–non-user comparisons, by study.

Study	Setting	Reported gradient $\Rightarrow$ frontier	Selection
Brynjolfsson et al. (2025)	Customer support	Larger gains for novices $\Rightarrow k$ above most agents	– or ≈ 0
Noy and Zhang (2023)	Professional writing	Gaps compressed $\Rightarrow k$ above the median	– or ≈ 0
Peng et al. (2023); Cui et al. (2025)	Software development	Larger gains for less experienced $\Rightarrow k$ above juniors	– or ≈ 0
Dell’Acqua et al. (2023)	Consulting, inside frontier	Larger gains below average $\Rightarrow k$ high	–
Otis et al. (2024)	Entrepreneurs, advice	Gains for high performers only $\Rightarrow$ judgment binds	+
Bick et al. (2024); Humlum and Vestergaard (2025)	Population surveys, early diffusion	Users more educated and better paid	+

Notes: “Selection” is the sign of $\mathbb{E}[Y_0 | D = 1] - \mathbb{E}[Y_0 | D = 0]$ in equation (9) implied by the model given the frontier’s location. A negative or near-zero sign means that a naive comparison of users with non-users understates the effect, because the lower tail has begun to adopt; a positive sign means it overstates it. The prediction applies to voluntary use within each setting, not to the randomised comparisons the studies report.

Designs.

- (a) *Staggered access with mandated use.* Rollouts across teams provide exogenous timing. Mandated use, or measured take-up by ability bin, separates the gain profile $\Delta(a)$ from selection into use (Proposition 6). The estimand is the effect by ability decile, and the test is for a V with a minimum that is interior and located where the share of problems the tool solves alone equals the worker’s own solve rate.
- (b) *Model upgrades as frontier shocks.* A change in the underlying model within an unchanged interface moves k while holding λ roughly fixed. Proposition 2 predicts that the interaction between AI use and baseline ability falls after the upgrade, and Proposition 1 predicts that the valley moves up the ability distribution.
- (c) *Randomised verification.* Randomly assigning review layers or automated checks shifts $\pi(a)$ without changing k , which tests Proposition 5: losses on tasks beyond the frontier should shrink, most for the least able.
- (d) *Employer diffusion and wages.* Matched employer–employee data with employer-level adoption allow a test of Proposition 7. The pass-through of a worker’s measured productivity gain to wages should rise with the adoption rate among employers in the worker’s local labour market and occupation.
- (e) *Learning.* Following cohorts of juniors hired before and after access, with performance measured on unaided tasks, tests Proposition 8. Randomising a tutoring mode, in which the tool explains rather than answers, would measure τ directly.
- (f) *Speed versus knowledge.* The sharpest contrast with Ide and Talamàs (2025) concerns co-pilots whose main effect is speed: drafting, formatting and search tools that rarely solve a problem the worker could not, but let them get through more. In this model such a tool has a low frontier and high leverage, so its gains rise with ability (Proposition 1(ii)). In theirs, a non-autonomous tool mainly benefits the least knowledgeable. Comparing, within the same job, the ability gradient of gains from a speed-oriented tool with that of

a knowledge-oriented one (or from the same tool before and after a model upgrade that raises accuracy but not speed) separates the two accounts.

None of these designs requires data beyond what large employers, platforms and AI vendors already record. Researchers can rarely obtain it. The model's value at this stage is to say what to ask for and which estimand each design delivers.

12 Conclusion

Generative AI has been described both as a technology that amplifies the advantage of the able and as one that lets the less able catch up. This paper argues that, within a job, both descriptions can be correct for different workers, and that this account can be told apart from an equilibrium account based on the tool's autonomy (Ide and Talamàs, 2025). Separating what the tool knows (its frontier) from how much it speeds up work (its leverage) gives a gain profile that is V-shaped in ability, with the minimum at the frontier. The average complementarity between AI and skill changes sign as the frontier moves. Relative and absolute inequality can move in opposite directions. Judgment becomes the skill-biased margin once the frontier is jagged. Adoption is two-tailed, and selection bias changes sign over time. Workers' share of the gain rises with diffusion. The return to skill below the frontier, and with it the incentive to learn, disappears.

