What Should We Tax When We Tax AI?*

Giacomo Brusco
Francesco Capozza

This version: August 28, 2026

Abstract

Generative AI is a metered intermediate input: concentrated in supply, powered by imperfectly priced electricity, and substitutable for some labor but not all. Each feature has its own tax proposal: a token tax, an excess profits tax, an electricity levy, a robot tax. We embed all four in one task model with Roy sorting across tasks and free entry into provision, and derive the optimal ad valorem tax on AI usage. It has three terms, all zero at the Diamond–Mirrlees benchmark. Predistribution rises with the substitutability of AI for routine labor but is scaled by an *exposure gap* that the literature sets to one by assumption and that no plausible grouping of workers delivers. A Pigouvian term gives the residual external cost as a share of the pre-tax price, a few percent for a token against sixty percent or more for a kilowatt-hour at the data center. And a negative market power term is scaled by three things that halve it relative to its textbook size: the rent share of the markup, which recurring training costs keep low; the welfare weight on shareholders, which is pinned by consistency with the weight placed on routine workers rather than chosen; and the fraction of the correction that free entry dissipates in duplicated training runs, which the demand system determines. Netting the three, the efficient wedge on AI usage is at or below zero across almost the whole parameter space, and above five percent in under one percent of it. The rates under discussion are an order of magnitude too high. Tax the kilowatt-hour, not the token; the rent, not the transaction.

Keywords: Generative AI; intermediate inputs; production efficiency; Pigouvian taxation; rent taxation; automation.
JEL: H21, H23, Q58, O33.

*Brusco: UB, IEB: gbrusco@ub.edu; Capozza: UB, IEB, CESifo, IZA@LISER: francesco.capozza@ub.edu. We do not report any conflict of interest.

1 Introduction

Generative AI enters the economy as an input to cognitive work. Earlier automation waves substituted for codifiable tasks, rules that could be written down and handed to a machine, and this made the incidence question relatively clean. Middle-skill routine occupations lost ground, and the polarization literature traced the consequences. Modern systems learn by example, acquire knowledge that resists codification, and scale across compute instead of across people (Ide and Talamàs, 2025). They therefore reach work that no earlier technology reached, which unsettles the standard account of who loses.

The productivity gains are real and measurable in deployed settings. Their incidence is not settled. The same model can compress or widen the earnings distribution depending on whether it advises a worker or replaces one, and firms decide which. Autor (2024) makes the case for compression. Because AI supplies expert judgment instead of executing rules, it can extend high-stakes decision work to people who lack the credential for it today, which would raise wages in the middle of the distribution. Autor and Thompson (2025) show that the answer depends on whether the technology takes over the low-expertise or the high-expertise part of a job.

Diffusion is also uneven and shallower than the public discussion suggests. Automating an automatable task is often not worth the compute it consumes, so adoption stops well short of exposure. The Census Bureau's firm survey finds AI in use at 18 percent of firms, 32 percent weighted by employment, and most adopters use it in three or fewer business functions (Bonney et al., 2026). Two of its findings bear on the case for a robot tax. Two thirds of firms reporting any task effect report augmentation alone, and AI-driven employment declines appear at 2 percent of firms; 16 percent of adopters report replacing existing software or equipment with AI. So far the technology displaces capital more often than labor.

An input this pervasive, this unevenly adopted, and this ambiguous in its distributional effect is a poor candidate for an instrument that assumes it knows who loses.

Four proposals to tax artificial intelligence are on the table. A levy on usage, charged per token or per call. An excess profits tax on the firms that sell it. A charge on the electricity its data centers consume. And a robot tax, levied on AI because it displaces workers. Each has a constituency and each has a story, but the four are usually discussed as variations on a single impulse—that a fast-growing and profitable sector should pay something—rather than as instruments aimed at different problems.

They are aimed at different problems, and the problems do not point the same way. A tax on tokens does nothing about the carbon intensity of the grid serving a data center.

A tax on profits does nothing about displacement. And one of the four runs backwards relative to the others: if AI is sold above marginal cost, it is consumed too little, and correcting that calls for a lower price, not a higher one. Whether a tax on AI usage should be positive at all is therefore not obvious, and cannot be settled by adding up arguments that were each developed in isolation.

The natural benchmark says the answer is zero. Generative AI is an intermediate input; firms buy it to get tasks done. Diamond and Mirrlees (1971) show that when the government can tax final goods and incomes optimally, production efficiency should be preserved and intermediate inputs left untaxed, whatever their distributional consequences downstream. The benchmark is not wrong. It is conditional, and its conditions are precisely the four features that make this case interesting: a missing market for the environmental cost of the electricity, market power in supply, a redistributive motive the income tax cannot fully serve, and an open economy in which some of the rent accrues abroad. Each is a reason the benchmark might fail. None of them has been priced against the others in a single model.

This paper builds that model and prices them. AI is a metered service produced from compute and electricity by a sector with a markup, and used in a subset of tasks that would otherwise be performed by labor. Workers sort across tasks on comparative advantage. Entry into provision is free but requires a large fixed cost, so the number of providers responds to policy. The government has a linear income tax, a usage tax on AI, a cash flow tax on provider rents, and a Pigouvian charge at the data center meter, and it cannot tag workers by type.

The optimal ad valorem tax on AI usage is the sum of three terms, one for each coherent argument in the debate:

$$t_U^\star \;=\; \underbrace{\text{predistribution}}_{\text{wage compression}} \;+\; \underbrace{\text{Pigouvian}}_{\text{external cost / price}} \;-\; \underbrace{\text{market power}}_{\text{underprovision}}.$$

All three vanish at the Diamond–Mirrlees benchmark, and each is a sufficient statistic that can in principle be measured. Three findings follow.

First, the externality is real but it is on the wrong base. The uninternalized external cost of a unit of AI is a few percent of the price of a token, dominated by unpriced carbon. The same cost expressed per kilowatt-hour at the data center is sixty percent or more of the price of that electricity. The ratio of harm to value also varies by an order of magnitude across workloads—a short text query and an agentic task differ hundredfold in energy per unit of delivered value—so no single ad valorem rate on usage can track

the harm, whereas a per-kilowatt-hour charge tracks it by construction. Pricing at the meter also reaches an efficiency margin that a usage tax cannot touch, which is where the observed order-of-magnitude improvements in energy per token have come from.

Second, the redistributive case is weaker than the robot-tax literature implies, for a reason that literature cannot express. That literature identifies the group a planner wants to protect with the occupation the technology reaches. Doing so fixes at one what we call the *exposure gap*: the difference between two groups' shares of employment in exposed tasks. Modelling task assignment as a Roy choice makes the exposure gap a parameter rather than a convention, and no plausible grouping of workers delivers a value near one. Carrying it explicitly cuts the redistributive term to roughly a fifth of the conventional figure. It also reverses a comparative static. The dispersion of comparative advantage is not a mobility parameter: it governs sorting and switching simultaneously and in opposite directions, so a world in which the losing group is sharply identifiable is a world in which nobody is near the switching margin. Perfect sorting is perfect segregation, and the two-type model is the limit of high dispersion rather than its opposite.

Third, the market-power term is negative, but about half the size the textbook expression gives it. Three scalings that the literature suppresses cut it. Only the pure rent in a markup signals underprovision, and at the frontier most of the markup recovers training costs that recur rather than having been sunk once. The marginal unit of foregone output is entirely provider profit, so the term is a transfer to shareholders and must be weighted as one—and a model that values a dollar to a routine worker above the average cannot value a dollar to a frontier-lab shareholder at the average. And with free entry, part of any correction to underprovision is dissipated in the fixed costs of the entry it induces; we derive that share in closed form rather than assume it, obtaining one under linear demand and a quarter under the constant-elasticity demand our own structure generates. The shareholder weight is derived too, from the requirement that it be consistent with the weight the same planner places on routine workers—a discipline that turns out to cut against the argument, since a planner only mildly averse to inequality discounts shareholders only mildly.

Netting the three, at our central calibration the efficient wedge on AI usage is about minus five percent, and across the full parameter space it is negative in ninety percent of draws, above five percent in under one percent, and above ten percent in none. *Whatever the sign, the debate is conducted over a rate an order of magnitude larger than anything the model supports.* That is the robust finding. The negative sign is the fragile one, and a reader who rejects it should note that setting the market-power term to zero altogether—

on the ground that underprovision is the rent tax's problem and not the usage tax's—does not deliver a usage tax worth legislating, but a wedge of four and a half percent at most.

One affirmative argument for a usage tax survives, and it is the one least often made. Compute is mobile: a unilateral charge on data center electricity relocates load rather than abating it, and a destination-based levy on AI *consumption* is the backstop that prevents this. That argument is about leakage, not about robots or rents, and it rests on a parameter nobody has estimated.

A sufficient-statistics expression can be read two ways and we mean the weaker one. The strong reading is a menu of legislation. The weak reading is a diagnostic: it says which measurements would change the answer, and which would not. On that reading the paper's practical contribution is a short list. The share of provider revenue absorbed by recurring training and capacity investment decides both how much rent a cash flow tax can reach and how large the market-power correction is, so it decides between the two live positive proposals; it is an accounting quantity a disclosure rule could produce. The exposure gap decides the size of the redistributive term and has never been reported, because it is set to one by assumption. Neither requires new econometrics.

The paper sits at the intersection of four literatures and takes something from each. From the theory of *intermediate-input taxation* we take the Diamond–Mirrlees benchmark and the two classical routes around it: Naito (1999), where restricted instruments and endogenous wages break production efficiency, and Sandmo (1975), whose additivity of Ramsey and Pigouvian terms is the structure of our main result. Scheuer and Werning (2016) show that in exactly the assignment and Roy models we use, general-equilibrium wage effects wash out of the optimal *nonlinear* income tax, so our redistributive term should be read throughout as an upper bound generated by a restricted instrument set.

From the *optimal taxation of automation* we take the robot-tax formulation of Costinot and Werning (2023), Thümmel (2023) and Guerreiro et al. (2022), whose routine and non-routine occupational choice is closest to our own, and depart from all three on the exposure gap as described above. Rothschild and Scheuer (2013) supply the template for optimal taxation with Roy sorting.

From the *economics of AI and energy* we take the positive facts—International Energy Agency (2026), Shehabi et al. (2024), Luccioni et al. (2024) on load and energy per query, and the contested evidence on grid cost allocation—and supply the normative half, which is largely missing. From the *taxation of the digital economy* we take the design of destination-based and cash flow instruments (Auerbach and Devereux, 2018) and offer a distinction current practice blurs: a cash flow tax on rents and a turnover tax on revenue

are not variants of one instrument but opposites, one neutral and one distortionary, and what gets legislated is usually the second.

To the *industrial organization of taxation* we contribute two results we have not seen stated for this setting. With free entry and fixed costs, an ad valorem tax on a concentrated sector raises the equilibrium markup, and we derive the amplification in closed form; in a sector where fixed costs run to billions and marginal costs to fractions of a cent, the tax meant to fall on incumbents partly protects them, which yields a testable political-economy prediction about who should be expected to support it. And the negative market-power term in every formula of this kind is scaled by the three quantities described above, whose product is under a third.

Section 2 sets out the four claimed departures from the benchmark and the evidence for each. Section 3 presents the model and derives the optimal usage tax. Section 5 reads the three terms in turn, which is where the exposure gap, the entry offset and the ownership weight are derived, and where the negative sign is explained. Section 6 calibrates. Section 7 scores the four proposals against each other on revenue, distortion, incidence and administrability, and asks why the worst instrument is the one most often proposed. Section 8 concludes. Proofs and extensions are in the appendices.

2 Four claimed departures from the benchmark

The case for treating AI differently from any other intermediate input rests on four claims. Each corresponds to a condition of the production efficiency theorem, and each is a claim about the world rather than about the model, so we state the evidence for each before building anything on it.

A missing redistributive instrument. Governments cannot observe whether a given worker is a substitute or a complement for AI, so they cannot levy a tax that depends on it. When exposure is unobservable, the income tax cannot condition on the margin along which AI redistributes, and a wedge on the input becomes a second-best substitute—the Naito (1999) channel. This is the robot-tax argument, and it is the oldest and least contested of the four. Its force depends on how imperfect the income tax is, and Section 5.3 shows the term collapses to zero under a fully nonlinear schedule.

A missing market. The electricity data centers consume carries two costs its price does not fully reflect. The first is environmental. At a social cost of carbon of $190 per ton (U.S.

Environmental Protection Agency, 2023) and a grid intensity of 0.4 kg per kWh serving U.S. data centers (International Energy Agency, 2025), unpriced carbon damage alone is roughly \$0.076 per kWh where no carbon price applies—close to the entire industrial electricity price of about \$0.08.

The second is grid cost shifting. Under cost-of-service regulation a utility recovers costs across customer classes on the principle of cost causation: each class should pay for the infrastructure built to serve it (Borenstein and Bushnell, 2022). Large new loads that require transmission upgrades and generation capacity strain that principle, because the incremental cost is lumpy and the allocation rules were not written with hundred-megawatt single customers in mind. Whether shifting has actually occurred is genuinely disputed—Energy and Environmental Economics (E3) (2026) finds large-load tariffs largely recover incremental cost, while Jacobs (2025) and Monitoring Analytics (2025) find substantial socialization—and Section 6 carries the wedge γ as a free parameter over $[0, 0.03]$ per kWh rather than taking a side. Figure D.2 shows the level conclusion survives the whole range, because carbon dominates; what γ governs is the distributional character of the status quo, which is the subject of Section 6.5.

Market power in supply. Frontier model provision is a natural oligopoly. Training a frontier model costs hundreds of millions to billions of dollars and the cluster required is comparable; marginal cost per token is fractions of a cent. Reported gross margins on frontier inference run to fifty or sixty percent, implying a markup around 2.5 against an economy-wide 1.25 (De Loecker et al., 2020). A good sold above marginal cost is consumed too little, which is the claim that runs opposite to the other three. Section 5.6 shows how much of that markup is genuine rent rather than fixed-cost recovery, and how much of the resulting correction survives free entry.

Foreign ownership of the rents. Those rents accrue disproportionately to shareholders in a small number of countries. For a jurisdiction that imports AI services, a tax on usage is partly a tariff that shifts rent across the border, and the optimal-tax calculus differs accordingly. This is why ownership rather than technology turns out to set the sign of the optimal wedge.

2.1 What makes AI different from steel

An input tax needs a reason, and production efficiency says the default reason is none. So why does compute clear the bar when a lathe, a truck, or a ton of steel does not?

Three features do the work, and it is their combination rather than any one of them that is unusual.

The externality is badly matched to the value of the good. Where abatement is possible at the point of harm, a tax on the input is a poor instrument for any externality, which is the standard targeting principle. What is unusual here is the *dispersion* of the harm-to-value ratio. Energy per query varies by more than two orders of magnitude across workloads, from a short text completion to an agentic task, while price varies far less, so the ratio of external cost to price differs several-fold across the queries a single ad valorem rate would cover. A ton of steel is a ton of steel; a token is not a token. Section 6 quantifies this.

The externality is spatially discretionary in the extreme. A steel mill's emissions are tied to where the mill is, and the mill is where the ore and the workforce are. An inference request can be served from anywhere with power and fibre, at a latency cost measured in milliseconds. The elasticity of the location of the harm with respect to a local charge is therefore far higher than for any conventional polluting activity, which is what makes leakage a first-order concern and motivates the destination-based backstop of Appendix B.3.

AI is an unusually clean case of a factor-replacing input. A truck complements a driver and a lathe complements a machinist. Generative AI is sold explicitly as a substitute for cognitive labor, which is what gives the redistributive claim its purchase and what the task structure of Section 3 is built to represent.

None of the three is unique to AI taken singly. Electricity has unpriced carbon; pharmaceuticals have market power; imported goods carry terms-of-trade considerations. What is unusual is their coincidence in one input, and that coincidence is why the four arguments have to be priced against each other rather than in turn.

3 Model

3.1 Environment

Tasks and production. A single final good Y is produced competitively from two task bundles. A share $\lambda \in (0, 1)$ of production value comes from *automatable* tasks, in which AI and routine labor are substitutes. The remaining share comes from *non-automatable*

tasks, performed by non-routine labor. With X_A and X_M the two bundles,

$$Y \;=\; \Lambda\, X_A^{\lambda}\, X_M^{1-\lambda}, \qquad X_A \;=\; \Big[\beta\, L_R^{\frac{\sigma-1}{\sigma}} + (1-\beta)\, a^{\frac{\sigma-1}{\sigma}}\Big]^{\frac{\sigma}{\sigma-1}}, \qquad X_M \;=\; L_N, \tag{1}$$

where a is AI services, L_R and L_N are routine and non-routine labor, and Λ is a constant.[1]

Write q for the buyer price of AI relative to the routine wage, c_A for the unit cost of the automatable bundle, and

$$s \;\equiv\; \frac{q\,a}{c_A X_A} \;=\; \frac{(1-\beta)^{\sigma} q^{1-\sigma}}{\beta^{\sigma} + (1-\beta)^{\sigma} q^{1-\sigma}} \qquad \text{so that} \qquad \frac{\mathrm{d}\ln s}{\mathrm{d}\ln q} \;=\; -(1-s)(\sigma-1), \tag{2}$$

for the AI cost share within that bundle. Exposure, the share of national income spent on AI, is $\theta \equiv \lambda s$.

The elasticity σ is the central technological parameter: values above one make AI a gross substitute for routine labor within the automatable bundle, which is what gives the displacement concern its content, and $\sigma = 1$ shuts the redistributive channel entirely. It is also the parameter we know least well, so Appendix D.1 reports every result across $\sigma \in [1.2, 3]$. The outer Cobb–Douglas nest is not innocuous either, and Lemma D.1 in Appendix A records what changes when it is relaxed.

The AI sector. AI services are supplied by a sector with market power. One unit requires compute at resource cost m_0 and e kilowatt-hours of electricity at the price p_e paid by data centers. Providers also choose an efficiency level $\xi \geq 0$, which stands for chip generation, model architecture, serving stack, cooling, and the carbon intensity of the site. Efficiency lowers energy per unit, $e = e(\xi)$ with $e' < 0$ and $e'' > 0$, at a resource cost $\varphi(\xi)$ with $\varphi' > 0$ and $\varphi'' > 0$. Marginal cost and the efficiency choice are

$$m(\xi) \;=\; m_0 + \varphi(\xi) + p_e\, e(\xi), \qquad\qquad \varphi'(\xi) \;=\; -\,p_e\, e'(\xi). \tag{3}$$

The second condition is the engine of the targeting result below. The efficiency margin responds to the price of energy and to nothing else. It does not respond to the price of AI output.

Providers are an oligopoly: n symmetric firms with the constant marginal cost $m(\xi)$ of (3) and a fixed cost F covering the training run and the compute cluster. They compete

[1]The Cobb–Douglas outer aggregator is for transparency: it fixes the value shares of the two bundles and lets every result be stated in cost shares. A CES outer nest changes magnitudes but not structure; Lemma D.1 bounds the correction.

in quantities, so the equilibrium buyer price carries a markup over marginal cost,

$$p \;=\; \mu\, m(\xi)\,(1+t_U), \qquad \mu \;=\; \frac{\sigma_q}{\sigma_q - 1/n} \;>\; 1, \tag{4}$$

decreasing in n. Entry is free, so n adjusts until profit is zero; Section 5.5 shows this makes μ a function of the tax rather than a parameter, which is where the entry amplification comes from.

The markup requires interpretation, and getting it wrong is what makes the market-power debate go in circles. A firm that charges above short-run marginal cost is not thereby earning rent: with a large fixed cost F, free entry drives profit to zero and the entire margin recovers F. Only the part of the markup in excess of fixed-cost recovery signals underprovision, so we write

$$r \;\equiv\; \frac{\text{pure rent per unit}}{(\mu - \bar{\mu})\, m} \;\in\; [0,1] \tag{5}$$

for the pure-rent share of the *excess* markup, measured against the economy-wide markup $\bar{\mu}$ rather than against marginal cost. Section 5.6 shows that r is not a free parameter but the complement of the share of revenue absorbed by recurring investment, and that it therefore governs the rent tax and the market-power correction jointly.

Because the same sector must deliver both the targeting result and the entry result, we fix one market structure and use it throughout rather than switching between a monopolist for one argument and free entry for another. Appendix C.4 records what this buys and what it costs.

Externalities. Electricity for AI produces carbon at intensity ζ kg per kWh, valued at social cost δ, and its delivery causes capacity and transmission costs of which a portion γ per kWh is recovered from other ratepayers rather than from the facility. Write $d \equiv \zeta\delta + \gamma$ for the residual external cost per kilowatt-hour net of any carbon price already in force, and $D \equiv d\, e(\xi)$ for the external cost per unit of AI. Two features of d matter. It is defined per kilowatt-hour, not per unit of output, so the mapping to a usage tax runs through $e(\xi)$ and therefore through an endogenous object.

The second feature is the hinge of the targeting argument. There are two distinct margins by which the external cost of a unit of AI can fall: the *scale* margin, fewer units bought, and the *efficiency* margin, fewer joules per unit. A wedge on output reaches only the first. A charge on energy reaches both, because it enters the provider's marginal cost through $p_e e(\xi)$ and so changes the efficiency ξ that minimizes it. This is why Proposition

3 is not merely the standard targeting principle restated: the instruments differ not only in precision but in the set of margins they can touch at all.

Households: types are ex ante, tasks are chosen. We depart here from the two-type structure used in most of this literature, and the departure changes the object welfare weights attach to. If workers are permanently assigned to tasks, "routine worker" is a type and a welfare weight can be attached to it. If workers sort across tasks in response to wages (which is the empirically relevant case, and the one Section 5.4 takes seriously) then "routine worker" is an *outcome*, its membership moves when the tax moves, and a welfare weight attached to it is not a well-defined object. We therefore put the sorting margin in the primitive model and define welfare on the objects that do not move.

There is a unit mass of workers, divided into two *ex ante* groups $i \in \{1, 2\}$ of equal mass. A worker of group i draws task-specific efficiencies (z_R, z_N) independently from a Fréchet distribution with shape $\kappa > 0$ and group-specific scales A_{iR}, A_{iN}, and supplies her labor to the task with the higher return, so the wage she commands is $\max_j A_{ij} z_j w_j$. Groups differ only in comparative advantage,

$$\varrho \;\equiv\; \ln \frac{A_{1R}}{A_{1N}} \;-\; \ln \frac{A_{2R}}{A_{2N}} \;>\; 0, \tag{6}$$

so group 1 is the group with comparative advantage in the tasks AI reaches. Group membership is the primitive: it is fixed, unobservable, and untaggable. Occupation is not. Hours are supplied with elasticity ε_L.

Two objects follow from the Fréchet assumption and carry the whole distributional analysis. The share of group i working in routine tasks and the group's expected wage index are

$$\ell_{iR} \;=\; \frac{(A_{iR} w_R)^\kappa}{(A_{iR} w_R)^\kappa + (A_{iN} w_N)^\kappa}, \qquad W_i \;=\; \Gamma\big(1 - \tfrac{1}{\kappa}\big) \Big[(A_{iR} w_R)^\kappa + (A_{iN} w_N)^\kappa\Big]^{1/\kappa}, \tag{7}$$

and differentiating W_i gives the statistic that replaces "the routine wage":

$$\mathrm{d} \ln W_i \;=\; \ell_{iR}\, \mathrm{d} \ln w_R + (1 - \ell_{iR})\, \mathrm{d} \ln w_N, \qquad \text{hence} \qquad \mathrm{d} \ln \frac{W_1}{W_2} \;=\; \Delta_\ell\, \mathrm{d} \ln \frac{w_R}{w_N}, \qquad \Delta_\ell \;\equiv\; \ell_{1R} - \ell_{2R}. \tag{8}$$

Equation (8) is an envelope result: the marginal switcher is indifferent between tasks, so reallocation has no first-order welfare effect and the entire burden of a relative task-wage movement falls on inframarginal workers, in proportion to the share of the group that is exposed. We call $\Delta_\ell \in [0, 1]$ the *exposure gap*. It is the fraction of a relative task-

wage movement that survives as a relative movement between the two groups the planner actually cares about, and it is in principle measurable.

Government and instruments. The government levies a linear income tax τ with a lump sum transfer and cannot condition either on group. It also has access to a usage tax t_U, a cash flow tax z on the supernormal profits of providers, and a Pigouvian charge t_E per kilowatt-hour at the data center. Social welfare is Bergson–Samuelson with weights g_i defined on *ex ante groups*, and we write $\hat{g}_1 \equiv g_1 - \bar{g}$ for the excess weight on group 1 relative to the income-weighted average, normalized so that a uniform proportional change in all wages has no first-order distributional value. Because the weights sit on groups rather than on occupations, they are invariant to the reform, and the normalization $\sum_i \hat{g}_i W_i = 0$ holds before and after.

We assume throughout that τ is set optimally. This permits the envelope argument of Lemma 3: a reform that raises all wages proportionally has no first-order welfare effect, because the income tax has already been adjusted to the point where the marginal value of redistribution equals its marginal cost. Only the *relative* wage movement survives into the optimal-tax formula, which is what makes the predistribution term a statement about the wage *gap* rather than about the wage level.

Assumption 1 (Restricted instruments)**.** The government cannot tag workers by ex ante group, cannot observe the AI intensity of individual jobs, and is restricted to a linear income tax, so it cannot implement a schedule that separates groups perfectly.

3.2 Equilibrium and comparative statics

All proofs are collected in Appendix A.

Factor payments follow from (1). The automatable bundle receives λY, of which a share $1-s$ goes to routine labor and s to AI; the non-automatable bundle receives $(1-\lambda)Y$, all of it to non-routine labor:

$$w_R L_R = \lambda(1-s)\,Y, \qquad w_N L_N = (1-\lambda)\,Y, \qquad q\,a = \lambda s\,Y = \theta Y. \tag{9}$$

Lemma 1 (Exposure)**.** *A small increase in the buyer price of AI reduces real output at a rate equal to the AI cost share:* $\mathrm{d}\ln Y/\mathrm{d}\ln q = -\theta = -\lambda s$.

Lemma 1 imposes a useful discipline. Because exposure is small, both the output cost of a usage tax and its revenue are small. Whatever the case for such a tax, it is not a fiscal case, and Section 6 puts numbers on this.

Lemma 2 (Relative wages and the mobility elasticity)**.** *Let $\mathcal{E}$ denote the elasticity of relative task labor supply to the relative task wage generated by reallocation. Then*

$$\frac{\mathrm{d}\ln(w_R/w_N)}{\mathrm{d}\ln q} \;=\; \frac{s\,(\sigma-1)}{1+\varepsilon_L+\mathcal{E}}, \qquad \textit{and under (7)} \qquad \mathcal{E} \;=\; \kappa\big(1-\Delta_\ell^2\big).$$

A higher price of AI compresses the task wage gap if and only if AI and routine labor are gross substitutes, and the compression is damped by hours adjustment and by reallocation.

The identity $\mathcal{E} = \kappa(1-\Delta_\ell^2)$ is the hinge of everything that follows. The Fréchet shape κ is not "mobility": it governs the dispersion of comparative advantage, and so determines both how sharply groups sort into tasks and how much probability mass sits near the switching margin. These pull opposite ways, and Section 5.4 works out the consequence, which is that the fixed-assignment two-type model is the limit $\kappa \to \infty$ rather than $\kappa \to 0$.

Corollary 1 (The labor share)**.** *Aggregate labor income is $(1-\theta)Y$, and $\mathrm{d}\ln(1-\theta)/\mathrm{d}\ln q = \lambda s(\sigma-1)(1-s)/(1-\lambda s) > 0$ if and only if $\sigma > 1$. Cheaper AI lowers the labor share when AI and routine labor are gross substitutes.*

Finally, the own price elasticity of AI demand combines the within-bundle substitution effect with the scale effect of Lemma 1:

$$\sigma_q \;\equiv\; -\frac{\mathrm{d}\ln a}{\mathrm{d}\ln q} \;=\; \underbrace{\sigma(1-s)}_{\text{substitution}} \;+\; \underbrace{s+\theta}_{\text{scale}} \;\approx\; \sigma \quad \text{when } s,\theta \text{ are small.} \tag{10}$$

Two results are needed before the tax formula, because each underpins a step that is otherwise a leap. The first makes precise the sense in which the income tax neutralizes the common wage movement. The claim is often asserted; here we state exactly what it requires.

Lemma 3 (Envelope neutrality of the common wage component)**.** *Suppose the government sets the linear income tax (τ, T) optimally, worker preferences are weakly separable between consumption and labor, and the wage response to a reform of t_U decomposes into a common proportional component $\mathrm{d}\ln\bar{w}$ and a relative component $\mathrm{d}\ln(w_R/w_N)$. Then the common component enters welfare only as a change in aggregate real income, valued at the marginal cost of public funds and already recorded on the resource side of the reform through $\mathrm{d}\ln Y$; it carries no first-order* distributional *value, because at the optimal linear tax the income-weighted average welfare weight equals that same shadow value. The distributional value of the reform is therefore carried entirely by the relative component,*

weighted by the excess weight $\hat{g}_1$. If instead the reform changes a moment of the earnings distribution that a linear tax cannot replicate, a residual common-component term of order $(\mathrm{d}\ln\bar{w}) \cdot \mathrm{Cov}_j(\hat{g}_j,\, rank_j)$ survives, and the shortcut is exact only when that covariance is zero.

The decomposition into a common and a relative part is a property of the technology rather than an assumption: with constant returns and competitive factor markets each wage equals its marginal product, and a Cobb–Douglas or CES task structure makes any shock resolve into a term that scales all marginal products together and a term that tilts them. The same structure underlies the general-equilibrium incidence analysis of Sachs et al. (2020), and the neutrality of the common component under an optimal linear tax is the Naito–Saez logic (Naito, 1999; Saez, 2004), itself a linear-tax specialization of Stiglitz (1982). The shortcut is not free: it holds exactly when the income tax spans the distributional content of the common wage movement, which for a linear tax means a proportional shift that τ and T can undo. A technology that fattens one tail of the earnings distribution violates this, and the residual term the lemma isolates is what a richer nonlinear analysis would carry. We flag at each use that the shortcut rests on Lemma 3 and is exact only under its covariance condition.

Exposure also cancels for reasons that do not depend on the Cobb–Douglas outer nest. Lemma D.1 in Appendix D.8 replaces it with a CES aggregator of elasticity η and shows the predistribution term is multiplied by $1 + (\eta - 1)\phi$ with $\phi = O(\theta)$, so at the observed exposure the correction is below two percent of the term for any $\eta \in [0.5, 2]$. The cancellation is exact to first order and accurate to second; we use Cobb–Douglas in the body for readability and rely on the lemma for generality.

3.3 The benchmark

Proposition 1 (Production efficiency)**.** *Suppose there are no externalities ($d = 0$), the AI sector is competitive, and the government can tag worker types, so that Assumption 1 fails. Then $t_U^\star = 0$. Any revenue raised by a tax on AI purchases can be raised at strictly lower welfare cost by income and final good taxes.*

Proposition 1 is the right starting point and it is frequently forgotten. A usage tax distorts the choice of how to perform a task, raises the cost of every unit of output that uses AI, collects revenue on a base that is currently trivial, and buys nothing that the income tax cannot buy more cheaply. Every argument for such a tax must therefore be

an argument that one of its hypotheses fails. There are exactly three ways for that to happen here, and the next section prices each of them.

One hypothesis deserves naming rather than leaving implicit, because the entire predistribution term rests on it. Assumption 1 restricts the government to a *linear* income tax, and that restriction is what makes the general-equilibrium wage response something the planner wants to manipulate—the Naito channel. It is not harmless. Scheuer and Werning (2016) show that in exactly the assignment and Roy models we use, general-equilibrium wage effects wash out of the optimal *nonlinear* income tax, which reverts to its fixed-wage Mirrlees form, and production efficiency is restored. Read through that result, our predistribution term measures not a feature of the technology but how far the income tax falls short of the nonlinear optimum. We keep the linear restriction because it describes most real tax systems' capacity to target occupational exposure, and because it is the assumption most favorable to the instrument we argue against. But the term belongs in the right category: an upper bound generated by a restricted instrument set.

4 The optimal tax on AI usage

Consider a marginal reform dt_U, with the income tax adjusted to balance the budget and set optimally. Welfare in money metric moves for four reasons. Firms adjust their AI use, and each unit they no longer buy had a private value q but a social resource cost that must be counted carefully: the short-run resource cost m, plus the fixed cost the unit helps defray, less the pure rent it embeds. The rent is a transfer, domestic or foreign, not a resource saving, so only the rent portion of the markup drives a wedge between private and social value on this margin. Each unit also carried an external cost D. Relative wages move, which has distributional value because types cannot be tagged. All purely mechanical transfers cancel, and by the envelope theorem, whose scope we make precise in Lemma 3, the common component of the wage response has no first-order *distributional* effect; its efficiency value is the change in aggregate real income already recorded by Lemma 1, so it is booked once, on the resource side, and not again in the distributional term.

Proposition 2 (The optimal usage tax)**.** *Under Assumption 1, with the income tax set*

optimally and market structure fixed, the optimal ad valorem tax on AI purchases satisfies

$$\boxed{t_U^\star \;=\; \underbrace{\frac{(1-\tau)\,\hat{g}_1\,\Delta_\ell\,(\sigma-1)(1-s)}{(1+\varepsilon_L+\mathcal{E})\,\sigma_q}}_{\text{predistribution, } \Omega/p} \;+\; \underbrace{\frac{D}{p}}_{\text{Pigouvian}} \;-\; \underbrace{(1-f)\,r\left(1-\frac{\bar{\mu}}{\mu}\right)}_{\text{market power (rent only)}}} \tag{11}$$

where $\bar{\mu}$ is the markup prevailing in the rest of the economy, r is the pure-rent share of the AI markup from (5)*, Δ_ℓ is the exposure gap of* (8) *and $\mathcal{E} = \kappa(1-\Delta_\ell^2)$ the reallocation elasticity of Lemma 2. When s and θ are small, $\sigma_q \to \sigma$ and the first term simplifies to $(1-\tau)\hat{g}_1\Delta_\ell(\sigma-1)/[\sigma(1+\varepsilon_L+\mathcal{E})]$. The two-type fixed-assignment formula obtains at $\Delta_\ell = 1$, $\mathcal{E} = 0$.*

Each term answers a different question in the debate.

The predistribution term. This is the robot tax. It is the product of three things, and separating them is the main analytical payoff of putting the Roy structure in the primitive model. It is proportional to the excess welfare weight $\hat{g}_1$ on the group the planner wants to help, which is a normative parameter. It is proportional to the exposure gap Δ_ℓ, which measures how much of a relative *task* wage movement survives as a relative movement between the groups the planner cares about, and which is a measurable feature of the sorting pattern rather than a modelling convention. And it is proportional to $(\sigma-1)$, so it vanishes at $\sigma = 1$, since a change in the price of AI then leaves the routine wage share of the automatable bundle unchanged. It is damped by the labor supply elasticity and the reallocation elasticity, because workers then absorb part of the shock in hours and in switching rather than in wages, and by the net of tax rate, because the income tax already absorbs part of any wage movement.

The literature almost universally sets $\Delta_\ell = 1$ by identifying the losing group with the occupation. That is not a conservative simplification but the corner of the parameter space that maximizes the term, and it asserts that every member of the disadvantaged group works in an exposed task and no member of the advantaged group does. Section 5.4 shows what replaces it and Section 6 shows that, of the three factors, the exposure gap is the tightest.

The Pigouvian term. The corrective component is the residual external cost as a share of the pre-tax price, exactly as in Sandmo (1975). It is not the external cost as a share of the electricity price, and that distinction does the analytical work in Section 6.

It also depends on the provider's efficiency choice through $D = d \cdot e(\xi)$, and Proposition 3 shows that the usage tax cannot move ξ.

The market power term. This term is negative and the largest of the three, so most of the argument about the sign of the optimal tax is an argument about it. As written it is also an overstatement, for three reasons developed in Section 5.6; the reader should hold it loosely until then. We note here only the first, which is the correction the debate usually misses. The term is negative only to the extent the markup is *rent* rather than fixed-cost recovery, and the factor r does that work. If the markup merely covers the fixed cost of training frontier models then $r \to 0$, the term vanishes, and there is no market-power case either way: the good is priced at its full social cost of production, fixed cost included, and is not underprovided in any sense a planner should want to fix. Only genuine rent signals underprovision. Section 5.6 adds the other two scalings—the welfare weight on shareholders, and the share of any correction that free entry dissipates—which together with r bring the term to under a third of its face value.

Does the formula call for a subsidy, and why does ownership matter? Both questions have clean answers and both are developed in Section 5.6, but the second is worth previewing because it looks paradoxical. If the problem is that a concentrated sector produces too little, ownership should seem irrelevant to the remedy. It is not, because the marginal unit of foregone output is worth p to a buyer who pays p: her surplus on it is zero, and the whole of the wedge m accrues to the provider. The market-power term is thus not a consumer-surplus correction but a statement about the social value of provider profit—and profit accruing abroad is not part of domestic welfare. At $f = 1$ underprovision has not stopped; its cost is simply borne by someone the domestic planner does not weight.

Corollary 2 (Ownership and the sign of the tax)**.** $\partial t_U^\star / \partial f = r\,(1-\bar{\mu}/\mu) \geq 0$*. An economy that imports all of its AI from foreign providers (*$f = 1$*) faces* $t_U^\star = (1-\tau)\hat{g}_1\Delta_\ell(\sigma-1)/[\sigma(1+\varepsilon_L + \mathcal{E})] + D/p > 0$*, while an economy that owns its providers (*$f = 0$*) faces a strictly lower rate, negative whenever the rent-adjusted excess Lerner index* $r(1-\bar{\mu}/\mu)$ *exceeds the other two terms.*

Corollary 2 rationalizes an observed asymmetry. Jurisdictions hosting the frontier labs are hostile to taxing AI usage; jurisdictions importing AI services are actively legislating digital levies. The model says both are behaving as the formula prescribes: for the importer the market-power term is absent, since the rent it would recover accrues abroad,

and the remaining terms are positive. That is a terms-of-trade argument rather than an environmental or redistributive one, and it should be defended as such—which also means it is a beggar-thy-neighbour policy in the standard sense, and would not survive a cooperative solution.

Remark 1 (Tax base erosion is not an argument). It is often argued that AI must be taxed because it shrinks the labor income tax base. Corollary 1 confirms the premise and Proposition 2 rejects the inference. The mechanical revenue loss from a smaller labor share is not a first-order welfare loss: the government raises τ, at a marginal cost of public funds already reflected in the optimal τ. A tax base is not a policy objective. What matters is who is made worse off, and that is what the term in $\hat{g}_1$ measures. The base erosion argument is a distributional argument wearing a fiscal costume.

5 Reading the three terms

Proposition 2 gives three terms and says nothing about how large they are or why one of them is negative. This section takes them in turn. The organizing idea is instrument assignment: each of the three problems the debate raises has an instrument built for it, and the question for the usage tax is what remains once those instruments are doing their work. Sections 5.1 to 5.3 make the assignment. Sections 5.4 and 5.6 then derive the two results that shrink the terms most—the exposure gap on the redistributive side, and the ownership weight and entry offset on the market-power side—and explain the negative sign that the last of them produces.

Proposition 2 describes the second best when the usage tax is the only instrument. It is not. This section shows that each term in (11) has a better instrument, then develops three extensions that the static formula omits and that change the answer in the same direction.

5.1 Emissions belong on the emissions node

The general principle is Kaplow (2012): when an income tax is available, the corrective instrument should be set at the Pigouvian level on the externality itself and the distributional consequences handled by the tax and transfer system, rather than folded into the corrective rate. What the following proposition adds is that in this setting there are two externality margins, not one, and only one instrument reaches both.

Proposition 3 (Targeting, with an efficiency margin)**.** *Let t_E be a charge per kilowatt-hour at the data center. Setting $t_E = d$ implements the first best on both the scale and the efficiency margins and drives the Pigouvian term in* (11) *to zero. A usage tax set to D/p replicates the first best on the scale margin only. The welfare gap per unit of AI is*

$$\Delta \;=\; \Big[\varphi(\xi^0) + (p_e + d)\, e(\xi^0)\Big] \;-\; \Big[\varphi(\xi^\star) + (p_e + d)\, e(\xi^\star)\Big] \;\geq\; 0,$$

where ξ^0 denotes the efficiency level satisfying the second condition in (3)*, the first-order condition $\varphi'(\xi) = -p_e\, e'(\xi)$, evaluated at the energy price p_e, and $\xi^\star$ the level satisfying that same condition when the perceived energy price is $p_e + d$. The gap is strictly positive whenever any efficiency response is technically possible.*

The efficiency margin is not a curiosity. It is where most of the action has been. Energy per query has fallen by roughly an order of magnitude in three years through better chips, better serving stacks, and smaller distilled models, and the carbon intensity of a query varies by a factor of several across grids. A tax on tokens is neutral with respect to all of it.

Corollary 3 (Location)**.** *With heterogeneous grid intensities ζ_i across sites, a uniform usage tax applies the same wedge to a query served from a clean grid and one served from a coal grid. An emissions charge applies wedges in the ratio ζ_i/ζ_j. The usage tax cannot implement the efficient allocation of load across sites at any positive rate.*

A note on tax form. Everything above concerns the tax *base*—energy versus the value of AI output—and is separate from the question of tax *form*, ad valorem versus specific. The two interact, because under imperfect competition the choice of form affects pass-through and the equilibrium markup. Appendix B.1 shows that the targeting conclusion is unaffected: a specific levy per token is a better instrument than an ad valorem one on several margins, but both are levied on a base that is the wrong physical quantity, and no choice of form repairs a base that mistracks the harm by an order of magnitude.

5.2 Rents belong on the rent

Proposition 4 (Neutrality of the rent tax)**.** *A cash flow tax at rate z on supernormal profits, with immediate expensing and full loss offset, leaves the provider's first order conditions for price, quantity, and efficiency unchanged. It raises $z\Pi$ with zero first-order deadweight loss and strictly dominates any usage tax raising the same revenue whenever $\sigma_q > 0$.*

Proposition 4 does not reduce the markup and so does not remove the negative term in (11). It transfers rent to the treasury without changing the allocation, which is a different question from the one Proposition 2 answers. Proposition 2 asks what wedge belongs on the *transaction*, and finds that genuine rent argues for a lower one; Proposition 4 asks how to collect the rent itself, and finds an instrument that collects it without moving any wedge at all. The two are complements: the rent tax raises revenue from the same rent whose existence made the usage tax look worse.