Several extensions are natural. The first is the organisation of work. Once the bottom of the ability distribution can solve what the middle used to, firms may reallocate problems and restructure their hierarchies, as Ide and Talamàs (2025) study in equilibrium. The second is the frontier's own dynamics. The AI's knowledge is partly built from the output of experts, and Proposition 8 suggests the supply of new experts may shrink, so the frontier and the pipeline of human expertise may be linked in a way that a model with exogenous k cannot capture. The third is empirical. The predictions in Section 11 concern the shape of effects across workers and over technology vintages. Testing them requires task-level data with usage logs and baseline ability, which large employers and AI providers already hold.

For policy and management, the model gives three practical points. Average productivity effects say little about who benefits, so evaluations of AI deployments should report effects by baseline ability and by task. Verification is a complement of AI that organisations can buy, and it matters most for their least experienced workers. Finally, early-career training needs explicit design. The tool that makes a junior productive on day one also removes the work through which juniors used to learn, and no individual employer has an incentive to replace it.

References

- Daron Acemoglu. The simple macroeconomics of AI. *Economic Policy*, 40(121):13–58, 2025.
- Daron Acemoglu and David Autor. Skills, tasks and technologies: Implications for employment and earnings. In Orley Ashenfelter and David Card, editors, *Handbook of Labor Economics*, volume 4B, pages 1043–1171. Elsevier, Amsterdam, 2011.
- Daron Acemoglu and Jörn-Steffen Pischke. Why do firms train? theory and evidence. *Quarterly Journal of Economics*, 113(1):79–119, 1998.

- Daron Acemoglu and Pascual Restrepo. The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, 108(6): 1488–1542, 2018.
- Ajay Agrawal, Joshua Gans, and Avi Goldfarb. *Prediction Machines: The Simple Economics of Artificial Intelligence*. Harvard Business Review Press, Boston, 2018.
- Ajay Agrawal, Joshua S. Gans, and Avi Goldfarb. Artificial intelligence: The ambiguous labor market impact of automating prediction. *Journal of Economic Perspectives*, 33(2):31–50, 2019.
- David Autor. Applying AI to rebuild middle class jobs. Working Paper 32140, National Bureau of Economic Research, 2024.
- David H. Autor, Frank Levy, and Richard J. Murnane. The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics*, 118(4):1279–1333, 2003.
- Gary S. Becker. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. National Bureau of Economic Research, New York, 1964.
- Alexander Bick, Adam Blandin, and David J. Deming. The rapid adoption of generative AI. Working paper, National Bureau of Economic Research, 2024.
- Timothy F. Bresnahan, Erik Brynjolfsson, and Lorin M. Hitt. Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *Quarterly Journal of Economics*, 117(1):339–376, 2002.
- Erik Brynjolfsson, Danielle Li, and Lindsey Raymond. Generative AI at work. *Quarterly Journal of Economics*, 140(2):889–942, 2025.
- Zheyuan Kevin Cui, Mert Demirer, Sonia Jaffe, Leon Musolff, Sida Peng, and Tobias Salz. The effects of generative AI on high-skilled work: Evidence from three field experiments with software developers. *Management Science*, 2025. doi:10.1287/mnsc.2025.00535.
- Fabrizio Dell’Acqua, Edward McFowland, Ethan R. Mollick, Hila Lifshitz-Assaf, Katherine Kellogg, Saran Rajendran, Lisa Kraye, François Cadelon, and Karim R. Lakhani. Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality. Working Paper 24-013, Harvard Business School, 2023.
- Tyna Eloundou, Sam Manning, Pamela Mishkin, and Daniel Rock. GPTs are GPTs: Labor market impact potential of LLMs. *Science*, 384(6702):1306–1308, 2024.
- Johan Fellman. The effect of transformations on Lorenz curves. *Econometrica*, 44(4):823–824, 1976.
- Edward Felten, Manav Raj, and Robert Seamans. Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal*, 42(12):2195–2217, 2021.

- Luis Garicano. Hierarchies and the organization of knowledge in production. *Journal of Political Economy*, 108(5):874–904, 2000.
- Luis Garicano and Esteban Rossi-Hansberg. Organization and inequality in a knowledge economy. *Quarterly Journal of Economics*, 121(4):1383–1435, 2006.
- Kunal Handa, Alex Tamkin, Miles McCain, Saffron Huang, Esin Durmus, et al. Which economic tasks are performed with AI? evidence from millions of Claude conversations. Technical Report 2503.04761, arXiv, 2025.
- James J. Heckman and Edward Vytlacil. Structural equations, treatment effects, and econometric policy evaluation. *Econometrica*, 73(3):669–738, 2005.
- Anders Humlum and Emilie Vestergaard. Large language models, small labor market effects. Working Paper 33777, National Bureau of Economic Research, 2025.
- Enrique Ide and Eduard Talamàs. Artificial intelligence in the knowledge economy. *Journal of Political Economy*, 133(12), 2025. doi:10.1086/737233.
- Ulf Jakobsson. On the measurement of the degree of progression. *Journal of Public Economics*, 5(1–2):161–168, 1976.
- Shakked Noy and Whitney Zhang. Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654):187–192, 2023.
- Nicholas G. Otis, Rowan Clarke, Solene Delecourt, David Holtz, and Rem Koning. The uneven impact of generative AI on entrepreneurial performance. Working paper, Harvard Business School, 2024.
- Sida Peng, Eirini Kalliamvakou, Peter Cihon, and Mert Demirer. The impact of AI on developer productivity: Evidence from GitHub Copilot. Technical Report 2302.06590, arXiv, 2023.
- A. D. Roy. Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2):135–146, 1951.

A Proofs

Every closed form below has been verified symbolically, and every general statement has been checked numerically on random draws of (G, H, v) , in `code/01_model_checks.py`. Throughout, $V'(a) = v(a)h(a) > 0$.

Proof of Proposition 1. (i) For $a < k$, $\Delta(a) = (1 + \lambda)V(k) - V(a)$, so $\Delta'(a) = -v(a)h(a) < 0$. For $a > k$, $\Delta(a) = \lambda V(a)$, so $\Delta'(a) = \lambda v(a)h(a) > 0$ when $\lambda > 0$. Both expressions equal $\lambda V(k)$ at $a = k$, so Δ is continuous with its minimum there. (ii) If $a < k$, $Y_1 = (1 + \lambda)V(k)$ and $\partial Y_1 / \partial k = (1 + \lambda)v(k)h(k)$. If $a > k$, Y_1 does not depend on k . Finally, $\partial Y_1 / \partial \lambda = V(\max\{a, k\})$, which is non-decreasing in a . $\square$

Proof of Proposition 2. Since $\text{Var}(a)$ does not depend on (k, λ) , it suffices to sign derivatives of $C(k, \lambda) = \text{Cov}(\Delta(a), a)$. Differentiating under the expectation, $\partial C / \partial k = (1 + \lambda)v(k)h(k)\text{Cov}(\mathbf{1}\{a < k\}, a)$. For $k \in (0, 1)$, $\text{Cov}(\mathbf{1}\{a < k\}, a) = G(k)(\mathbb{E}[a | a < k] - \mathbb{E}[a]) < 0$ because $0 < G(k) < 1$. Next, $\partial C / \partial \lambda = \text{Cov}(V(\max\{a, k\}), a)$. This is the covariance of two non-decreasing functions of a , one of which is strictly increasing on $(k, 1]$, a set of positive measure when $k < 1$, so it is strictly positive. At $k = 0$, $\Delta = \lambda V(a)$ and $C = \lambda \text{Cov}(V(a), a) > 0$. At $k = 1$, $\Delta = (1 + \lambda)V(1) - V(a)$ and $C = -\text{Cov}(V(a), a) < 0$. Since C is continuous and strictly decreasing in k , it has a unique root $\hat{k}(\lambda) \in (0, 1)$, and by the implicit function theorem $\hat{k}'(\lambda) = -C_\lambda / C_k > 0$. In the uniform benchmark, $\mathbb{E}[a \max\{a, k\}] = k^3/2 + (1 - k^3)/3$ and $\mathbb{E}[\max\{a, k\}] = (1 + k^2)/2$, so $\text{Cov}(a, \max\{a, k\}) = (1 - k)^2(1 + 2k)/12$. With $\text{Var}(a) = 1/12$ this gives (4). $\square$