None of this is new as a matter of tax design, and the paper is stronger for saying so. Cash flow bases with expensing and full loss offset are the standard neutral instrument (Auerbach and Devereux, 2018), and the contribution here is the mapping to this sector: the base is the excess of revenue over expensed cost, the sunk compute stock cannot flee, and the elasticity that makes the usage tax expensive is absent by construction. The known failure mode is political rather than technical. What gets legislated is typically a levy on gross revenue—a digital services tax—which falls on turnover rather than cash flow and so lands on marginal as well as inframarginal units. A revenue tax and a cash flow tax are not variants of one instrument but opposites: one distorts the quantity margin exactly as a usage tax does, the other does not.

One caveat matters in practice: neutrality requires that the base be economic rent and not the normal return.[2]

5.3 Redistribution belongs on the income tax

Corollary 4 (When predistribution is worth nothing)**.** *If the government can implement a nonlinear income tax that separates groups, or can tag on any variable correlated with group membership at low cost, then at the optimum* $\hat{g}_1 \to 0$ *and the predistribution term vanishes. The same conclusion follows without tagging if the income tax is unrestricted, by Scheuer and Werning (2017).*

The predistribution term is a statement about missing instruments, not about technology, and there are three independent routes to zero. Tag, or use a nonlinear schedule, and $\hat{g}_1$ collapses. Or, without tagging, Scheuer and Werning (2016) show that in a Roy economy the general-equilibrium wage channel is not a lever a fully optimizing income tax wants to pull, because the schedule already prices every earnings level at which the wage change lands. Or wait: Guerreiro et al. (2022), in a model with routine and non-routine

[2]A tax on gross revenue, which is what most digital services taxes actually are, is a usage tax in disguise and inherits every one of its distortions. The distinction between a revenue tax and a cash flow tax is the distinction between Proposition 2 and Proposition 4, and it is not a detail.

occupational choice close to ours, find a positive robot tax survives only while the cohort that cannot retrain is still working. Any AI tax justified on this basis is transitional by construction and should be legislated with a sunset.

5.4 The predistribution term: mobility and identification are one parameter

Section 3 put the sorting margin in the model rather than bolting it on, and this subsection collects what that buys. The question it answers is the one the robot-tax literature usually settles by assumption: how much of the predistribution term survives when workers can move?

The intuitive answer is that mobility damps the term. A fall in the routine wage induces exit from routine tasks, which raises the routine wage back toward where it started; in the limit of frictionless reallocation, relative task wages cannot be moved at all by an input tax and the term is zero. That argument is correct about *task wages* and it is the reason $\mathcal{E}$ appears in the denominator of (11). It is incomplete as a statement about welfare, and the incompleteness reverses a comparative static.

Proposition 5 (Mobility and identification are jointly constrained)**.** *Under the Roy structure of* (7)*, the predistribution term is*

$$\frac{\Omega}{p} \;=\; \frac{(1-\tau)\,\hat{g}_1\,\Delta_\ell\,(\sigma-1)(1-s)}{\big(1+\varepsilon_L+\kappa(1-\Delta_\ell^2)\big)\,\sigma_q}, \qquad \Delta_\ell \;=\; \tanh\!\big(\tfrac{\kappa\varrho}{4}\big).$$

Then, holding the comparative-advantage gap ϱ fixed:

(i) the reallocation elasticity $\mathcal{E} = \kappa(1-\Delta_\ell^2)$ is hump-shaped in κ, with $\mathcal{E} \to 0$ as $\kappa \to 0$ and as $\kappa \to \infty$;

(ii) the exposure gap Δ_ℓ is strictly increasing in κ, from 0 to 1;

(iii) consequently Ω/p is strictly increasing *in κ, from zero at $\kappa = 0$ to the fixed-assignment value at $\kappa \to \infty$, and this holds for every $\varrho > 0$ without further restriction. The two-type model of the robot-tax literature is the $\kappa \to \infty$ limit, not the $\kappa = 0$ limit;*

(iv) for any fixed measured exposure gap $\Delta_\ell < 1$, the term is bounded above by Δ_ℓ times its fixed-assignment value, with equality only if $\kappa = 0$, which is inconsistent with $\Delta_\ell > 0$.

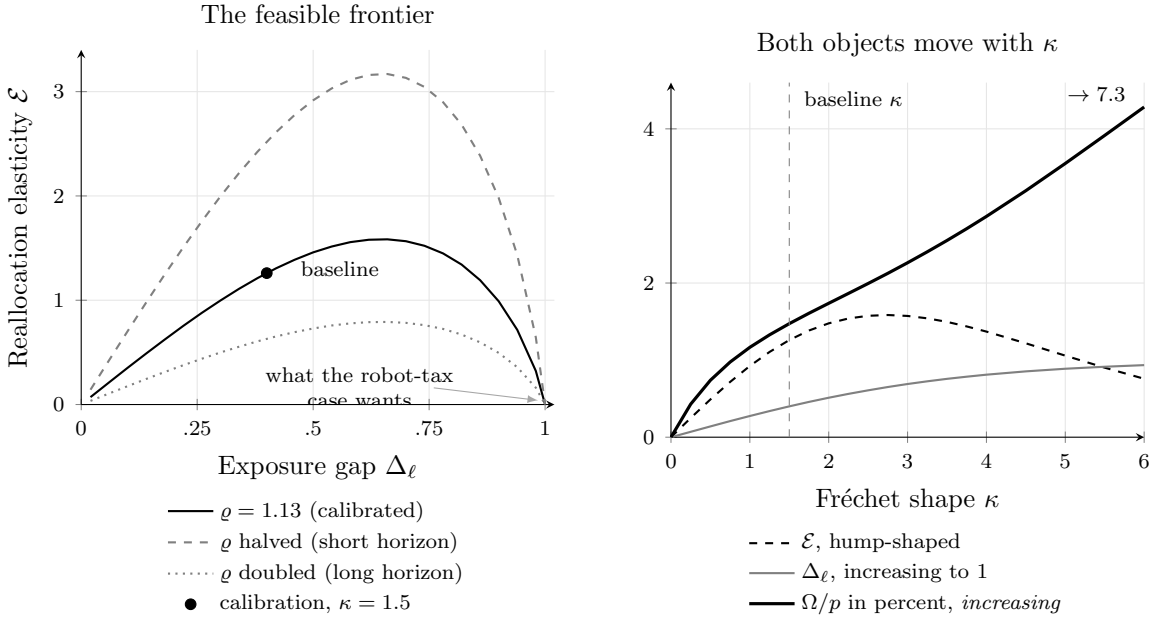


Figure 1: Proposition 5 drawn. Left: eliminating κ between $\Delta_\ell = \tanh(\kappa\varrho/4)$ and $\mathcal{E} = \kappa(1-\Delta_\ell^2)$ leaves the frontier $\mathcal{E} = (4/\varrho)(1-\Delta_\ell^2)\,\mathrm{artanh}(\Delta_\ell)$, which is single-peaked and vanishes at both ends. Everything above a curve is unavailable at that comparative-advantage gap. The robot-tax case implicitly wants a sharply identified group of losers together with slow reallocation, so that the wage effect is not competed away. That is the top-right corner, where the frontier collapses to zero. Widening ϱ shifts the whole frontier down, so the two objects cannot be improved together by appeal to the horizon either. Right: why part (iii) inverts the usual comparative static. Raising κ speeds reallocation, but it also sharpens sorting, and the second effect dominates. $\mathcal{E}$ is hump-shaped and eventually falls, Δ_ℓ rises monotonically toward one, and Ω/p therefore *rises* in κ throughout, approaching the fixed-assignment value of 7.3 percent instead of falling to zero. A reader who raises κ to damp the term is moving the wrong way. Drawn at $\varrho = 1.13$, the value for which the baseline $\kappa = 1.5$ delivers the measured $\Delta_\ell = 0.4$; other parameters as in Table 4.

The intuition. Think of the Fréchet shape parameter as answering a single question: how different are workers from one another *within* a group? When dispersion is wide, everyone is a coin flip between tasks, the two groups look alike, and there is no identifiable loser for the planner to protect—but there is a great deal of switching, because half the workforce sits near the margin. When dispersion is narrow, comparative advantage is nearly deterministic: group 1 works routine tasks and group 2 does not, so the loser is perfectly identified—and nobody switches, because nobody is near the margin. The same parameter that makes workers mobile is the parameter that makes them indistinguishable. It cannot be turned up to damp the wage effect while the welfare weight on "routine workers" is held fixed, because turning it up is what creates a routine worker in the first place.

Parts (i) to (iii) invert the comparative static usually reported in this literature. Treat-

ing the Fréchet shape as a mobility parameter, so that raising it damps the predistribution term while the welfare weight on "routine workers" stays fixed, is not a valid exercise: it varies the reallocation response while holding constant an object—the identity of the losing group—that the same parameter determines. Once both move together, the sign flips. Perfect sorting is perfect segregation. A world in which comparative advantage is sharp enough for the exposed group to be cleanly identifiable is also a world in which nobody is near the switching margin, and the fixed-assignment model is recovered exactly.

Part (iv) is where the economics is. The right way to discipline the term is not to argue about κ but to measure the two objects the formula actually contains. The exposure gap Δ_ℓ is the difference between two employment shares, the fraction of the disadvantaged group working in AI-exposed tasks, less the fraction of the advantaged group doing so, and it is estimable from any occupational exposure measure crossed with any ex ante grouping the planner might have in mind. The reallocation elasticity $\mathcal{E}$ is then pinned down by $\mathcal{E} = \kappa(1 - \Delta_\ell^2)$ given a switching estimate. The literature's convention $\Delta_\ell = 1$ is not a conservative simplification; it is the corner of the parameter space that maximizes the term, and it is empirically false, since exposure to AI is not remotely segregated by education, income, or any other candidate for an ex ante group. Section 6 measures it, carries $\Delta_\ell \approx 0.4$—the value on the exposure measure most favourable to the robot-tax case, with every direct AI measure giving a negative number—and finds the term falls to about a fifth of the fixed-assignment number. The reduction comes chiefly from the exposure gap rather than from mobility.

Two implications for policy survive intact from the earlier treatment, and one does not. What survives: the term is transitional, because ϱ itself falls as cohorts choose skills knowing what AI can do, which is the parameter behind the sunset that Corollary 4 calls for; and the correction runs in the same direction as every other extension in this section, lowering the optimal usage tax. What does not survive is the claim that a reader who believes in fast reallocation should read the predistribution column as zero. The right statement is that a reader who believes the losing group is sharply identifiable should also believe reallocation is slow, and vice versa, and that neither belief rescues the term, because the product is bounded by the exposure gap either way. Figure 1 draws both halves of the proposition; Appendix B.7 develops the frontier, the horizon dependence of ϱ, and two caveats that cut in opposite directions.

5.5 Endogenous entry: the usage tax amplifies market power

Taxation with an endogenous entry margin is studied by Scheuer (2014) in the context of entrepreneurship, where the same tension appears: an instrument that looks like it falls on rents partly falls on the decision to be in the market at all.

The markup was taken as given above. It is not. Provision has large fixed costs in research and in compute clusters, so the number of providers is determined by free entry, and a tax that shrinks the market shrinks the number of firms.

Let n symmetric providers compete in quantities facing the demand elasticity σ_q of (10), so the equilibrium markup is $\mu = \big(1 - 1/(n\sigma_q)\big)^{-1}$, and let each incur a fixed cost F. Free entry drives per-firm profit to zero.

The result below is the one that most surprises people, so the mechanism is worth stating before the statement. A tax on usage raises the price buyers pay and shrinks the market. A smaller market supports fewer firms, because each must still cover the same fixed cost. Fewer firms compete less hard. So the tax, which was meant to fall on incumbents, ends by handing the survivors a wider margin on the sales they retain. The tax is a barrier to entry that the incumbents did not have to build.

Proposition 6 (Entry amplification)**.** *With Cournot competition and free entry,*

$$\frac{\mathrm{d}\ln\mu}{\mathrm{d}\ln(1+t_U)} \;=\; \tfrac{1}{2}\,(\mu-1)\,\sigma_q \;>\; 0,$$

so a usage tax raises the equilibrium markup. The buyer price responds by

$$\frac{\mathrm{d}\ln q}{\mathrm{d}\ln(1+t_U)} \;=\; \Phi \;\equiv\; 1 + \tfrac{1}{2}(\mu-1)\sigma_q \;>\; 1,$$

and the optimal usage tax is strictly below the level in Proposition 2.

The amplification is large, and Figure 2 shows how it scales. At $\mu = 2.5$ and $\sigma_q = 2$, $\Phi = 2.5$: a usage tax of one percent raises the buyer price of AI by two and a half percent, because it also drives out entrants and lets the survivors raise their margins. A usage tax on a concentrated sector with heavy fixed costs is not a tax on rent but a barrier to entry erected on the incumbents' behalf, part of which the incumbents capture.[3]

[3]The Cournot structure is doing work here: with contestable markets or a competitive fringe $\Phi \to 1$. What survives is the direction: any model in which fixed costs are large and entry responds to market size delivers $\Phi > 1$, so the static formula understates the distortion. The other caveat (that free entry can be socially excessive) is not a caveat at all but a force that changes the sign of a whole term, and §5.6 takes it up properly.

Training compute. Inference and training are different activities with different externalities, and a levy on tokens taxes almost exclusively the former. If frontier training runs generate social risk convex in the compute used, the corrective instrument is a Pigouvian levy on training compute, which dominates the compute thresholds already written into regulation for the usual reason that a threshold is a zero marginal rate followed by an infinite one. Separately, training burns electricity but produces no tokens, so its implicit rate under a token tax is not low but exactly zero, while a charge at the meter reaches it automatically. Appendix B.4 states both results.

The gap grows over time. Proposition 3 compares instruments at a fixed efficiency technology, and in doing so understates its own conclusion. Efficiency is the outcome of durable investment, and only an instrument that raises the perceived price of energy enters the investment condition at all. Appendix B.5 extends the model to a horizon T and shows that the emissions charge induces strictly more efficiency investment at every horizon, that the gap widens as assets live longer, and that the cumulative welfare gap is increasing and convex in T. The practical implication is worth stating here: the common defence of the token tax as a simple approximation to be tightened later has the sequencing backwards, because the approximation deteriorates with time rather than improving.

5.6 The market-power term, and why it is negative

The market-power term is the largest object in the formula and the only one that is negative, so it governs the sign of everything reported later. It is also the term hardest to accept at face value, since it appears to say that a planner should subsidize the most profitable firms in the economy. Three things need saying, and the third is a constraint the calibration must satisfy.

First: the sign is not a claim about AI. A good sold above marginal cost is consumed too little, and the textbook corrective for underconsumption is a lower price. Equation (11) inherits that from Cournot (1838) and would return the same sign for bread. The formula prices underprovision on the only margin it has been given, because Proposition 2 solves for a single wedge. It does not know that the planner also has a cash flow tax, an antitrust authority, and a procurement policy. Section 5.2 and Section 5.5 give the planner those instruments, and once they are assigned the term leaves the usage tax's job description entirely. Nothing in this paper recommends paying frontier labs; Appendix

C.1 says so at length. What the term does is measure the size of one distortion relative to another, and that comparison is informative whatever one thinks of the instrument.

Second: free entry offsets it, and possibly by a lot. The term as written treats a reduction in output as a pure loss. With free entry and large fixed costs that is not right. Each entrant duplicates a fixed cost (at the frontier, a training run costing hundreds of millions) and captures business from incumbents that would have been served anyway. Mankiw and Whinston (1986) show that with homogeneous products this business-stealing effect makes the free-entry number of firms exceed the constrained optimum, so an output subsidy that induces entry destroys resources at the margin even as it corrects the price.

Proposition 7 (Business stealing offsets the market-power term)**.** *In the symmetric Cournot free-entry equilibrium of Proposition 6, the welfare effect of a marginal reduction in the usage wedge is a fraction* $(1-\chi)$ *of the naive allocative gain* $m\,dQ$*, where* $m \equiv p - mc$ *and*

$$\chi \;=\; 1 \;+\; \left.\frac{\mathrm{d}\ln m}{\mathrm{d}\ln Q}\right|_{\textit{free-entry locus}},$$

the derivative being taken along the locus $n = Q\sqrt{-P'(Q)/F}$*,* $m = \sqrt{-P'(Q)\,F}$*. The market-power term in* (12) *is therefore scaled by* $(1-\chi)$*. Two demand systems bracket it:*

(i) Linear demand*,* $P = a - bQ$*: the equilibrium margin is* $m = \sqrt{bF}$*, independent of the subsidy, so* $\chi = 1$ *exactly. The allocative gain is fully dissipated in the fixed costs of induced entry, and a marginal output subsidy has* zero *first-order welfare effect.*

(ii) Constant-elasticity demand *at the elasticity* σ_q *of* (10)*, which is the demand system this paper's CES structure actually generates:*

$$\chi \;=\; 1 - \frac{\sigma_q + 1}{2\sigma_q} \;\in\; \left[0, \tfrac{1}{2}\right),$$

increasing in σ_q*, equal to zero at* $\sigma_q = 1$ *and approaching one half as* $\sigma_q \to \infty$*.*

The intuition. A subsidy to output looks like a pure gain: each extra unit is worth p to a buyer and costs mc to make, so society banks the difference. But with free entry the extra output does not come from existing firms working harder. It comes partly from a new firm, and the new firm must first pay the fixed cost—at the frontier, a training run. The proof shows that this is not a partial offset but an exact one: along the free-entry locus the gain $m\,dQ$ and the entry cost $F\,dn$ cancel *term for term*, and all that survives

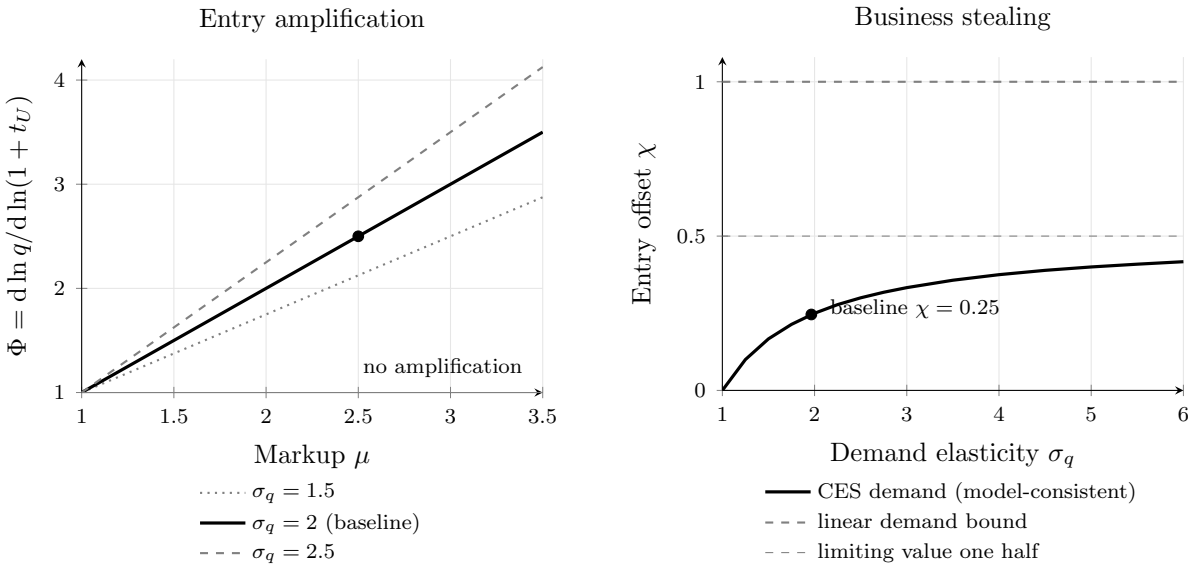


Figure 2: Two forces that the single-wedge formula omits, and that pull the optimal usage tax in opposite directions. Left, Proposition 6: with Cournot competition and free entry a usage tax shrinks the market, the market supports fewer firms, and the survivors widen their margins, so the buyer price moves by $\Phi = 1 + \frac{1}{2}(\mu - 1)\sigma_q > 1$ per unit of tax. At the baseline $\mu = 2.5$, $\sigma_q = 2$ the marked point gives $\Phi = 2.5$: a one percent tax raises the price of AI by two and a half percent. The amplification is increasing in the markup, so it is largest exactly where the tax is most often proposed. Right, Proposition 7: the offsetting force. Each entrant induced by a lower wedge duplicates a fixed cost, so a fraction χ of the allocative gain is dissipated rather than banked. Under linear demand the offset is complete and a marginal output subsidy is worthless; under the CES demand this paper's structure generates, χ rises from zero at unit elasticity toward one half, and the model-consistent value at $\sigma_q = 1.965$ is 0.25. The paper's own demand system therefore puts the offset near the *bottom* of the bracketed range, which is worth noting because it means χ is derived and not chosen to shrink the term.

is the second-order effect of the subsidy on the equilibrium margin itself. Whether that residual is large depends entirely on how the margin moves as the market grows, which is a property of the demand curve and nothing else. Under linear demand the margin does not move at all, so the offset is complete and the subsidy is worthless. Under constant-elasticity demand the margin falls as output expands, which is a genuine gain to inframarginal buyers, and the offset is partial.