Proof of Proposition 4. (i) Write $X = Y_0 = V(a)$ and $Y_1 = t(X)$ with $t(x) = (1 + \lambda) \max\{x, V(k)\}$. The map t is non-decreasing and $t(x)/x = (1 + \lambda) \max\{1, V(k)/x\}$ is non-increasing for $x > 0$. By Jakobsson (1976) and Fellman (1976), the Lorenz curve of $t(X)$ lies weakly above that of X , and all Lorenz-consistent indices weakly fall. For quantile ratios, monotonicity gives $Q_p(t(X)) = t(Q_p(X))$. Because $t(x)/x$ is non-increasing and $Q_p \geq Q_{p'}$, we have $t(Q_p)/Q_p \leq t(Q_{p'})/Q_{p'}$, which rearranges to $t(Q_p)/t(Q_{p'}) \leq Q_p/Q_{p'}$. For logs, $\log Y_1 = \log(1 + \lambda) + f(\log X)$ with $f(u) = \max\{u, \log V(k)\}$, which is 1-Lipschitz. With $U = \log X$ and U' an independent copy of U , $\text{Var}(f(U)) = \frac{1}{2}\mathbb{E}[(f(U) - f(U'))^2] \leq \frac{1}{2}\mathbb{E}[(U - U')^2] = \text{Var}(U)$. The inequality is strict when censoring binds on a set of positive probability and does not bind everywhere, that is when $0 < G(k) < 1$. The same condition makes the Lorenz ranking strict, so the Gini and coefficient of variation fall strictly. (ii) $\text{Var}(Y_1) = (1 + \lambda)^2 \text{Var}(\max\{Y_0, V(k)\})$ follows directly. In the uniform benchmark, $\mathbb{E}[\max\{a, k\}^2] = (1 + 2k^3)/3$, so $\text{Var}(\max\{a, k\}) = (1 - k)^3(1 + 3k)/12$, and dividing by $\text{Var}(a) = 1/12$ gives (7). Its derivative in k is $-12(1 + \lambda)^2 k(1 - k)^2 < 0$. It equals $(1 + \lambda)^2 > 1$ at $k = 0$ and 0 at $k = 1$, so $\bar{k}(\lambda)$ exists and is unique. At $k = \hat{k}$, $(1 + \lambda) = 1/[(1 - k)^2(1 + 2k)]$, so the variance ratio equals $(1 + 3k)/[(1 - k)(1 + 2k)^2]$. Since $(1 - k)(1 + 2k)^2 = 1 + 3k - 4k^3 < 1 + 3k$, the ratio exceeds one at $\hat{k}$, and because it is decreasing in k , $\bar{k} > \hat{k}$. $\square$

Proof of Proposition 5. (i) $r'(a) = (1 - q)d\pi'(a) > 0$. With $\pi(a) = a$, $r(a) = 0$ gives $a_r = 1 - q/((1 - q)d)$. (ii) A reaching worker obtains $r(a)$ on each problem in $(\max\{a, k\}, 1]$, of which they process $(1 + \lambda)$ times the mass $1 - H(\max\{a, k\})$. In the uniform benchmark with $a \geq k$ the reach gain is $(1 + \lambda)(q - (1 - q)(1 - a)d)(1 - a)$, a concave quadratic in a with first-order condition $2(1 - q)d(1 - a) = q$, that is $a = 1 - q/(2(1 - q)d) = (1 + a_r)/2$. (iii) A worker with $\tilde{\pi}(a) > \pi(a)$ reaches whenever $\tilde{r}(a) \geq 0$ and realises $\Delta(a) + (1 + \lambda)r(a)(1 - H(\max\{a, k\}))$. For $a \geq a_r$, $r(a) \geq 0$ and the realised gain is at least $\Delta(a) \geq 0$. For $a < a_r$, $r(a) < 0$, and the gain is negative whenever $(1 + \lambda)|r(a)|(1 - H(\max\{a, k\})) > \Delta(a)$. For $a < k$, the derivative of the realised gain in k is $(1 + \lambda)h(k)(1 - r(a)) > 0$ since $r(a) \leq q < 1$. For $a \geq k$, the gain does not depend on k . $\square$