The proof is in Appendix B.6. The point worth emphasizing is that χ is *not* a free parameter here. The paper's own demand structure pins it: at the baseline $\sigma_q = 1.965$, part (ii) gives $\chi = 0.25$, so a quarter of the underprovision correction is dissipated in duplicated training runs and the market-power term is three quarters of its textbook size. The value at $\sigma_q = 1$ is a useful check, unit-elastic demand holds revenue constant, so free entry induces no entry response and there is nothing to dissipate, and the linear case is the natural upper benchmark, in which the offset is complete. We use $\chi = 0.25$

throughout and report $\chi = 1$ as the linear-demand bound, which is the honest range: the model-consistent value is at the bottom of it.

Third: whose dollar is the subsidy, and what is it worth? The reason a planner would want a lower price of AI is that the last unit is worth p to a buyer and costs mc to make, so the unexploited gain from trade is $m = p - \text{mc}$. But at the margin the buyer pays exactly what the unit is worth to her, so her surplus on it is zero: the *entire* marginal gain m accrues to the provider. The market-power term is therefore not a consumer-surplus correction at all. It is a statement that a dollar of foregone provider profit is a dollar of lost social surplus, and whether that is true depends on two things the formula has so far assumed away.

The first is who owns the provider, which is why f appears. Surplus that would have accrued to foreign shareholders is not part of domestic welfare, so at $f = 1$ there is nothing at the margin for a domestic planner to recover and the term vanishes, not because underprovision has stopped, but because its cost is borne abroad. The second is what a dollar to a domestic shareholder is worth to the planner, and here the paper had been silently inconsistent. Frontier AI equity is held overwhelmingly at the top of the wealth distribution. A planner who values a dollar to a routine worker at 1.3 times the average, which is what $\hat{g}_1 = 0.3$ asserts, cannot coherently value a dollar to a frontier-lab shareholder at exactly the average. Writing g_S for that weight,

Proposition 8 (The market-power term is a distributionally weighted transfer)**.** *With foreign ownership share f, a welfare weight g_S on domestic AI shareholders relative to the income-weighted average, and the entry offset χ of Proposition 7, the market-power term in* (12) *is*

$$-(1-f)\, g_S\, r\Big(1 - \frac{\bar{\mu}}{\mu}\Big)(1-\chi).$$

It vanishes if providers are foreign-owned ($f = 1$), if the planner attaches no value to shareholder income ($g_S = 0$), if the markup is pure fixed-cost recovery ($r = 0$), or if entry fully dissipates the correction ($\chi = 1$). The unweighted formula is the special case $g_S = 1$, $\chi = 0$.

This resolves what we take to be the real source of discomfort with the negative sign. Read at $g_S = 1$ the formula says a planner should hand money to the most concentrated firms in the economy and count the transfer as costless. That is not a proposition anyone holds; it is an artifact of a utilitarian aggregation applied to workers and not to shareholders. g_S is a normative parameter, but it is not free: it is disciplined by consistency

with $\hat{g}_1$, and Lemma 4 performs that calibration rather than asserting a value.

Lemma 4 (The shareholder weight implied by $\hat{g}_1$)**.** *Let welfare weights be isoelastic in income, $g(y) \propto y^{-\eta}$, normalized so that the income-weighted mean weight is one, and let log income be normal with dispersion s. Then $g(y) = \exp\big(-\eta s z + \tfrac{1}{2}\eta(2-\eta)s^2\big)$, where z is the income z-score. The requirement $g = 1.3$ at the routine worker's position pins η; evaluating g at the equity-weighted holder's position then delivers g_S.*

The calibration is disciplined by two observable positions. Routine occupations sit near the 36th percentile of the earnings distribution, $z_R \approx -0.35$, and the dispersion of log earnings is $s \approx 0.8$; these give $\eta \approx 0.32$. Corporate equity is held overwhelmingly at the top—the top one percent hold roughly half of it—so the equity-weighted dollar sits near $z_S \approx 2.3$, which gives

$$g_S \;\approx\; 0.65, \qquad \text{range } 0.45 \text{ to } 0.85$$

across $s \in [0.7, 1.0]$, $z_R \in [-0.5, -0.25]$ and $z_S \in [2.0, 3.0]$.

The calibration is a constraint rather than a licence. A planner who values a routine worker only thirty percent above the average is not strongly redistributive, and the same modest curvature that keeps the predistribution term small also keeps shareholders from being heavily discounted: $\hat{g}_1 = 0.3$ and $g_S = 0.2$ are not jointly available under any isoelastic welfare function. Combined with the derived $\chi = 0.25$, the scaling on the textbook term is $g_S(1-\chi) \approx 0.49$, so the market-power term is *about half* of what (11) reports at face value.

It also answers the objection that the deadweight loss from market power might simply be smaller than that from a tax, so the comparison should run the other way. That compares levels; the formula is a marginal condition. What matters is not the size of the Harberger triangle under the monopoly wedge but that a wedge already exists where the tax is levied: a tax imposed on top of a pre-existing wedge has a *first-order* welfare cost rather than the second-order cost it would have in an undistorted market. That is why the term enters linearly in m rather than quadratically, and the sense in which the two distortions compound rather than compete.

Fourth: r is not free, and it is the same parameter that governs the rent tax. Section B.2 already pins the rent share to the recurring-investment share through $r = 1 - \phi\,\mu/(\mu - \bar{\mu})$, so that $\phi = 0.2, 0.3, 0.4$ imply $r = 0.6, 0.4, 0.2$. What that section

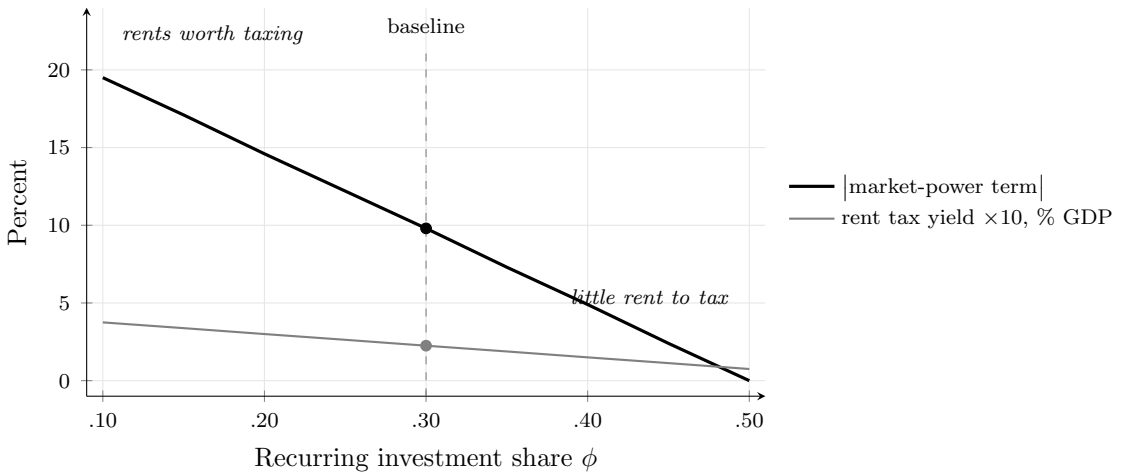


Figure 3: Corollary 5: the two live positive proposals are governed by one number. Recurring training and capacity investment ϕ fixes the pure-rent share of the markup through $r = 1 - \phi\mu/(\mu - \bar{\mu})$, and therefore fixes both curves at once. Move left and rent is abundant, so a cash flow tax collects well—but the same abundance makes the efficient wedge on usage strongly negative, because more of the price a buyer pays is a transfer rather than a resource cost. Move right and the negative wedge shrinks toward zero, but so does the rent tax base, until at $\phi = 0.5$ the markup is entirely fixed-cost recovery and neither instrument has anything to reach. No value of ϕ makes AI rents large enough to be worth taxing while leaving the market-power correction small enough to ignore. They are the same number read from opposite ends. The rent-tax series is a fifty percent cash flow tax on the stationary base $[(1 - 1/\mu) - \phi]\theta$, scaled by ten so that both series fit one axis; the market-power series is $(1 - f)g_S r(1 - \bar{\mu}/\mu)(1 - \chi)$ at $f = 0$, $g_S = 0.65$, $\chi = 0.25$, $\mu = 2.5$, $\bar{\mu} = 1.25$, $\theta = 1.5\%$.

does not do, and what resolves most of the puzzlement the negative term provokes, is trace the same ϕ through to the *other* instrument.

Corollary 5 (The rent tax yield and the market-power term move together)**.** *The stationary cash flow base of Section B.2 is* $[(1 - 1/\mu) - \phi]\theta$ *and the market-power term is* $-(1 - f)r(1 - \bar{\mu}/\mu)(1 - \chi)$ *with* $r = 1 - \phi\mu/(\mu - \bar{\mu})$*. Both are strictly decreasing in* ϕ*. Hence the revenue a rent tax raises and the magnitude of the negative wedge on usage are governed by one parameter and cannot be chosen separately: an economy in which the rent tax is lucrative is exactly an economy in which the efficient wedge on AI usage is strongly negative, and an economy in which the market-power term is negligible is one in which the rent tax collects little beyond the normal markup.*

Table 3 traces the trade-off and converts the sign question into something checkable against a single accounting fact. Applying the textbook correction unweighted, the efficient wedge on AI usage is non-negative for a domestic economy only if recurring investment exceeds about 46 percent of revenue—at which point a fifty percent cash flow tax raises about 0.11 percent of GDP, half the headline figure of Table 5. Weighting the term at the g_S of Lemma 4 and netting the entry dissipation of Proposition 7 moves that

threshold only to about 41 percent, so the conclusion does not turn on the weighting: on either reading the wedge is negative at every ϕ for which a rent tax collects anything worth collecting. *One cannot simultaneously hold that AI rents are large enough to be worth taxing and that the market-power correction is small enough to ignore. They are the same number read from opposite ends.*

Weighting shareholder income at the $g_S = 0.65$ of Lemma 4 and netting the derived dissipation $\chi = 0.25$ scales the term to about half of that, which moves the threshold to roughly 41 percent of revenue and leaves the baseline wedge at -5.2 rather than -15.5 percent; under the linear-demand bound the term vanishes and the wedge is $+4.5$ percent at every ϕ. The trade-off is therefore about magnitude within a narrow band rather than about sign. We regard this as the most useful form of the result, because ϕ is an accounting quantity a disclosure requirement could produce whereas r is a construct—and it is one more argument for the instrument unaffected by any of it: the emissions charge does not care what fraction of the markup is rent, who owns it, or how entry responds.

5.7 Diffusion spillovers, and the general formula

If a marginal unit of AI use generates learning or diffusion benefits the buyer does not capture, the Pigouvian term is netted of them and the optimum falls further. The sign of these spillovers is unsettled in the literature and they enter against a corrective term worth only a few percent, so we carry them as a free parameter and note that every plausible value moves the optimum in the same direction as everything else in this section: down. Appendix B.9 develops the point. Collecting the results of this section, the general formula is

$$t_U^\star \;=\; \underbrace{\frac{(1-\tau)\hat{g}_1\Delta_\ell(\sigma-1)(1-s)}{(1+\varepsilon_L+\mathcal{E})\sigma_q}}_{\text{predistribution}} \;+\; \underbrace{\frac{D-\psi}{p}}_{\text{net externality}} \;-\; \underbrace{(1-f)\,g_S\,r\left(1-\frac{\bar{\mu}(\cdot)}{\mu(t_U)}\right)(1-\chi)}_{\text{market power, weighted and net of entry}}, \tag{12}$$

with $\mu(t_U)$ increasing at the rate given in Proposition 6, with g_S and χ as in Propositions 8 and 7, and with the training margin handled by a separate instrument $t_K^\star = R'(K)$. Proposition 2 is the special case $g_S = 1$, $\chi = 0$, $\Delta_\ell = 1$, $\mathcal{E} = 0$: welfare weights that do not distinguish shareholders from anyone else, entry that is socially costless, and workers welded to their tasks. Each of the four is an assumption the literature makes silently, and relaxing them is what separates (12) from (11). Note that (12) is an implicit equation rather than a formula: the markup on the right-hand side is itself a function of the tax

on the left, so the optimum is the fixed point at which μ is evaluated at $t_U^\star$. We solve it numerically in the calibration; at our parameters the fixed point is unique and the iteration converges immediately, because the market-power term depends on μ only through the ratio $\bar{\mu}/\mu$ and is therefore flat in the relevant range. Every extension moves (12) in the same direction, which is down.

6 Calibration

The formula is only as useful as the numbers in it. This section calibrates each term, then runs five back of the envelope calculations through the model's own lemmas, then reads the policy debate against the results. Parameters are chosen, where there is doubt, to favor the case for taxing AI.

6.1 The three terms

The residual external cost per kilowatt-hour. Section 2 gave the two components. Unpriced carbon at \$190 per ton and $\zeta = 0.4$ kg per kWh is \$0.076 per kWh where no carbon price applies; the grid-shifting wedge γ we carry over $[0, 0.03]$; water and local damages add a few tenths of a cent. A central residual, with a modest carbon price already in place and γ near the middle of its range, is

$$d \;\approx\; \$0.05 \text{ per kWh}, \qquad \text{range } \$0.02 \text{ to } \$0.10,$$

and higher with no carbon price. The residual wedge is therefore between roughly sixty percent and more than one hundred percent of the price of the input on which it should be levied—an extraordinary number by the standards of environmental taxation. Figure D.2 shows it is robust across the whole range of γ, because carbon dominates. This separates the two roles of γ: it barely moves the *level* of the wedge, and what it governs instead is the distributional character of the status quo, which is Section 6.5.

The residual external cost per unit of AI, across workloads. Here the point estimate is misleading and the distribution is the object that matters. A median short text query consumes about 0.3 Wh, but reasoning and agentic workloads run one to two orders of magnitude higher (International Energy Agency, 2026; Luccioni et al., 2024; de Vries, 2023), and these are where load growth is concentrated. Because both energy and price scale with query length, D/p is roughly constant *within* a workload type but

Table 1: The externality ratio D/p varies across workloads

Workload	Energy/query	Price/query	D/p	Share of queries	Share of energy
Simple text	0.3 Wh	\$0.003	0.5%	0.60	0.02
Standard chat	3 Wh	\$0.010	1.5%	0.18	0.05
Reasoning	30 Wh	\$0.050	3.0%	0.19	0.52
Agentic	150 Wh	\$0.150	5.0%	0.03	0.41
Count-weighted mean			1.3%		
Energy-weighted mean			**3.7**%		

Notes. Energy and price per query are order-of-magnitude figures drawn from International Energy Agency (2026), Luccioni et al. (2024) and de Vries (2023); D/p evaluates $d \cdot e/p$ at $d = \$0.05$ per kWh. The query mix is illustrative rather than measured: no public source reports the composition of inference by workload type, so we impose a plausible mix in which short queries dominate by count while reasoning and agentic workloads dominate by energy. The share of energy is derived from the query shares and energy per query, not assumed separately. The two means are the D/p column weighted by the two share columns respectively. Nothing in the paper's argument depends on the particular mix, only on the fact that the ratio of harm to value rises across the rows.

differs *across* types through the price per joule of delivered value. Table 1 works the ratio for four representative workloads at $d = \$0.05$ per kWh.

The two means differ by a factor of three, and the energy-weighted mean is the policy-relevant one, because a corrective instrument should track the harm, which is proportional to energy. A tax calibrated to the modal query, the naive 1.3%, understates the externality of the marginal kilowatt-hour by nearly threefold, since the energy-hungry tail dominates the total footprint. The token tax therefore not only sits on the wrong base, but on one whose relationship to the harm is itself unstable across the workloads driving the load growth. We carry the energy-weighted central value $D/p \approx 2$–4% forward, and note it would be higher still with no carbon price.

Market power and the rent share. Reported gross margins on frontier inference of fifty to sixty percent give $\mu \approx 2.5$ against $\bar{\mu} \approx 1.25$ (De Loecker et al., 2020), so the excess Lerner index $1 - \bar{\mu}/\mu \approx 0.5$. But the term is scaled by r, and by Corollary 5 r is the complement of the recurring-investment share: at $\phi = 0.3$, $r = 0.4$, so the naive use of the gross markup as if it were all rent overstates the term by more than a factor of two. Scaling further by the $g_S = 0.65$ of Lemma 4 and $1 - \chi = 0.75$ gives an effective market-power term of about ten percent rather than fifty. We report ϕ over $[0.2, 0.4]$ rather than a point value, because it is not directly observed and it is the parameter the sign turns on.

Predistribution. The formula has three inputs and only one of them is normative. Take $\tau = 0.35$, $\varepsilon_L = 0.3$, $\sigma = 2$, and a welfare weight $\hat{g}_1 = 0.3$, so that the planner values a dollar to the exposed group thirty percent above the average. At the fixed-assignment corner $\Delta_\ell = 1$, $\mathcal{E} = 0$ (the configuration the robot-tax literature works in) the exact term of (11), with $(1-s)$ in the numerator and $\sigma_q = \sigma(1-s)+s+\theta \approx 1.965$ in the denominator, gives $\Omega/p \approx 7.3\%$. We report that number only as the corner it is.

The exposure gap is not one. Δ_ℓ is a difference of two employment shares—the share of the disadvantaged ex ante group working in AI-substitutable occupations, less the share of the advantaged group—so it requires an exposure measure, a grouping, and a threshold, and given those three it is a descriptive statistic computable from one year of the American Community Survey. It is not a correlation, not a regression coefficient, and not bounded below by zero. Appendix D.7 states the estimand formally and the replication package includes the code.

We compute it. Table 2 reports Δ_ℓ for seven public occupation-level exposure measures crossed with four ex ante groupings, on an employment-weighted panel of 765 detailed SOC occupations; Appendix D.7 gives the construction and its limitations. Three findings come out of the table.

First, the exposure gap is nowhere close to one. The largest absolute value in the table is 0.50 and the largest positive value is 0.45; the left panel of Figure 4 plots these against the value the literature assumes. The predistribution term is proportional to Δ_ℓ and is damped further through $\mathcal{E} = \kappa(1-\Delta_\ell^2)$, so the convention $\Delta_\ell = 1$ overstates it by a factor of about five at the most favourable measure in the table, and by more at every other one. The right panel shows why the overstatement is so large. The map from Δ_ℓ to Ω/p is convex, since the damping through $\mathcal{E}$ relaxes as the gap widens, and almost all of the term is generated in the stretch between the measured values and the assumed corner.

Second, the measures disagree about the sign, and they disagree by family. The robot-exposure measure of Webb (2020) is the closest available proxy for the routine displacement that σ describes, and it gives $\Delta_\ell = +0.41$ for a less-than-BA group 1, with $+0.34$ to $+0.45$ across the other groupings. The software measure of the same paper gives a uniformly smaller $+0.18$ to $+0.23$. On all four direct AI and generative-AI measures, Felten et al. (2021), Eloundou et al. (2024) under both the human and the GPT-4 rubric, and Eisfeldt et al. (2023), the gap is negative for every grouping, between -0.18 and -0.50. The employment-weighted correlation between exposure and the share of an occupation's incumbents without a bachelor's degree is $+0.52$ on the robot measure and -0.25 to -0.64 on the four AI measures. The choice of grouping does not move the sign. The two families

of measure describe different technologies.

Evidence of a different kind supports the negative side of the split. Tomlinson et al. (2025) do not rate occupations for exposure. They classify two hundred thousand logged conversations with a deployed assistant into the work activities it was asked to perform, then aggregate to occupations. Their applicability score correlates with occupational wages at +0.13 when occupations are weighted by employment, and it spreads broadly across education requirements. This is what our negative Δ_ℓ also says, obtained without any expert judgment about what a model could do. We weight it more heavily than we would agreement among rating-based measures, which share a method and could share its errors.

Third, the gap does not survive a tighter definition of exposure. Figure 5 sweeps the threshold from the fortieth to the ninetieth percentile of employment. Six of the seven measures collapse toward zero, and at the ninetieth percentile all seven lie between -0.22 and $+0.10$. The exception is Webb software, which stays flat inside a band around $+0.1$ at every cut. This is the signature of an exposure distribution that is tilted across the education distribution but not segregated by it. A measure capturing genuine segregation would hold its gap open at a tight cut, since the occupations surviving such a cut would be held overwhelmingly by one group. None of the seven does.

We carry $\Delta_\ell = 0.4$ as the central case. This is the value on the single measure most favourable to the robot-tax case, and we carry it for that reason, following this section's practice of resolving doubt in favour of taxing AI. The range the measurement supports is $[-0.50, +0.45]$, which contains zero.

Remark 2 (The sign of Δ_ℓ is not guaranteed). Every treatment of the robot tax we know of imposes $\Delta_\ell = 1$ and therefore cannot express the possibility that the exposure gap is negative. Nothing in the theory rules it out, and Table 2 finds it. On all four direct AI and generative-AI measures, and on all four groupings, Δ_ℓ lies between -0.18 and -0.50. When $\Delta_\ell < 0$ the predistribution term in (11) changes sign and joins the market-power term. A usage tax then *worsens* the distribution across ex ante groups, because it protects the tasks performed disproportionately by the better-off, and the redistributive argument turns into an argument against the instrument. At the measured LLM gap of -0.37 the term is -1.3 percent against the $+7.3$ percent of the fixed-assignment corner, a swing of nearly nine points on the same formula. Which sign applies depends on what σ describes. If it is routine displacement then the robot measure is the right one and $\Delta_\ell \approx +0.4$; if it is direct LLM substitution then it is not. We do not settle that here, and the frontier is moving in any case. What matters is that the question is empirical and has a measured

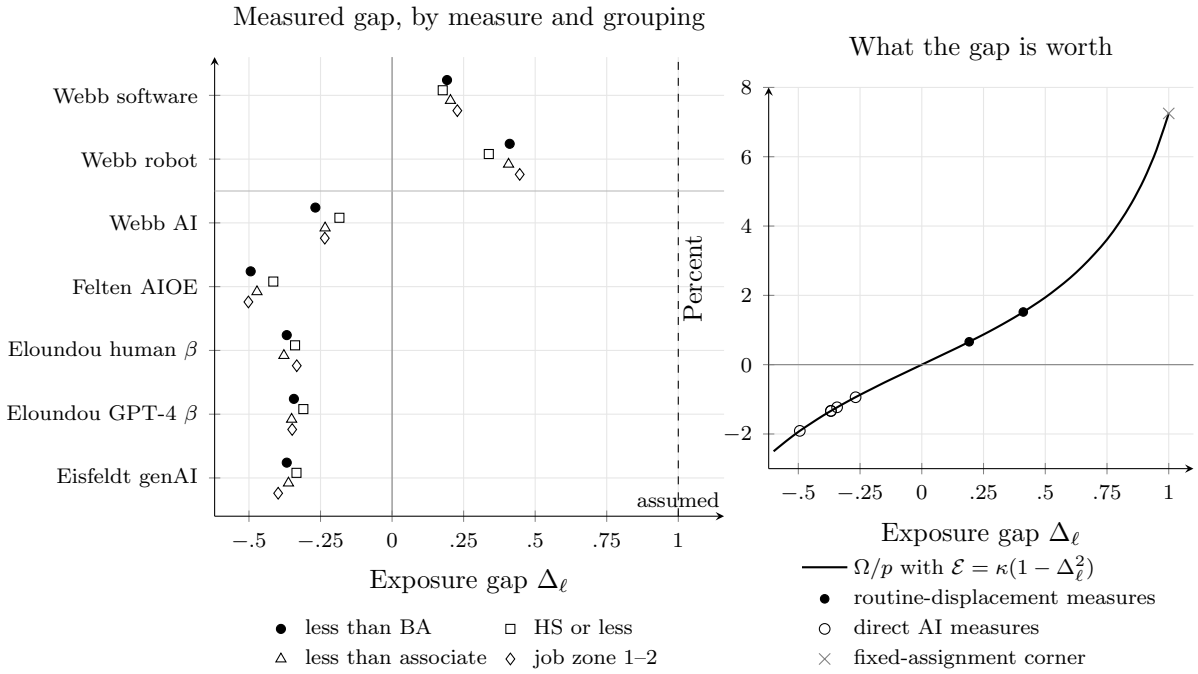


Figure 4: The exposure gap, measured, and what it is worth. Left: Δ_ℓ for seven public occupation-level exposure measures crossed with four ex ante groupings, from Table 2; the horizontal rule separates the two routine-displacement measures above from the four direct AI measures below, and the dashed vertical line is the value $\Delta_\ell = 1$ that the robot-tax literature imposes by assumption. One of the twenty-eight cells reaches half that value in absolute terms and none is positive beyond 0.45. The four markers move together within every row, so the sign is a property of the measure and not of the grouping. Right: the predistribution term Ω/p of (11) as a function of Δ_ℓ, with the reallocation elasticity moving along it through $\mathcal{E} = \kappa(1 - \Delta_\ell^2)$, at $\tau = 0.35$, $\varepsilon_L = 0.3$, $\sigma = 2$, $\hat{g}_1 = 0.3$, $\kappa = 1.5$. The curve is convex over the positive range, so the last stretch toward the corner is worth far more than the first. Moving from the measured +0.41 to the assumed 1 multiplies the term by nearly five. Markers give the measured values from the left panel at the less-than-BA grouping.

answer under either reading, and that a literature imposing $\Delta_\ell = 1$ cannot pose it.

Given Δ_ℓ, the reallocation elasticity is not a free parameter: Lemma 2 pins $\mathcal{E} = \kappa(1 - \Delta_\ell^2)$. At the Roy–Fréchet estimates of Hsieh et al. (2019) and Galle et al. (2023), $\kappa \approx 1.5$ to 2, and at $\Delta_\ell = 0.4$ this gives $\mathcal{E} \approx 1.3$ to 1.7, so the denominator rises from 1.3 to roughly 2.6 to 3.0. Combining,

$$\Omega/p \;\approx\; 1.5\%, \qquad \text{range } 0.8\% \text{ to } 2.2\%,$$

which is the value we carry forward as central. It sits at the bottom of the one to four percent that Costinot and Werning (2023) report for industrial robots. Roughly two thirds of the reduction from the 7.3% corner comes from the exposure gap and one third from reallocation, which reverses the emphasis of the earlier literature: what disciplines the

Table 2: The exposure gap Δ_ℓ, measured

	Ex ante grouping (group 1 = less educated)				
Exposure measure	$<$BA	HS or less	$<$assoc.	job zone 1–2	Ω/p
Webb (2020) software	+0.19	+0.18	+0.20	+0.23	+0.7
Webb (2020) robot	**+0.41**	+0.34	+0.41	+0.45	+1.5
Webb (2020) AI	−0.27	−0.18	−0.23	−0.24	−0.9
Felten et al. (2021) AIOE	−0.49	−0.42	−0.47	−0.50	−1.9
Eloundou et al. (2024), human β	−0.37	−0.34	−0.38	−0.33	−1.3
Eloundou et al. (2024), GPT-4 β	−0.34	−0.31	−0.35	−0.35	−1.2
Eisfeldt et al. (2023) genAI	−0.37	−0.33	−0.36	−0.40	−1.3
Memo: the literature's convention	+1.00	+1.00	+1.00	+1.00	+7.3

Notes. $\Delta_\ell = \ell_{1R} - \ell_{2R}$, the employment-weighted share of group 1 in exposed occupations less that of group 2. Group 1 is the less-educated or lower-preparation group in each column. An occupation is exposed if its score is above the employment-weighted 66th percentile for the two Webb (2020) routine-displacement measures, the Autor–Dorn convention, and above the employment-weighted median otherwise. The panel is 765 detailed SOC-2018 occupations; within-occupation education distributions are the incumbent-reported Required Level of Education from O*NET 30.3 (National Center for O*NET Development, 2026), and employment weights are Dorn (2009) labor-supply weights allocated to SOC. The last column is the predistribution term Ω/p of (11) in percent, evaluated at the $<$BA gap with $\tau = 0.35$, $\varepsilon_L = 0.3$, $\sigma = 2$, $\hat{g}_1 = 0.3$ and $\kappa = 1.5$, so that $\mathcal{E} = \kappa(1-\Delta_\ell^2)$ moves with each row. Appendix D.7 reports the threshold sweep and the construction in full.

Table 3: The sign of the optimal usage tax reduces to one accounting ratio

Recurring	Implied	Rent tax	Mkt.-power	Optimal usage tax $t_U^\star$, domestic		
investment ϕ	rent share r	yield (% GDP)	term, unweighted	unweighted g_S=1, χ=0	**baseline** g_S=.65, χ=.25	linear-dem. bound χ=1
0.10	0.80	0.375	−40%	−35.5%	−15.0%	+4.5%
0.20	0.60	0.300	−30%	−25.5%	−10.1%	+4.5%
0.30	0.40	0.225	−20%	−15.5%	**−5.2**%	+4.5%
0.40	0.20	0.150	−10%	−5.5%	−0.4%	+4.5%
0.50	0.00	0.075	0%	+4.5%	+4.5%	+4.5%

Notes. ϕ is the share of provider revenue absorbed by recurring training and capacity investment; it determines the pure-rent share of the excess markup through $r = 1 - \phi\mu/(\mu - \bar{\mu})$ (Corollary 5). g_S is the welfare weight on AI shareholders relative to the income-weighted average (Proposition 8); χ is the share of the underprovision correction dissipated in the fixed costs of induced entry, derived rather than assumed and equal to $1 - (\sigma_q + 1)/2\sigma_q = 0.25$ at the baseline (Proposition 7). The last column is the linear-demand bound, in which the offset is complete. All three right-hand columns add the predistribution and Pigouvian terms, +1.5% and +3%. The baseline $\phi = 0.3$ is the third row.

robot-tax term is not primarily that workers move but that the losing group was never cleanly identified in the first place.

Table 4 collects the results, and its most striking feature needs addressing head-on rather than left for the reader to notice.

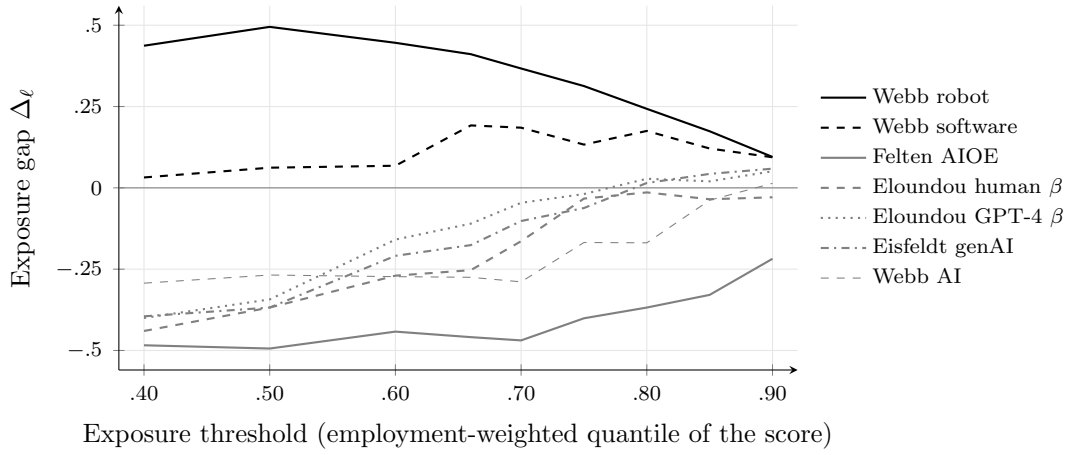


Figure 5: The exposure gap as the definition of "exposed" is tightened, at the less-than-BA grouping. Black lines are the two routine-displacement measures, grey the four direct AI measures. A measure describing genuine segregation of exposure across the education distribution would hold its gap open, or widen it, as the cut moves into the most exposed tail, since the occupations surviving a tight cut would be held overwhelmingly by one group. Six of the seven collapse toward zero instead, and at the ninetieth percentile every gap lies between −0.22 and +0.10. The seventh, Webb software, never leaves a narrow band around +0.1 at any threshold. The vertical scale is the same as the horizontal scale of the left panel of Figure 4, and both stop well short of the value the literature assumes.

6.2 Why the market-power term dominates

Table 4 puts most entries of the last column below zero, and Table 3 shows what governs them. Both deserve a plain reading, because a column of negative numbers invites the objection that the model recommends subsidizing the most profitable firms in the economy, and it does not.

The arithmetic is simple. Predistribution and Pigouvian terms sum to about four and a half percent and are positive in every row. The market-power term, after the three scalings of Section 5.6, is about ten percent at the baseline. It dominates, and the residual is a wedge of about minus five percent. The substantive point is that *a single wedge on the usage margin is being asked to do two incompatible jobs*—to correct an externality and a wage distortion, both of which want a higher price of AI, and to correct underprovision, which wants a lower one—and at defensible parameters the second is the larger.

The right response is not to accept the residual as a recommendation but to stop asking one instrument to do both. Section 5 assigns underprovision to the instruments built for it: a cash flow tax that reaches rent without touching price, and entry and interoperability policy that attacks the concentration directly. Do that, and column (3) is no longer the usage tax's job. Column (1)+(2) is then the operative number, and it is the one we would put to a finance ministry: *a positive wedge of about four and a half*

Table 4: The optimal AI usage tax, by scenario (percent)

	Predistr.	Extern.	Mkt. power	Wedge $t_U^\star$	
Scenario	(1)	(2)	(3)	(1)+(2)	(1)+(2)+(3)
Baseline, domestic (r=0.4)	+1.5	+3	−9.8	+4.5	**−5.2**
Domestic, low rent (r=0.2)	+1.5	+3	−4.9	+4.5	**−0.4**
Domestic, all rent (r=1)	+1.5	+3	−24.4	+4.5	**−19.9**
Importing economy (f=1)	+1.5	+3	0	+4.5	**+4.5**
Emissions priced at source	+1.5	0	−9.8	+1.5	**−8.2**
Unweighted ownership (g_S=1)	+1.5	+3	−15.0	+4.5	**−10.5**
Linear demand (χ=1)	+1.5	+3	0	+4.5	**+4.5**
Wide exposure gap (Δ_ℓ=.6)	+2.2	+3	−9.8	+5.2	**−4.5**
Narrow exposure gap (Δ_ℓ=.2)	+0.8	+3	−9.8	+3.8	**−6.0**
Measured, robot exposure (Δ_ℓ=+.41)	+1.5	+3	−9.8	+4.5	**−5.3**
Measured, LLM exposure (Δ_ℓ=−.37)	−1.3	+3	−9.8	+1.7	**−8.1**
Measured, most negative (Δ_ℓ=−.50)	−2.0	+3	−9.8	+1.0	**−8.8**
Tagging available ($\hat{g}_1$=0)	0	+3	−9.8	+3.0	**−6.8**
All instruments available	0	0	−9.8	0	**−9.8**
Memo: fixed-assignment corner	+7.3	+3	−9.8	+10.3	**+0.6**

Notes. All entries in percent. Column (1) is the predistribution term Ω/p, column (2) the net externality term $(D-\psi)/p$ at the energy-weighted central value, column (3) the market-power term of (12). Two totals are reported and the distinction is the point of the table: (1)+(2) is the efficient wedge once underprovision is addressed by the instruments assigned to it in Section 5, so the usage margin carries only what it is for; (1)+(2)+(3) is the second best of Proposition 2, in which the usage tax is the only instrument available. Parameters: $\tau = 0.35$, $\varepsilon_L = 0.3$, $\sigma = 2$, $\hat{g}_1 = 0.3$, $\Delta_\ell = 0.4$, $\kappa = 1.5$ so that $\mathcal{E} = 1.26$, $\mu = 2.5$, $\bar{\mu} = 1.25$, $\psi = 0$, $d = \$0.05$ per kWh, $p_e = \$0.08$ per kWh, and for column (3) $g_S = 0.65$, $\chi = 0.25$ per Propositions 7 and 8; the two rows that switch these off in turn show what each is worth. The three *measured* rows evaluate column (1) at the values of Table 2 with $\mathcal{E} = \kappa(1-\Delta_\ell^2)$ recomputed row by row; they are the rows to read if one wants the calibration to carry no stipulated value of the exposure gap at all. The final row is the fixed-assignment corner used throughout the robot-tax literature, the $\kappa \to \infty$ limit of our model rather than a baseline.

percent, and under three percent for an economy that has already priced carbon at source, against the ten percent in public discussion. The two readings differ in sign but agree on the thing that matters, which is that no defensible configuration supports a double-digit rate.

Figure 3 and Table 3 add the discipline that makes this checkable. By Corollary 5 the rent share is the complement of the recurring-investment share, so a reader who wants the market-power term to be small is committing to a rent tax that raises correspondingly less. The two live positive proposals—tax the rents, tax the usage—are in tension at the level of the accounting, and the quantity that decides between them, ϕ, is something a disclosure rule could produce. The probabilities below read the same way: they say the rent share

is unlikely to be small enough for the corrective terms to dominate, not that subsidy is the modal recommendation. On the same 200,000 draws, $\Pr(\Omega/p + D/p > 0) = 0.95$ and $\Pr\left(0 < \Omega/p + D/p < 10\%\right) = 0.95$, with a median of 2.9 percent. That pair is the paper's quantitative conclusion stated without reference to the market-power term at all.

6.3 Three calculations

The table prices the tax. The model also prices the things the tax is supposed to achieve, and the orders of magnitude are not what the debate assumes. Throughout we take exposure $\theta = \lambda s = 1.5\%$ of GDP, which is generous given that global spending on AI services is currently well under one percent of world output, and we set $\lambda = 0.3$ and $s = 0.05$.

1. What a usage tax raises, and what it costs. A ten percent tax on a base worth 1.5 percent of GDP raises 0.15 percent of GDP before behavioral response and, by Lemma 1, destroys 0.15 percent of output: about a dollar of output per dollar raised, before Proposition 6 multiplies the output cost by $\Phi = 2.5$. A fifty percent cash flow tax on the operating surplus raises about 0.45 percent of GDP with no output cost, or 0.23 percent on the stationary base that expenses recurring investment first. And at $d = \$0.05$ per kWh the emissions charge has a base of \$9 billion today on U.S. data center load of roughly 180 TWh, rising to \$20–30 billion by 2030 (Shehabi et al., 2024; International Energy Agency, 2026). Section 7 scores all four instruments properly; the point here is only that they are not close, and not in the direction the political debate assumes.

2. What a usage tax does for routine workers. By Lemma 2, a ten percent usage tax raises the routine wage relative to the non-routine wage by

$$\frac{s(\sigma - 1)}{1 + \varepsilon_L + \mathcal{E}} \times 10\% \;=\; \frac{0.05 \times 1}{1.3} \times 10\% \;\approx\; 0.4\%.$$

Four tenths of one percent, and that is the movement in the relative *task* wage at $\mathcal{E} = 1.26$ rather than at the fixed-assignment corner, where it would be 0.39 percent. What the exposed *group* gains is smaller again by the exposure gap: $\Delta_\ell \times 0.2 \approx 0.08$ percent, eight hundredths of one percent. Keeping these two objects apart matters, because the first is what an incidence regression would recover and the second is what enters welfare. This is the entire predistributive payoff of a policy that reduces national output by 0.15 percent, and it accrues as a relative wage movement rather than as income to any identifiable

person. An earned income tax credit of the same fiscal size delivers more to the same workers, without touching the price of compute, and without depending on the value of σ.

There is reason to think even this is an upper bound. **?** study a large, long-lasting employer payroll tax cut for young workers in Sweden and find no effect on their net-of-tax wages even six years on; the response appears as employment, and the windfall accrues to firms that happen to employ many young workers—a rent-sharing rather than a relative-price channel. The mapping is not exact, since their variation taxes labor rather than a substitute for it, but the direction is uncomfortable for the predistribution term. A model in which a factor wedge's incidence is fully expressed in relative wages is the model that maximizes the term in $\hat{g}_1$; if much of it instead shows up in quantities or in firm rents, the wage compression a usage tax buys is smaller still. We flag this as a bound rather than a correction, and note it moves the conclusion in only one direction.

3. Why the wedge is on the wrong base. Take the standard-chat workload, energy 3 Wh and price one cent. Electricity is $p_e e/p = 2.4\%$ of the price of the query, and the residual external cost is sixty to a hundred percent of the price of that electricity, so the external cost of the query is between 1.4 and 2.4 percent of its price. Read backwards the same fact is more striking. A usage tax at a politically meaningful rate, say ten percent, applies a wedge of \$0.001 per query against an external cost near \$0.00015: it over-corrects several-fold for the median query while, by Table 1, *under*-correcting for the agentic tail that drives load growth. A charge of \$0.05 per kWh at the data center prices the harm exactly for every workload at once and would be invisible on any invoice. The point is not merely that the token tax mis-sets the level; it is that no single ad valorem rate can track a harm whose ratio to value varies by an order of magnitude across workloads, whereas a per-kilowatt-hour charge tracks it by construction. That is the whole argument in one line, and the reason is arithmetic rather than ideological: an ad valorem tax on an information good is a poor instrument for a physical externality.

Two further calculations are relegated to Appendix D: what a ninety percent fall in the price of AI does to the wage distribution, which bears on the size of the problem rather than on the instrument; and the welfare cost of choosing the wrong instrument to raise a given sum, which is about eight cents of deadweight per dollar for the usage tax against zero to first order for the rent tax and a first-order gain for the emissions charge.

6.4 A probabilistic calibration

Table 4 reports scenarios, and a scenario table invites the objection that the conclusion is an artifact of one column. Since several parameters are uncertain at once, we take the ranges of Table D.3 at face value, place independent uniform priors on all of them, and report the induced distribution of the optimal tax. This is not an estimate; it is a statement about how much of the parameter space the conclusion occupies. We draw $\sigma \sim U[1.2, 3]$, $\phi \sim U[0.2, 0.4]$, $\hat{g}_1 \sim U[0.1, 0.5]$, $\varepsilon_L \sim U[0.1, 0.5]$, $\kappa \sim U[0, 3]$, $\Delta_\ell \sim U[0.2, 0.6]$, $\mu \sim U[2, 3]$, $g_S \sim U[0.45, 0.85]$, $d \sim U[0.02, 0.10]$, an energy share of price $\sim U[0.02, 0.06]$, and $\psi/p \sim U[0, 0.03]$, with 200,000 draws; Appendix D.4 reports the full distribution and a correlated-prior variant.

Two internal restrictions are imposed draw by draw rather than left to independent priors, and both are consequences of Section 5: $\mathcal{E} = \kappa(1 - \Delta_\ell^2)$, so no draw combines a mobility elasticity and an exposure gap that are jointly infeasible, and $r = 1 - \phi\mu/(\mu - \bar{\mu})$ by Corollary 5, so the rent share is not drawn independently of the recurring investment that determines it. The entry offset χ is not drawn at all: Proposition 7 determines it from σ_q draw by draw, with a resulting median of 0.26.