Proof of Proposition 6. (i) By Proposition 1, Δ is strictly decreasing then strictly increasing, hence quasi-convex. The set $\{a : \Delta(a) < c\}$ is therefore an interval, empty or containing k , and the set of adopters is its complement in $[0, 1]$. (ii) For $a < k$, $\Delta(a) \geq c$ if and only if

$V(a) \leq (1 + \lambda)V(k) - c$. This set contains $a = 0$ if and only if $(1 + \lambda)V(k) \geq c$. For $a \geq k$, $\Delta(a) = \lambda V(a) \geq c$ if and only if $a \geq V^{-1}(c/\lambda)$. If $\lambda V(k) \geq c$, the minimum gain clears the cost and everyone adopts. (iii) Writing $Y = Y_0 + D\Delta$ gives (9). If only the upper tail adopts, adopters are $[a_H, 1]$ and non-adopters $[0, a_H)$, and $\mathbb{E}[Y_0 | D = 1] > \mathbb{E}[Y_0 | D = 0]$ because V is increasing. Section 10 gives an example in which the selection term becomes negative once the lower tail appears. Finally, $\Delta \geq c$ on adopters and $\Delta < c$ on non-adopters by definition. Under random offers with voluntary take-up, the intention-to-treat effect at a is $\mathbb{E}[D\Delta | a] = D(a)\Delta(a)$. $\square$

Proof of Proposition 7. Substituting $\omega(a) = V(a) + \varphi\Gamma(a)$ into the Nash wage gives $w = V + \varphi\Gamma + \beta(1 - \varphi)\Gamma$, which is (10). The firm receives $V + \Gamma - w = (1 - \beta)(1 - \varphi)\Gamma$, and the vendor receives p . The share $s = \varphi + \beta(1 - \varphi)$ has $\partial s / \partial \varphi = 1 - \beta$. With $\varphi = 1$, $w = V + \Gamma = Y_1 - p$, whose derivative in a is 0 for $a < k$ and $(1 + \lambda)v(a)h(a)$ for $a > k$. Without AI the derivative is $v(a)h(a)$ everywhere. $\square$

Proof of Proposition 8. (i) Without AI, the objective is $-\kappa e + \delta(a_1 + \gamma e)$, which is linear in e with slope $\delta\gamma - \kappa$. (ii) With AI, the objective $-\kappa e + \delta[(1 + \lambda)\max\{a_1 + (1 + \tau)\gamma e, k\} - p]$ is the maximum of two affine functions minus an affine function. It is convex in e and is maximised at $e \in \{0, 1\}$. If $a_1 \geq k$, the comparison of the two corners reduces to $\delta(1 + \lambda)(1 + \tau)\gamma \geq \kappa$. If $a_1 < k$, practice yields $-\kappa + \delta(1 + \lambda)\max\{a_1 + (1 + \tau)\gamma, k\}$ against $\delta(1 + \lambda)k$ without practice. Because $\kappa > 0$, practice is chosen if and only if $\delta(1 + \lambda)(a_1 + (1 + \tau)\gamma - k) \geq \kappa$, which is (12). (iii) $\partial a_1^* / \partial k = 1$, $\partial a_1^* / \partial \tau = -\gamma$ and $\partial a_1^* / \partial \lambda = -\kappa / (\delta(1 + \lambda)^2)$. $\square$