Four probabilities summarize the exercise, and they are the numbers we would ask a skeptical reader to engage with rather than any single column of Table 4. The optimal usage tax is negative in 76 percent of draws for an economy that owns its providers and in five percent for one that imports, so *ownership rather than technology determines the sign*; but the sharper finding is about magnitude, since the wedge exceeds five percent in under one percent of draws and ten percent in none. The market-power term exceeds the net externality term in nineteen draws out of twenty. The predistribution term stays below five percent in 98 percent of draws, with a median of 1.3 percent once the exposure gap and the reallocation elasticity vary jointly over their full range. And the electricity wedge d/p_e exceeds fifty percent in three quarters of draws, which is why the claim that the wedge on the *kilowatt-hour* is large survives the contested range of γ. Only the last of these is robust to every parameter we cannot measure, which is one more reason the paper's recommendation rests on the meter rather than on the formula.

One prior deserves comment. The draw $\Delta_\ell \sim U[0.2, 0.6]$ is one-sided in a way Table 2 does not support, and we keep it because it is the configuration favourable to the instrument. Widening it to the measured $U[-0.50, 0.45]$, and leaving every other prior alone, moves the predistribution term from a median of $+1.3$ percent to -0.1 percent, with a fifth percentile of -2.3 and a ninety-fifth of $+1.9$, and takes $\Pr(\Omega/p < 0)$ from zero to 0.53. Conditioning instead on the routine-displacement measures alone, $\Delta_\ell \sim U[0.18, 0.45]$,

gives a median of +1.0 percent and $\Pr(\Omega/p > 5\%) = 0.003$. Every probability we report moves against the usage tax under the measured range. The figures in the text are the generous ones.

Appendix D.4 reports the full distribution of each term, the priors, and a robustness check in which the parameters are drawn from a Gaussian copula rather than independently. We impose $\rho(\mu, \phi) = -0.4$, $\rho(\sigma, \hat{g}_1) = 0.3$, $\rho(d, \text{energy share}) = 0.3$, and $\rho(\kappa, \Delta_\ell) = -0.4$, the last because sharp sorting and fast reallocation are substitutes even beyond the deterministic frontier. Correlation fattens the lower tail, the fifth percentile of the domestic optimum falling from -15.2 to -15.9 percent, but leaves the median close to unchanged and moves $\Pr(t_U^\star < 0 \mid f{=}0)$ only from 0.90 to 0.89. The conclusion is not an artifact of the independence assumption.

6.5 Who pays, and who benefits

The paper has so far kept efficiency and distribution in separate compartments. But the missing-market wedge γ has a distributional character of its own, and once it is made explicit the environmental and the equity cases for taxing at the meter stop being separate arguments and become the same argument.

Recall what γ is: the part of the capacity and transmission cost caused by a data center that is recovered from all ratepayers rather than from the facility. It is collected through ordinary electricity bills, so households bear it roughly in proportion to their electricity consumption. Consumption rises with income but far less than proportionately, so the charge is steeply regressive as a share of income. The median U.S. household spends about 2.9 percent of income on energy, low-income households between 6 and 15 percent, and a quarter of low-income households above 15 percent (**??**). A socialized grid cost is therefore borne, relative to income, something like three to five times as heavily by the poorest households as by the median one.

The other side of the ledger runs the opposite way. The benefits of the buildout—rents captured by providers and surplus enjoyed by users of AI—accrue disproportionately to the affluent: to shareholders of a small number of firms, and to professional and business users who are the early adopters of paid services. The socialization of γ is thus a transfer up the income distribution, with a large and growing base.

The consequence for policy is that the welfare weight on the grid-cost term is doing real work, and it points the same way as the efficiency argument. Charging the data center for the capacity it causes, through a dedicated large-load rate class or a load-based Pigouvian charge, corrects the externality and stops the regressive transfer in a single

stroke. The token tax does neither: it cannot reach the grid cost, and its revenue is recycled with no particular distributional tilt. It is unusual for a Pigouvian instrument to have its efficiency and equity cases coincide—the standard worry about carbon taxes is that they are regressive in use even when efficient (**?**). Here the regressivity is in the status quo of socialized cost, and pricing the externality removes it rather than creating it.[4]

7 Policy assessment

The preceding sections characterize an optimum. A finance ministry choosing among live proposals needs something narrower: what each would raise, what it would cost, who would bear it, and whether it can be administered. This section puts the framework to that use. We score stylized versions of the four instruments under discussion, express the comparison in a common welfare metric, ask which of them can be administered, then turn to why the political system may nonetheless reach for the worst one, and close with the measurements that would settle the remaining disagreements.

7.1 Revenue, distortion, and output

Table 5 runs each instrument through the model's own lemmas at the baseline calibration. The rates are chosen to correspond to instruments that have been proposed rather than to be optimal: a ten percent usage tax as the round number that appears in public discussion, three percent as the modal rate among existing digital services taxes, a fifty percent cash flow tax on rents, and a charge of $d = \$0.05$ per kilowatt-hour on data center load. Revenue is net of the behavioral response; the output cost applies the entry amplification Φ of Proposition 6; deadweight loss is of Harberger form.

Three things in the table are worth stating. The rent tax raises between three and five times what a ten percent usage tax raises, depending on whether recurring investment is expensed first, at little or no output cost, because the base is large where the usage base is small and the elasticity near zero where the usage elasticity is large. The usage tax costs forty-four cents of deadweight per dollar raised, an unusually poor ratio for a tax on a growing base, and the reason is entry amplification rather than the size of σ_q. And

[4]Two qualifications. The incidence of the benefit side is inferred from adoption patterns rather than measured, so the transfer is regressive with confidence on the cost side and less on the benefit side; and if a large load lowers average rates by improving asset utilization, the sign can flip. Both point to measuring γ rather than assuming it, and neither disturbs the central point that the grid-cost term carries a distributional charge the token tax cannot touch.

Table 5: Scoring stylized proposals at the baseline calibration

Instrument	Rate	Revenue (% GDP)	DWL (% GDP)	DWL per $ raised	Δ Output (%)
Token / usage tax	10%	0.083	0.037	\$0.44	−0.375
DST-style levy on AI revenue	3%	0.039	0.003	\$0.09	−0.112
Cash flow tax, sunk-stock base	50%	0.450	$\approx 0^{\dagger}$	$\approx \$0^{\dagger}$	$\approx 0^{\dagger}$
Cash flow tax, stationary base (ϕ=0.3)	50%	0.225	$\approx 0^{\dagger}$	$\approx \$0^{\dagger}$	$\approx 0^{\dagger}$
Data center emissions charge	\$0.05/kWh	0.03–0.09	*negative*	*gain*	≈ 0

Notes. Baseline: $\theta = 0.015$, $\sigma_q = 1.965$, $\Phi = 2.5$, $\mu = 2.5$. Revenue for the usage tax and DST is $t\theta(1-\Phi\sigma_q t/(1+t))$; DWL is $\frac{1}{2}\Phi\sigma_q\theta t^2$; output uses Lemma 1 with pass-through Φ. The rent tax is reported on two bases: the sunk-stock reading, in which the operating surplus $(1-1/\mu)\theta$ is fully taxable, and the stationary reading of §B.2, in which recurring training and capacity investment of $\phi = 0.3$ of revenue is expensed first, leaving $[(1-1/\mu)-\phi]\theta$. The second is the more defensible if frontier retraining is permanent rather than a phase. The emissions charge is applied to U.S. data center load of roughly 180 TWh currently and 400–600 TWh by 2030 (Shehabi et al., 2024; International Energy Agency, 2026), giving \$9bn today and \$20–30bn at the end of the decade; its "deadweight loss" is negative because the first \$0.05 per kWh corrects a pre-existing externality worth more per unit than the revenue it raises. †The rent-tax entries are the *theoretical* neutrality of Proposition 4, which holds for an ideal cash flow base and does not survive contact with a real statute; §7.3 puts a magnitude on the departure. Read these entries as "an order of magnitude smaller than the usage tax" rather than as exact zeros.

the emissions charge is not in the same column as the others: it raises real money while improving allocative efficiency, so its entry in the deadweight column is a gain.

7.2 A common metric: the marginal value of public funds

Revenue and deadweight loss are not comparable across instruments whose burdens fall on different people. The marginal value of public funds of Hendren and Sprung-Keyser (2020) supplies a common denominator, willingness to pay of those who bear a reform, over its net cost to government, and for a revenue-raiser a *lower* MVPF means a cheaper dollar. Table D.2 in Appendix D.5 reports it for each instrument; the ranking is the paper's recommendation restated in a metric finance ministries already use.

The emissions charge is cheapest under either ownership structure, at 0.38 to 0.69 in the long run, because part of what looks like a burden is a correction. The cash flow tax follows at 1.00 domestically and 0.00 for an importing economy, which is the terms-of-trade point of Corollary 2 in MVPF form. The usage tax is dearest under both, at 1.18 to 4.43 depending on the rate and on who owns the rents, and its cost rises with the rate because the entry channel compounds.

Two qualifications cut against our preferred instrument and belong in the text rather

than the notes. The emissions charge's advantage is a *long-run* advantage: at the short-run elasticity its MVPF is 0.94 to 0.97, barely below the rent tax, and the gap opens only as the efficiency and siting margins operate. And the corrective benefit must be netted of relocation, since Burke and Abayasekara (2018) find the large long-run industrial elasticity arises precisely because electricity-intensive activity clusters in low-price states, which through our model is leakage rather than abatement, delivering no global benefit. One asymmetry runs against the tariff intuition and is easy to conflate: the usage tax scores *worse* for an importing economy, because the buyer price rises by Φ while the incumbent margin gain leaves the country. Corollary 2 operates through the social marginal cost of the rent, not through the MVPF of the levy.

7.3 The zeros are not honest zeros

Proposition 4 delivers exact neutrality, and Table 5 therefore carries zeros for the cash flow tax. We do not believe those zeros and neither should a finance ministry. They are a property of an idealized base—immediate expensing of all investment, full loss offset carried forward with interest, no distinction between normal return and rent, no ability to relocate the taxable entity—and no enacted statute has had all four. Appendix B.8 lists the four channels through which a wedge creeps back in, each with its drafting remedy.

The claim we defend is weaker and still decisive: the distortion is smaller than a usage tax's by an order of magnitude, not nil, and the magnitudes are second order in the residual wedge, which is what makes the comparison robust even though the zero is not. Suppose the channels together leave an effective wedge on the marginal return of a tenth of the statutory rate, and an investment elasticity of unity—both pessimistic. On a base of 0.9 percent of GDP the deadweight loss is of order 0.005 percent of GDP, or roughly *two cents per dollar raised*, against forty-four cents for the usage tax and nine for the DST. The ordering survives any plausible value, for a structural reason: the usage tax distorts an inframarginal-heavy base at a rate amplified by entry, while the cash flow tax distorts only whatever fraction of the normal return the statute fails to expense. Closing the gap is a drafting problem, not an economic one.

Administrability. Optimal tax formulae assume the base is observable, and following Keen and Slemrod (2017) we treat administrability as part of instrument choice. The emissions charge is the easy case: the base is metered by the utility at the point of delivery for billing purposes that already exist, tons follow from kilowatt-hours and a published grid factor, and electricity delivered to a physical site cannot be booked in

another jurisdiction. The rent tax is demanding but familiar, and its failure mode is political rather than technical. The usage tax is trivially easy to administer, and that is part of the problem: administrative convenience is precisely why it attracts legislatures, and a low enforcement cost is a genuine advantage that must be weighed rather than dismissed—our claim is that it does not come close to offsetting a base that is wrong by an order of magnitude on the harm-to-value ratio. The border adjustment is the hard case and we do not overclaim it; Appendix E.1 sets out why, and the three fallbacks in descending order of fidelity.

7.4 Why we might get the wrong instrument

If the usage tax is poorly suited to every stated goal, why is it the instrument most often proposed? The model contains the beginning of an answer, and it comes from the entry result rather than from any appeal to confusion among policymakers.

Proposition 9 (Incumbents prefer the usage tax)**.** *Compare a usage tax with a revenue-equivalent cash flow tax on rents. The cash flow tax lowers incumbent profit one for one with the revenue it raises. The usage tax, by Proposition 6, raises the equilibrium markup through the entry channel, so incumbents recover a share* $\Phi - 1$ *of the statutory burden through higher margins on retained sales; for sufficiently high fixed cost the surviving incumbents can be left better off than under no tax at all. Incumbents rank the instruments in the opposite order to the planner.*

The content is a positive prediction, not a normative claim: the instrument that is worst for welfare is the one the most concentrated and best-organized interest has least reason to oppose, because it raises rivals' entry cost and lifts margins on the sales that remain. This yields a testable implication that distinguishes the account from mere pessimism. If incumbents prefer the usage tax for the entry-barrier reason, their revealed positions should track firm size and sunk cost: the largest labs least opposed to usage levies and most opposed to rent taxes and interoperability mandates, smaller entrants holding the reverse ranking. That is a pattern one can look for in lobbying disclosures. Appendix E.2 develops two further forces that push the same way and the normative implication, which is not defeatism but sequence: because the worst instrument is the path of least resistance, the reforms most worth defending are the ones the political system will fight hardest.

7.5 What would settle this

The framework's practical value is not that it delivers rates but that it says which measurements would change the answer. We list them in descending order of how much a credible estimate would move the conclusion.

First, the recurring-investment share ϕ, and through it the rent share r. By Corollary 5 this one quantity fixes both how much rent a cash flow tax can reach and how large the market-power correction is, so it decides between the two live positive proposals. Our range rests on an accounting decomposition rather than data. What is needed is a decomposition of frontier inference margins into marginal cost, recurring investment, and residual rent, buildable from public sources—accelerator rental rates against measured throughput, published API prices, training-run costs amortized over plausible inference volume—and exactly the kind of thing a disclosure requirement could produce. We list it first because it is decisive for instrument choice, not because the sign of the usage tax is the interesting question: Section 6.4 finds the level far more robust than the sign, and no value of ϕ delivers a usage tax at the rates proposed.

Second, the exposure gap Δ_ℓ at the individual level. This is what the size of the predistribution term chiefly depends on, and the literature sets it to one by assumption rather than reporting it. Table 2 gives the occupation-level version, in which the ex ante grouping enters through the education distribution of an occupation's incumbents rather than through the workers themselves. That is enough to establish the two facts that matter for the argument—that $|\Delta_\ell|$ is nowhere near one, and that the sign flips between routine and direct-AI measures—but it is not the whole estimand. One year of American Community Survey microdata merged to the same occupation-level scores would deliver the individual-level version, would allow the earnings and age×education groupings we cannot construct from occupational aggregates, and would supply the sampling variance of two proportions, which we cannot. The same exercise should then be run on realized adoption, using the occupation-level index of Bick et al. (2026) once it is released. Exposure measures what AI could reach. The model's σ concerns what it does reach, and the two differ enough to be worth separating (Appendix D.7). We expect it to move the numbers of Table 2 toward zero, since our construction averages out within-occupation heterogeneity in education, and doing so can only make groups look more segregated than they are. That expectation is what wants checking.

Third, the grid-cost wedge γ and its incidence. The staggered adoption of dedicated large-load tariffs across U.S. states is close to an ideal setting for a quasi-experimental estimate of the pass-through of data center load into other customers' rates, and the same

variation would identify the distributional incidence of Section 6.5, which we currently infer from expenditure shares rather than measure. Note that γ barely moves the level of the electricity wedge; what it governs is the equity argument.

Fourth, the leakage rate ℓ. Our single affirmative case for a usage tax rests on a parameter never estimated for compute. The same tariff variation would show whether announced siting shifts toward non-adopting jurisdictions, which is leakage measured directly. Until then Proposition B.1 is a conditional statement rather than a recommendation.

Fifth, the concentration of AI equity. Lemma 4 derives the shareholder weight from consistency with $\hat{g}_1$, and the one input it takes from data is where the equity-weighted dollar of frontier-AI ownership sits in the income distribution. We use the position implied by aggregate corporate equity holdings, but frontier AI is held more narrowly than the market—by founders, early employees and concentrated venture positions—so the relevant z_S is plausibly higher and g_S lower than we assume. Ownership registries and Form 4 filings would sharpen this, and it enters the largest term in the formula.

Two further inputs—the workload composition behind Table 1 and the split of AI energy between training and inference—are not model parameters but calibration inputs no public source reports. Both would be resolved by disclosure rather than by econometrics.

8 Conclusion

The debate about taxing AI has run four different arguments together. Once separated, they point to different instruments and, in one case, in the opposite direction to the one usually assumed.

The externality is real, but it lives in the electricity system rather than in the token. The residual wedge there is between sixty percent and more than the whole of the price of a kilowatt-hour, driven mainly by unpriced carbon, against a few percent for a tax on AI usage—and no single ad valorem rate can track a harm whose ratio to value varies by an order of magnitude across workloads, whereas a charge per kilowatt-hour tracks it by construction. Pricing at the meter also reaches the efficiency margin, so its advantage grows with the horizon, and its efficiency case coincides with its equity case, since the socialized grid cost it corrects is itself a regressive transfer.

The redistributive argument is real and small. It is smaller than the robot-tax literature implies because that literature identifies the losing group with the exposed occupation, which is a modelling convention rather than a fact, and because the parameter usually read as mobility is the same parameter that makes the losing group identifiable in the first

place. The market-power argument runs the other way, but at a third of its textbook size once one asks how much of the markup is rent, what a dollar to a frontier-lab shareholder is worth to a planner who weights routine workers above average, and how much of any correction free entry dissipates in duplicated training runs.

Netting the three, they nearly cancel. Placing priors on every uncertain parameter at once, and imposing the restrictions the model itself generates—mobility tied to the exposure gap, the rent share tied to recurring investment, the entry offset tied to the demand elasticity—the optimal usage tax on a domestic sector is negative in ninety percent of the parameter space, above five percent in under one percent of it, and above ten percent in none. That asymmetry is the practical message. The sign is not the robust finding, and we do not recommend a subsidy: underprovision is the rent tax's problem, and assigning it there leaves an efficient wedge on usage of four and a half percent at most. What is robust is the magnitude. The instrument being legislated is calibrated an order of magnitude above anything the model supports, and the wedge that is unambiguously large sits on the kilowatt-hour rather than on the token.

Each of the three corrections removes an assumption rather than adding evidence, and what survives all of them untouched is the electricity wedge. That is why the recommendation rests on the meter rather than on the formula. Tax the kilowatt-hour, not the token; the rent, not the transaction.

References

Acemoglu, D., Manera, A., and Restrepo, P. (2020). Does the US tax code favor automation? *Brookings Papers on Economic Activity*, 51(1):231–300.

Atkinson, A. B. and Stiglitz, J. E. (1976). The design of tax structure: Direct versus indirect taxation. *Journal of Public Economics*, 6(1–2):55–75.

Auerbach, A. J. and Devereux, M. P. (2018). Cash-flow taxes in an international setting. *American Economic Journal: Economic Policy*, 10(3):69–94.

Autor, D. (2024). Applying AI to rebuild middle class jobs. Working Paper 32140, National Bureau of Economic Research. Published in *NOEMA Magazine* as "How AI Could Help Rebuild The Middle Class".

Autor, D. and Thompson, N. (2025). Expertise. *Journal of the European Economic Association*, 23(4):1203–1271.

Autor, D. H. and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review*, 103(5):1553–1597.

Bick, A., Blandin, A., Deming, D. J., and Schumacher, T. (2026). What work does generative AI do? Working Paper 2026-017, Federal Reserve Bank of St. Louis. Task- and occupation-

level generative-AI adoption indexes from the Real-Time Population Survey. The replication package containing the indexes had not been released at the time of writing.

Bonney, K., Breaux, C., Dinlersoz, E., Foster, L., Haltiwanger, J., and Pande, A. (2026). The microstructure of AI diffusion: Evidence from firms, business functions, and worker tasks. CES Working Paper 26-25, U.S. Census Bureau, Center for Economic Studies. Based on the 2026 AI supplement to the Business Trends and Outlook Survey (BTOS).

Borenstein, S. (2025). What will data centers do to your electric bill? Energy Institute at Haas, Blog Post.

Borenstein, S. and Bushnell, J. B. (2022). Do two electricity pricing wrongs make a right? cost recovery, externalities, and efficiency. *American Economic Journal: Economic Policy*, 14(4):80–110.

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

Burke, P. J. and Abayasekara, A. (2018). The price elasticity of electricity demand in the United States: A three-dimensional analysis. *The Energy Journal*, 39(2):123–146.

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

Cournot, A. A. (1838). *Recherches sur les principes mathématiques de la théorie des richesses*. Hachette, Paris.

De Loecker, J., Eeckhout, J., and Unger, G. (2020). The rise of market power and the macroeconomic implications. *Quarterly Journal of Economics*, 135(2):561–644.

de Vries, A. (2023). The growing energy footprint of artificial intelligence. *Joule*, 7(10):2191–2194.

Delipalla, S. and Keen, M. (1992). The comparison between ad valorem and specific taxation under imperfect competition. *Journal of Public Economics*, 49(3):351–367.

Diamond, P. A. and Mirrlees, J. A. (1971). Optimal taxation and public production I: Production efficiency. *American Economic Review*, 61(1):8–27.

Dorn, D. (2009). Essays on inequality, spatial interaction, and the demand for skills. *Dissertation 3613, University of St. Gallen*. Source of the `occ1990dd` harmonized occupation scheme and the associated labor-supply weights; files at `https://www.ddorn.net/data.htm`.

Eisfeldt, A. L., Schubert, G., Taska, B., and Zhang, M. B. (2023). Generative ai and firm values. *NBER Working Paper 31222*. Occupation-level generative-AI exposure scores `genaiexp_estz` by SOC 2010.

Eloundou, T., Manning, S., Mishkin, P., and Rock, D. (2024). GPTs are GPTs: Labor market impact potential of LLMs. *Science*, 384(6702):1306–1308. Occupation-level α, β, ζ exposure scores under both the human and the GPT-4 rubric.

Energy and Environmental Economics (E3) (2026). Understanding the drivers of rising electricity rates and the role of data centers. Technical report, E3 White Paper.

Felten, E. W., Raj, M., and Seamans, R. (2021). Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal*, 42(12):2195–2217. AI Occupational Exposure (AIOE) scores by SOC 2010, Data Appendix A.

Fischer, C. and Fox, A. K. (2012). Comparing policies to combat emissions leakage: Border

carbon adjustments versus rebates. *Journal of Environmental Economics and Management*, 64(2):199–216.

Galle, S., Rodríguez-Clare, A., and Yi, M. (2023). Slicing the pie: Quantifying the aggregate and distributional effects of trade. *The Review of Economic Studies*, 90(1):331–375.

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

Hendren, N. and Sprung-Keyser, B. (2020). A unified welfare analysis of government policies. *The Quarterly Journal of Economics*, 135(3):1209–1318.

Hoel, M. (1996). Should a carbon tax be differentiated across sectors? *Journal of Public Economics*, 59(1):17–32.

Hsieh, C.-T., Hurst, E., Jones, C. I., and Klenow, P. J. (2019). The allocation of talent and U.S. economic growth. *Econometrica*, 87(5):1439–1474.

Ide, E. and Talamàs, E. (2025). Artificial intelligence in the knowledge economy. *Journal of Political Economy*, 133(12):3761–3800.

International Energy Agency (2025). Energy and AI. Technical report, IEA, Paris.

International Energy Agency (2026). Key questions on energy and AI. Technical report, IEA, Paris.

Jacobs, M. (2025). Data centers and the connection costs loophole: Billions in transmission costs passed to PJM customers. Technical report, Union of Concerned Scientists.

Joint Legislative Audit and Review Commission (2024). Data centers in Virginia. Technical report, Commonwealth of Virginia.

Kaplow, L. (2012). Optimal control of externalities in the presence of income taxation. *International Economic Review*, 53(2):487–509.

Keen, M. (1998). The balance between specific and ad valorem taxation. *Fiscal Studies*, 19(1):1–37.

Keen, M. and Slemrod, J. (2017). Optimal tax administration. *Journal of Public Economics*, 152:133–142.

Kydland, F. E. and Prescott, E. C. (1977). Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy*, 85(3):473–491.

Labandeira, X., Labeaga, J. M., and López-Otero, X. (2017). A meta-analysis on the price elasticity of energy demand. *Energy Policy*, 102:549–568.

Luccioni, A. S., Jernite, Y., and Strubell, E. (2024). Power hungry processing: Watts driving the cost of AI deployment? In *ACM Conference on Fairness, Accountability, and Transparency*, pages 85–99.

Mankiw, N. G. and Whinston, M. D. (1986). Free entry and social inefficiency. *RAND Journal of Economics*, 17(1):48–58.

Markusen, J. R. (1975). International externalities and optimal tax structures. *Journal of International Economics*, 5(1):15–29.

Monitoring Analytics (2025). Analysis of the 2025/2026 RPM base residual auction. Technical report, Independent Market Monitor for PJM.

Myles, G. D. (1996). Imperfect competition and the optimal combination of ad valorem and specific taxation. *International Tax and Public Finance*, 3(1):29–44.

Naito, H. (1999). Re-examination of uniform commodity taxes under a non-linear income tax

system and its implication for production efficiency. *Journal of Public Economics*, 71(2):165–188.

National Center for O*NET Development (2026). O*net 30.3 database. U.S. Department of Labor, Employment and Training Administration. `https://www.onetcenter.org/database.html`. Used here: `education.csv` (Required Level of Education, scale RL), `job_zones.csv`, `occupation_data.csv`. Licensed CC BY 4.0.

Natural Resources Defense Council (2025). Data center growth and the cost to PJM ratepayers. Technical report, NRDC Issue Brief.

Nordhaus, W. D. (2017). Revisiting the social cost of carbon. *Proceedings of the National Academy of Sciences*, 114(7):1518–1523.

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

Riu, I., Patel, K., Mettetal, L., et al. (2025). Tailored for scale: Designing electric rates and tariffs for large loads. Technical report, Energy and Environmental Economics (E3).

Rothschild, C. and Scheuer, F. (2013). Redistributive taxation in the Roy model. *The Quarterly Journal of Economics*, 128(2):623–668.

Sachs, D., Tsyvinski, A., and Werquin, N. (2020). Nonlinear tax incidence and optimal taxation in general equilibrium. *Econometrica*, 88(2):469–493.

Saez, E. (2004). Direct or indirect tax instruments for redistribution: Short-run versus long-run. *Journal of Public Economics*, 88(3–4):503–518.

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

Scheuer, F. (2014). Entrepreneurial taxation with endogenous entry. *American Economic Journal: Economic Policy*, 6(2):126–163.

Scheuer, F. and Werning, I. (2016). Mirrlees meets Diamond–Mirrlees. NBER Working Paper 22076, National Bureau of Economic Research.

Scheuer, F. and Werning, I. (2017). The taxation of superstars. *Quarterly Journal of Economics*, 132(1):211–270.

Shehabi, A., Smith, S. J., Hubbard, A., et al. (2024). 2024 united states data center energy usage report. Technical report, Lawrence Berkeley National Laboratory.

Stiglitz, J. E. (1982). Self-selection and Pareto efficient taxation. *Journal of Public Economics*, 17(2):213–240.

Suits, D. B. and Musgrave, R. A. (1953). Ad valorem and unit taxes compared. *Quarterly Journal of Economics*, 67(4):598–604.

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

Tomlinson, K., Jaffe, S., Wang, W., Counts, S., and Suri, S. (2025). Working with AI: Measuring the applicability of generative AI to occupations. *arXiv preprint arXiv:2507.07935*. Microsoft Research. Version 6, 22 December 2025. Occupation-level AI applicability scores derived from 200,000 Bing Copilot conversations classified into O*NET work activities; aggregated measures at `https://github.com/microsoft/working-with-ai`.

U.S. Bureau of Labor Statistics (2018). 2010 to 2018 SOC crosswalk. Standard Occupational Classification. `https://www.bls.gov/soc/2018/soc_2010_to_2018_crosswalk.xlsx`.

U.S. Environmental Protection Agency (2023). Report on the social cost of greenhouse gases: Estimates incorporating recent scientific advances. Technical report, EPA.

Webb, M. (2020). The impact of artificial intelligence on the labor market. *SSRN Working Paper 3482150*. Occupation-level software, robot and AI exposure percentiles by `occ1990dd`, with 2010 labor-supply weights; replication file `exposure_by_occ1990dd_lswt2010.csv`.

Wiser, R., Barbose, G., Gorman, W., et al. (2026). Retail electricity price trends and drivers: Data update, 2026 edition. Technical report, Lawrence Berkeley National Laboratory.

Yin, M., Vu, H., and Persico, C. (2026). How (un)Stable are LLM occupational exposure scores? evidence from multi-model replication. Working Paper 35110, National Bureau of Economic Research, Cambridge, MA. Author forenames not given in the source we transcribed this entry from; complete before submission.

A Proofs

Proof of Lemma 1

The final good is the numeraire, so its price equals unit cost, which by (1) is $c_A^{\lambda} w_N^{1-\lambda}/\Lambda$. By (2), $\partial \ln c_A / \partial \ln q = s$, so the cost index rises by $\lambda s \, \mathrm{d}\ln q$ at fixed factor prices. With the output price fixed at one, quantity falls by the same proportion to first order. ■

Proof of Lemma 2

From (9), $w_R/w_N = [\lambda(1-s)/(1-\lambda)] \cdot (L_N/L_R)$. Let $x \equiv \mathrm{d}\ln(w_R/w_N)$. Differentiation and (2) give $\mathrm{d}\ln(1-s) = -\frac{s}{1-s}\mathrm{d}\ln s = s(\sigma-1)\mathrm{d}\ln q$, so $x = s(\sigma-1)\mathrm{d}\ln q - \mathrm{d}\ln(L_R/L_N)$. Labor supply satisfies $\mathrm{d}\ln L_j = \varepsilon_L \mathrm{d}\ln[(1-\tau)w_j]$, hence $\mathrm{d}\ln(L_R/L_N) = \varepsilon_L x$. Solving for x gives the result. ■

Proof of Lemma 2 and Proposition 5

Relative wages. From (9), $w_R/w_N = [\lambda(1-s)/(1-\lambda)]\cdot(L_N/L_R)$, so with $x \equiv \mathrm{d}\ln(w_R/w_N)$ and using (2),

$$x \;=\; s(\sigma-1)\,\mathrm{d}\ln q \;-\; \mathrm{d}\ln(L_R/L_N).$$

Relative task labor responds through two margins. The intensive hours margin contributes $\varepsilon_L x$. The reallocation margin contributes $\mathcal{E}x$, where $\mathcal{E}$ is defined by $\mathrm{d}\ln(L_R/L_N)\big|_{\text{reallocation}} = \mathcal{E}\,x$. The two are additive because they operate on distinct decisions: how many hours to supply, and to which task. Substituting and solving gives $x = s(\sigma-1)\mathrm{d}\ln q/(1+\varepsilon_L+\mathcal{E})$.

The identity $\mathcal{E} = \kappa(1-\Delta_\ell^2)$. Differentiating the Fréchet share in (7) gives $\mathrm{d}\ln \ell_{iR} = \kappa(1-\ell_{iR})\,x$ and $\mathrm{d}\ln(1-\ell_{iR}) = -\kappa\,\ell_{iR}\,x$. Let s_{iR} denote group i's share of routine employment and s_{iN} its share of non-routine employment. Aggregating,

$$\mathcal{E} \;=\; \kappa\Big[\textstyle\sum_i s_{iR}\,(1-\ell_{iR}) \;+\; \sum_i s_{iN}\,\ell_{iR}\Big].$$

Under the symmetric specification of (6), with equal group masses and $A_{1R}/A_{1N} = e^{\varrho/2}$, $A_{2R}/A_{2N} = e^{-\varrho/2}$, we have $\ell_{2R} = 1-\ell_{1R}$, hence $s_{1R} = \ell_{1R}$, $s_{2R} = 1-\ell_{1R}$, $s_{1N} = 1-\ell_{1R}$, $s_{2N} = \ell_{1R}$, and the bracket collapses to $4\ell_{1R}(1-\ell_{1R})$. Since $\Delta_\ell = \ell_{1R} - \ell_{2R} = 2\ell_{1R} - 1$, we have $4\ell_{1R}(1-\ell_{1R}) = 1-\Delta_\ell^2$, giving $\mathcal{E} = \kappa(1-\Delta_\ell^2)$. Evaluating ℓ_{1R} at $w_R = w_N$ gives $\ell_{1R} = (1+e^{-\kappa\varrho/2})^{-1}$ and therefore $\Delta_\ell = \tanh(\kappa\varrho/4)$.

The distributional value. By (8), $\mathrm{d}\ln(W_1/W_2) = \Delta_\ell\, x$: the switching margin drops out by the envelope theorem, since the marginal switcher is indifferent, and the relative welfare

movement between groups is the relative task-wage movement scaled by the exposure gap. Substituting into the evaluation of Ω in the proof of Proposition 2, with $\hat{g}_1$ defined on ex ante groups so that $\sum_i \hat{g}_i W_i = 0$ holds throughout the reform, replaces $s(\sigma-1)/(1+\varepsilon_L)$ by $\Delta_\ell\, s(\sigma-1)/(1+\varepsilon_L+\mathcal{E})$ and leaves every other step unchanged, since neither Δ_ℓ nor $\mathcal{E}$ enters the resource cost of a unit of AI or the external cost D.

Comparative statics. It is convenient to change variable to $u \equiv \kappa\varrho/4 \geq 0$, so that $\Delta_\ell = \tanh u$ and, writing $t \equiv \tanh u$ and $\operatorname{sech}^2 u = 1 - t^2$,

$$\mathcal{E} \;=\; \kappa\,(1-t^2) \;=\; \frac{4u}{\varrho}\operatorname{sech}^2 u.$$

For (i): $\mathcal{E} \to 0$ as $u \to 0$ because $u \to 0$, and $\mathcal{E} \to 0$ as $u \to \infty$ because $u\operatorname{sech}^2 u \leq 4u\,e^{-2u} \to 0$; since $\mathcal{E} > 0$ in between, it is hump-shaped, with the interior maximum at the unique root of $2u\tanh u = 1$. Part (ii) is immediate from the strict monotonicity of tanh on $[0,\infty)$.

For (iii), write $A \equiv 1+\varepsilon_L > 0$ and $f(u) \equiv \Delta_\ell/(A+\mathcal{E}) = t/(A+(4u/\varrho)\operatorname{sech}^2 u)$, so that Ω/p is $f(u)$ times a positive constant. Differentiating, and using $d\mathcal{E}/du = (4/\varrho)\operatorname{sech}^2 u\,(1-2u\tanh u)$,

$$f'(u) \;=\; \frac{\operatorname{sech}^2 u}{(A+\mathcal{E})^2}\,\Big[\,A \;+\; \tfrac{4}{\varrho}\big(u - t + u\,t^2\big)\Big].$$

Every term in the bracket is nonnegative: $u\,t^2 \geq 0$, and $u \geq \tanh u = t$ for all $u \geq 0$. Hence the bracket is bounded below by $A > 0$ and $f'(u) > 0$ for all $u > 0$, strictly and without any auxiliary condition. The limits are $f(0) = 0$ and $f(\infty) = 1/A = 1/(1+\varepsilon_L)$, the latter being exactly the fixed-assignment value, which establishes (iii). We note that the bound is tight in the sense that the bracket attains A only in the limit $u \to 0$, so the monotonicity does not depend on the size of ϱ: it holds uniformly over the comparative-advantage gap.

Part (iv) follows from $\mathcal{E} \geq 0$, which gives $\Omega/p \leq \Delta_\ell \cdot (1-\tau)\hat{g}_1(\sigma-1)(1-s)/[(1+\varepsilon_L)\sigma_q]$, with equality only when $\mathcal{E} = 0$, that is at $\kappa \in \{0,\infty\}$; and $\Delta_\ell > 0$ rules out $\kappa = 0$. □

Proof of Proposition 1

Under these conditions the economy satisfies the hypotheses of Diamond and Mirrlees (1971). Producer prices can be held equal to social marginal costs while any feasible revenue is raised on final uses and factor incomes. A positive t_U places a wedge on the intermediate margin, which by Lemma 1 raises the unit cost of the economy by $\lambda s\, t_U$ to first order and lowers output. Replace t_U by zero and adjust the income tax to

hold revenue constant: the tagged income tax attains the same distribution at the same revenue with strictly larger output. Separability and type-contingent taxes also make AI intensity uninformative once earnings are observed, so by Atkinson and Stiglitz (1976) no differentiation of labor taxes by AI usage is desirable. ■

Proof of Proposition 2

Write the domestic social marginal cost of a unit of AI. The short-run resource cost is m. Netting out the economy-wide benchmark $\bar{\mu}$, the excess markup $(\mu - \bar{\mu})m$ splits, by (5), into a fixed-cost-recovery part $(1-r)(\mu-\bar{\mu})m = F/y$ and a pure-rent part $r(\mu-\bar{\mu})m$. The fixed-cost part is a genuine resource cost: it pays for F, without which the good is not produced, so it belongs in social marginal cost on the same footing as m. The rent part is a transfer; of it, the share f leaves the country and is a domestic cost, while the share $1-f$ is a domestic transfer that cancels. A uniform markup $\bar{\mu}$ elsewhere acts like a uniform consumption tax and is not distortionary at the margin, so the correct benchmark nets it out. Collecting terms, the domestic social marginal cost is

$$\text{SMC} \;=\; \bar{\mu}\, m \;+\; (1-r)(\mu-\bar{\mu})\, m \;+\; f\, r\, (\mu-\bar{\mu})\, m \;+\; D,$$

where the second term restores fixed-cost recovery above the benchmark and the third counts foreign rent. Firms equate the marginal value of AI to $q = (1+t_U)\mu m$. Optimality requires the buyer price to exceed social marginal cost by the shadow value Ω, per unit of AI, of the wage compression a marginal reduction in a delivers:

$$(1+t_U^{\star})\,\mu\, m \;=\; \bar{\mu}\, m + (1-r)(\mu-\bar{\mu})m + f\, r\, (\mu-\bar{\mu})m + D + \Omega. \tag{13}$$

Divide by $\mu m = p$. The first two terms on the right combine to $\mu m - r(\mu-\bar{\mu})m$ over μm, that is $1 - r(1-\bar{\mu}/\mu)$, and adding the foreign-rent term gives $1-(1-f)\, r(1-\bar{\mu}/\mu)$. Rearranging,

$$t_U^{\star} \;=\; \frac{\Omega + D}{p} \;-\; (1-f)\, r \left(1 - \frac{\bar{\mu}}{\mu}\right),$$

which is (11) once Ω/p is evaluated. The rent share r now scales the market-power term, so a markup that is entirely fixed-cost recovery ($r = 0$) contributes nothing, as claimed.

To evaluate Ω: with the income tax optimal, the distributional value of the reform is the sum of net welfare weights times changes in net income, and by Lemma 3 the common component drops under its covariance condition. Using $\sum_i \hat{g}_i W_i = 0$ over ex ante groups, so that $\hat{g}_1 W_1 \;=\; -\hat{g}_2 W_2$, and (8) to convert task-wage movements into group-welfare

movements,

$$\begin{aligned}\frac{dW^{\text{dist}}}{\mathrm{d}\ln q} &= (1-\tau)\Big[\hat{g}_1 w_R L_R \mathrm{d}\ln w_R + \hat{g}_N w_N L_N \mathrm{d}\ln w_N\Big]\Big/\mathrm{d}\ln q \\ &= (1-\tau)\,\hat{g}_1\, w_R L_R \cdot \frac{\mathrm{d}\ln(w_R/w_N)}{\mathrm{d}\ln q},\end{aligned}$$

which by Lemma 2 and (8) equals $(1-\tau)\hat{g}_1 w_R L_R\, \Delta_\ell\, s(\sigma-1)/(1+\varepsilon_L+\mathcal{E})$. The associated fall in AI use is $-da = \sigma_q a\, \mathrm{d}\ln q$ by (10), so

$$\frac{\Omega}{p} = \frac{1}{p}\cdot\frac{dW^{\text{dist}}/\mathrm{d}\ln q}{\sigma_q\, a} = \frac{(1-\tau)\hat{g}_1\,\Delta_\ell\, s(\sigma-1)}{(1+\varepsilon_L+\mathcal{E})\sigma_q}\cdot\frac{w_R L_R}{p\,a}.$$

By (9), $w_R L_R/(pa) = \lambda(1-s)/(\lambda s) = (1-s)/s$. The factor s cancels, leaving the stated expression. ■

Proof of Proposition 3

Under t_E the provider's energy price becomes $p_e + t_E$ and the efficiency condition becomes $\varphi'(\xi) = -(p_e + t_E)e'(\xi)$. At $t_E = d$ the provider internalizes the full social cost of a kilowatt-hour, chooses $\xi^\star$, and passes the resulting marginal cost into the price, so the scale margin is correct as well. Under t_U the provider's energy price is unchanged, the efficiency condition is unchanged, and $\xi = \xi^0$ regardless of the rate: a usage tax is a wedge on output, not on energy, and the energy intensity of provision is not a function of the output price. The residual loss is the excess of the social cost of provision at ξ^0 over its minimized value at $\xi^\star$, which is Δ. Strict convexity makes $\Delta > 0$ unless $e' \equiv 0$. ■

Proof of Proposition 4

The provider maximizes $(1-z)[P(Q)Q - m(\xi)Q - F]$. Since z scales the objective, the maximizing Q, P, and ξ are unchanged, and revenue $z\Pi$ is raised without distortion. The usage tax raises $t_U pa$ and induces $\mathrm{d}\ln a = -\sigma_q t_U$ to first order, generating a Harberger loss of order $\frac{1}{2}\sigma_q\, pa\, t_U^2 > 0$. ■

Proof of Proposition B.1

(i) A marginal unit of domestic abatement reduces global emissions by $1-\ell$, so its marginal social benefit is $(1-\ell)d$ per kilowatt-hour while its marginal cost is borne domestically at full value. This is the leakage adjustment of Markusen (1975) and Hoel (1996). (ii)

The share ℓ of domestically consumed AI is served by compute abroad and escapes $t_E^\star$; its embodied residual damage is ℓD per unit. A levy on the buyer at $\ell D/p$ applies exactly that wedge and is invariant to the location of the server. (iii) Adding the pass-through of $(1-\ell)d \cdot e$ to the levy ℓD gives D. Fischer and Fox (2012) establish the analogue for border carbon adjustment in goods trade. ■

Proof of Proposition 6

We work with the conduct model of (4) and hold the pass-through discipline of (17). At fixed n, the constant-elasticity markup is multiplicative, so $\mathrm{d}\ln q = \mathrm{d}\ln(1+t_U) + \mathrm{d}\ln\mu + \mathrm{d}\ln m$; with m fixed and μ fixed at fixed n, short-run pass-through is $\rho_U = 1$, establishing the first claim of (17). The amplification is therefore entirely an entry effect. Per-firm profit under the symmetric conduct (4) is $\pi = (p-m)y - F$ with $y = Q/n$; at the Cournot value $\chi = 1/n$ this is $\pi = pQ/(n^2\sigma_q) - F$. Free entry $\pi = 0$ gives $n = (pQ/(\sigma_q F))^{1/2}$, hence $\mathrm{d}\ln n = \frac{1}{2}\mathrm{d}\ln(pQ)$. Producer revenue is $pQ = qQ/(1+t_U)$ and, by constant-elasticity demand, $\mathrm{d}\ln(qQ) = (1-\sigma_q)\mathrm{d}\ln q$, so around $\mathrm{d}\ln q \approx \mathrm{d}\ln(1+t_U)$ we have $\mathrm{d}\ln(pQ) = -\sigma_q \mathrm{d}\ln(1+t_U)$ and $\mathrm{d}\ln n = -\frac{1}{2}\sigma_q \mathrm{d}\ln(1+t_U)$. Writing $x \equiv \chi/\sigma_q = (\mu-1)/\mu$ with $\chi = 1/n$, we have $\mathrm{d}\ln x = -\mathrm{d}\ln n$ at fixed σ_q, and $\mathrm{d}\ln\mu/\mathrm{d}\ln x = \mu - 1$, so $\mathrm{d}\ln\mu/\mathrm{d}\ln(1+t_U) = (\mu-1)\cdot\frac{1}{2}\sigma_q$. Total pass-through including entry is $\Phi = \mathrm{d}\ln q/\mathrm{d}\ln(1+t_U) = 1 + \frac{1}{2}(\mu-1)\sigma_q$. Since the optimality condition (13) equates the buyer price to social marginal cost, and the buyer price now rises by $\Phi > 1$ per unit of tax while social marginal cost is unchanged, the tax satisfying (13) is strictly below its fixed-structure value. The distinction from Proposition 3 is clean: there, at fixed n, both instruments have unit pass-through and the emissions charge wins on the efficiency margin; here the entry margin adds a penalty that falls on the usage tax alone, because the emissions charge raises m symmetrically and does not shrink the number of firms relative to the market it serves. ■

Proof of the pass-through identities (17)

The equilibrium price is $p = \mu(n)\, m$ with $\mu(n) = (1-\chi/\sigma_q)^{-1}$ and χ, σ_q constant at fixed n. An emissions charge raises the marginal cost multiplicatively, $m \to m(1+\hat{t}_E)$ with $\hat{t}_E$ the proportional cost increase, so $\mathrm{d}\ln p = \mathrm{d}\ln m$ and $\rho_E = 1$. The usage tax enters the buyer price as $q = (1+t_U)p$, so at fixed n (hence fixed μ and m) $\mathrm{d}\ln q = \mathrm{d}\ln(1+t_U)$ and $\rho_U = 1$. Both identities hold exactly for constant-elasticity demand and any conduct χ; they fail only through the entry channel $\mathrm{d}\ln n \neq 0$, which is the subject of Proposition 6.
■

Proof of Lemma 3

Let the reform induce $\mathrm{d}\ln w_j = \mathrm{d}\ln \bar{w} + \varepsilon_j$, with the relative component normalized by $\sum_j w_j L_j\, \varepsilon_j = 0$. Evaluate welfare at the optimal (τ, T): by the envelope theorem the induced re-optimization of the tax has no first-order effect, so we may compute the direct effects of the wage change holding (τ, T) fixed. Measured in units of public funds, welfare moves through two channels. By Roy's identity the private valuation of the wage change is $(1-\tau)\sum_j g_j\, w_j L_j\, \mathrm{d}\ln w_j$, where g_j is the social marginal welfare weight normalized so that its population mean is one. The induced change in revenue is $\tau \sum_j w_j L_j (1+\varepsilon_L)\, \mathrm{d}\ln w_j$, since a wage change of $\mathrm{d}\ln w_j$ raises earnings both mechanically and through the labor-supply response ε_L that weak separability isolates; this revenue is returned lump sum.

The optimal linear tax satisfies the first-order condition

$$\sum_j g_j\, w_j L_j \;=\; \sum_j w_j L_j \Big(1 - \tfrac{\tau}{1-\tau}\, \varepsilon_L\Big),$$

i.e. the income-weighted average welfare weight equals $1 - \frac{\tau}{1-\tau}\varepsilon_L$. Apply this to the common component $\mathrm{d}\ln w_j = \mathrm{d}\ln \bar{w}$: its total contribution is

$$\mathrm{d}\ln \bar{w} \left[(1-\tau)\sum_j g_j\, w_j L_j + \tau(1+\varepsilon_L)\sum_j w_j L_j\right] = \mathrm{d}\ln \bar{w}\, \sum_j w_j L_j,$$

where the first-order condition collapses the bracket to $\sum_j w_j L_j$. This is precisely the marginal value of the change in aggregate labor income, the same object the resource accounting of Lemma 1 already books through $\mathrm{d}\ln Y$; it is an efficiency term, not a distributional one, and to avoid double counting it does not re-enter the distributional shadow value Ω. What remains for Ω is the relative component, $(1-\tau)\sum_j \hat{g}_j\, w_j L_j\, \varepsilon_j$, in which the common part has vanished because $\sum_j \hat{g}_j\, w_j L_j = 0$ by the definition of the income-weighted excess weights $\hat{g}_j \equiv g_j - \bar{g}$.

If the reform also shifts a distributional moment orthogonal to a proportional shift, a linear tax cannot neutralize it; writing its projection on the rank statistic as $\mathrm{Cov}_j(\hat{g}_j, \mathrm{rank}_j)$, a residual of order $(\mathrm{d}\ln \bar{w})\,\mathrm{Cov}_j(\hat{g}_j, \mathrm{rank}_j)$ remains, vanishing iff the covariance is zero. ■

Proof of Lemma D.1

With a CES outer aggregator of elasticity η, the automatable value share is $\lambda(q) = \lambda_0 c_A^{1-\eta}/(\lambda_0 c_A^{1-\eta} + (1-\lambda_0) w_N^{1-\eta})$, whose elasticity with respect to c_A is $(1-\eta)\lambda(1-\lambda)$, and $\partial \ln c_A / \partial \ln q = s$. The predistribution term is $\Omega/p \propto \lambda(1-s)/\sigma_q$ times the share-weighting

that in the Cobb–Douglas case is constant; differentiating, the CES case multiplies it by $1 + (\eta - 1)\phi$ with $\phi = s\lambda(1-\lambda)/\lambda = s(1-\lambda)$. Since $s = \theta/\lambda$ and $\theta < 0.02$, $\phi < 0.02$ and $|(\eta - 1)\phi| < 0.02$ for $\eta \in [0.5, 2]$. Hence the correction is below two percent and the cancellation of θ is exact to first order. ■

Proof of Proposition B.2

The first part is the standard Pigouvian argument applied to K, together with the observation that the firm's choice of K is a decision about model capability, taken before and separately from the volume of inference sold, so the two are not proportional in general. For the second part, a threshold sets an implicit marginal tax of zero below $\bar{K}$ and infinity at $\bar{K}$. With heterogeneous providers, the marginal value of the last unit of compute differs across firms at the threshold, so the allocation of aggregate training compute across firms is not cost-minimizing for the risk attained. A uniform levy equalizes marginal value across firms at R', which is the standard equimarginal principle. ■

Proof of Proposition B.3

Write $H(T) \equiv \sum_{s=1}^{T} \rho^s$, which is positive and increasing in T. The efficiency first order condition (15) is $\varphi'(\xi) = -H(T)(p_e + t)e'(\xi)Q$. Define $G(\xi; a) \equiv \varphi'(\xi) + a\, e'(\xi)$ with $a \equiv H(T)(p_e + t)Q > 0$; the optimum solves $G = 0$. Since φ and e are strictly convex with $\varphi' > 0$, $\varphi'' > 0$, $e' < 0$, $e'' > 0$, we have $\partial G/\partial \xi = \varphi'' + a\, e'' > 0$, so the optimal ξ is unique, and $\partial G/\partial a = e' < 0$. By the implicit function theorem $d\xi/da = -(\partial G/\partial a)/(\partial G/\partial \xi) = -e'/(\varphi'' + ae'') > 0$. Thus ξ is increasing in a.

(i) The emissions charge sets $t = d > 0$, giving $a^E = H(T)(p_e + d)Q$, while the usage tax leaves the perceived energy price at p_e, giving $a^U = H(T)p_eQ < a^E$. Since ξ increases in a, $\xi^E(T) > \xi^U(T)$ for every $T \geq 1$.

(ii) a is increasing in T through $H(T)$, so both efficiencies rise with the horizon. For the gap, write $\xi^E - \xi^U = \int_{a^U}^{a^E} \xi'(a)\, da$ with $\xi'(a) = -e'/(\varphi'' + ae'') > 0$ and $a^E - a^U = H(T)dQ$, so the interval of integration widens in T while its left endpoint $a^U = H(T)p_eQ$ also moves right. Two forces therefore act in opposite directions, and the gap is increasing in T if and only if the widening dominates the decay of ξ'. A sufficient condition is that $a\,\xi'(a)$ be nondecreasing, i.e. that the elasticity of ξ with respect to the effective energy price not fall as that price rises; this holds, for instance, in the isoelastic case $\varphi(\xi) = \varphi_0 \xi^{1+\nu}$ with $e(\xi) = e_0 \xi^{-\epsilon}$, for which $\xi \propto a^{1/(1+\nu+\epsilon)}$ and the gap scales as $H(T)^{1/(1+\nu+\epsilon)}\big[(p_e + d)^{1/(1+\nu+\epsilon)} - p_e^{1/(1+\nu+\epsilon)}\big]Q^{1/(1+\nu+\epsilon)}$, manifestly increasing in T. We adopt this condition;

absent it, part (i) and part (iii) below are unaffected, and only the monotonicity of the efficiency gap itself is at issue.

(iii) The per-period social cost of provision at efficiency ξ is $(p_e + d)e(\xi)Q$, and the one-time investment cost is $\varphi(\xi)$. Total social cost over the horizon is $C(\xi; T) = \varphi(\xi) + H(T)(p_e + d)e(\xi)Q$, minimized at $\xi^E(T)$. Evaluated at the usage-tax choice $\xi^U(T)$, which is not the minimizer, the excess $\Delta(T) = C(\xi^U(T); T) - C(\xi^E(T); T) \geq 0$. Differentiating and applying the envelope theorem at ξ^E, $d\Delta/dT = (p_e + d)Q\left[e(\xi^U) - e(\xi^E)\right] H'(T) > 0$ since $e(\xi^U) > e(\xi^E)$ by (i) and $H'(T) > 0$. Convexity follows because the bracket $e(\xi^U(T)) - e(\xi^E(T))$ is itself increasing in T by (ii), so $d\Delta/dT$ is increasing. ■

Proof of Proposition E.1

Under the cash flow tax of Proposition 4, the provider's optimum is unchanged and after-tax profit is $(1-z)\Pi$, so incumbent profit falls by $z\Pi$, exactly the revenue raised. Under a usage tax raising the same revenue, incumbents' pre-entry position changes through two channels. Statutory incidence lowers the effective price they receive, but by Proposition 6 the equilibrium markup rises as marginal entrants exit, at rate $\mathrm{d}\ln\mu/\mathrm{d}\ln(1+t_U) = \frac{1}{2}(\mu-1)\sigma_q$. The surviving incumbents' per-unit margin is $(\mu-1)m$, which rises with μ; their retained volume falls, but the margin increase applies to the inframarginal sales they keep. Let $\Phi = 1 + \frac{1}{2}(\mu-1)\sigma_q$ be the buyer-price pass-through. The share of the statutory burden recovered through higher margins is $\Phi - 1 = \frac{1}{2}(\mu-1)\sigma_q > 0$, so incumbents bear strictly less than the full burden, whereas under the cash flow tax they bear all of it. Hence incumbents strictly prefer the usage tax to the revenue-equivalent cash flow tax. For the stronger claim, note that as F rises the number of firms falls, μ rises, and $\Phi - 1$ rises without bound in the relevant range; for Φ large enough the margin gain on retained sales exceeds the statutory burden and surviving incumbents are better off than with no tax. The planner's ranking from Propositions 2–4 is the reverse. ■

B Extensions

This appendix collects extensions referenced in Section 5. None changes the sign of any term in (11), and all move the optimal usage tax in the same direction, which is down.

B.1 A caveat on the form of the tax

The argument so far concerns the tax *base*: energy versus the value of AI output. There is a separate question about the tax *form*, namely whether a levy is specific (a fixed amount per kilowatt-hour or per token) or ad valorem (a percentage of price), and market power bears on it in a direction that partly cuts against our recommendation. Since we have leaned on market power to argue against the usage tax, we owe the other half of the ledger.

The classical result is due to Suits and Musgrave (1953) and, for oligopoly, Delipalla and Keen (1992): for equal revenue, an ad valorem tax on an imperfectly competitive market yields a lower consumer price than a specific tax, and is therefore welfare superior. The comparison is in fact one of the oldest questions in public finance, going back to Cournot (1838) and surveyed by Keen (1998), who traces the line from Cournot's monopoly analysis through Wicksell to the modern competitive statement of Suits and Musgrave (1953). The mechanism is that an ad valorem tax scales with price, so the treasury absorbs a fixed share of any price increase the firm engineers, which blunts the firm's incentive to restrict output. A specific tax has no such property; the firm keeps the full price increase net of a constant per-unit charge. Under perfect competition the two forms are equivalent, but the AI sector is a case where they are not.

This matters because our preferred instrument, a per-kilowatt-hour charge at the data center, is specific, while the usage tax we argue against is ad valorem. On the market-power margin alone, then, the instrument we favor is the less well behaved of the two. The tension is real, but it does not overturn the ranking, for four reasons. First, the two considerations operate on different objects: the base question is about *targeting* the externality, which is worth the order of magnitude documented in Section 6, while the form question is a second-order refinement of pass-through within a given base. A large gain on targeting dominates a smaller loss on form. Second, the Suits–Musgrave advantage of the ad valorem form is largest when it applies to the marked-up good, and the data center's electricity is not the marked-up good; the markup sits downstream, on the token. A charge on the upstream input inherits little of the downstream market power, so the form penalty is muted where we actually propose to levy. Third, the efficiency-margin and location results of Proposition 3 and Corollary 3 have no counterpart for the ad valorem form: an ad valorem tax on tokens still cannot move the carbon intensity of a kilowatt-hour, whatever its pass-through properties. Fourth, and although we do not model electricity use in other sectors, a charge on data center power would in practice be a step toward (or an extension of) a charge on industrial electricity generally. Applied broadly it corrects

the same externality without introducing a wedge between AI and other users of the same kilowatt-hour, which is precisely the inter-sectoral distortion an AI-specific instrument risks creating. Section 6.5 takes up how AI's electricity use interacts with existing rate structures, and the point there reinforces this one: the reforms that matter most are changes to how large loads are charged in general, not a levy carved out for AI.

The clean statement is therefore as follows. If the choice is between a specific and an ad valorem tax *on the same base*, market power favors the ad valorem form (Keen, 1998). But the first-order choice is the base, not the form, and there the externality argument points decisively at energy. Where the two principles can both be honored, they should be: a levy on data center energy can itself be structured as a percentage of the wholesale power price rather than as a flat per-kilowatt-hour charge, capturing the Suits–Musgrave benefit on the small amount of market power present in generation while keeping the base correct. Indeed Myles (1996) shows that a suitable *combination* of a specific and an ad valorem instrument can reproduce the Ramsey price and undo the imperfect-competition welfare loss outright, which is the design-space version of the same point. We flag this as a design refinement rather than build it into the model, since it does not affect any of the magnitudes.

B.2 Training is not sunk once: recurring investment and the size of the rent

Section 3 decomposes the markup into fixed-cost recovery and pure rent, and the discussion of r has so far leaned on a sunk-stock reading: the training runs and clusters already paid for cannot flee, so the quasi-rents they generate can be taxed without discouraging an investment that has already happened. That reading is too convenient, and taken literally it is not credible. It requires that no further training ever be necessary. Frontier providers retrain continuously: each model generation requires a new run, post-training and evaluation are ongoing, and clusters depreciate both physically and through obsolescence as accelerator generations turn over. Investment in facilities and training does not merely dwarf operating cost once; it dwarfs it *every period*. The right frame is therefore ex ante and stationary rather than ex post and one-shot.

Formalize it minimally. Let providers spend I per period on training and capacity to remain at the frontier, and write $\phi \equiv I/(p\,a)$ for recurring investment as a share of revenue. Under a cash flow tax with immediate expensing, I is deducted as incurred, so

the base is the operating surplus net of investment:

$$\text{cash flow base} \;=\; \Big(1-\tfrac{1}{\mu}\Big)p\,a \;-\; I \;=\; \Big[\Big(1-\tfrac{1}{\mu}\Big)-\phi\Big]\,p\,a. \tag{14}$$

Two consequences follow, and they pull in opposite directions for the two halves of the paper.

Two consequences follow, and Appendix B.10 derives both. The rent tax raises less than the sunk-stock reading suggests: at $\mu = 2.5$ the operating surplus is sixty percent of revenue but the taxable residual is $60 - 100\phi$ percent, so at $\phi = 0.3$ a fifty percent cash flow tax raises about 0.23 rather than 0.45 percent of GDP, half again rather than triple the usage tax. What does not change is the ranking, since (14) is precisely the base that preserves neutrality: expensing means the marginal training run is never taxed.

More importantly, recurring investment pins down r, and pins it low. In a stationary equilibrium the fixed cost recovered per unit is a permanent claim on the margin rather than a legacy item, so (5) gives $r = 1 - \phi\mu/(\mu - \bar{\mu})$, and at $\mu = 2.5$, $\bar{\mu} = 1.25$ a recurring investment share of $\phi = 0.2, 0.3, 0.4$ implies $r = 0.6, 0.4, 0.2$. The calibrated range $r \in [0.2, 0.4]$ that Section 6 obtains from an accounting decomposition is therefore reproduced, independently, by the observation that training recurs. Both routes say the frontier markup is mostly the price of staying at the frontier rather than rent, and the limiting case $r \to 0$ is not an artifact but what a stationary industry with permanent retraining requirements looks like.

B.3 Mobile compute and the case for a destination base

The one argument that survives is the one least often made. Compute is mobile. A charge on data center emissions can be avoided by moving the load, and the emissions then occur abroad, often on a dirtier grid.

Proposition B.1 (Leakage and border adjustment)**.** *Let $\ell \in [0, 1]$ be the leakage rate, the fraction of the reduction in domestic data center emissions that reappears abroad in response to a unilateral charge, and suppose the planner values global damages. Then:*

(i) the optimal origin-based charge is $t_E^\star = (1 - \ell)\, d$, short of the full Pigouvian rate;

(ii) a destination-based levy on domestic purchases of AI services, at the ad valorem rate $t_U^{BA} = \ell\, D/p$, restores the missing wedge and cannot be avoided by relocation, since its base is domestic consumption of AI rather than domestic production of compute;

(iii) with both in place, the total wedge equals the full Pigouvian wedge D/p.

This is the strongest available defence of a tax on AI usage, and it is narrow. The tax is optimal not because AI is an input worth discouraging, not because it displaces labor, and not because its providers are rich. It is optimal because it is the only administratively feasible way to price a physical externality that has been shipped offshore. It should therefore be designed as a border adjustment: benchmarked to the embodied energy and grid intensity of the compute, creditable against carbon prices paid abroad, and set to zero for compute served from a jurisdiction with an equivalent charge. A usage tax without these features is not implementing Proposition B.1.

There is a working precedent, and it is instructive in both directions. The European Union's Carbon Border Adjustment Mechanism applies exactly this logic to goods: importers of cement, steel, aluminium, fertiliser, electricity and hydrogen must surrender certificates matching the embedded emissions of what they bring in, with credit for carbon prices already paid in the country of origin. The design principles we list above are the ones CBAM adopts, which is some evidence they are administrable rather than merely coherent. But CBAM also shows what the theory costs in practice. Embedded emissions must be reported, verified, and defaulted when unverifiable; the definitional perimeter is contested; the mechanism took years to phase in and remains disputed by trading partners. Those difficulties arise for physical goods that cross a border on a ship with a bill of lading. An AI query has no such paper trail, and, as Section E.1 argues, attributing embodied energy to a service routed dynamically across data centers is materially harder than attributing it to a tonne of steel. We take the CBAM experience as evidence that a border adjustment on embodied emissions is possible, and as a warning that our version of it is the least implementable recommendation in this paper.

B.4 Training compute and risk

Inference and training are different activities with different externalities, and an instrument levied on tokens taxes almost exclusively the former. Suppose the largest training runs generate an external social risk $R(K)$, convex in the compute K used in a single run, whether from misuse, from concentration, or from failures of control. Let training compute be purchased at price p_K.

Proposition B.2 (Training levy)**.** *The corrective instrument is a Pigouvian levy on training compute at the rate $t_K^\star = R'(K)$, evaluated at the post-tax level of K. A usage tax on inference implements $t_K^\star$ only in the degenerate case in which training compute is used in fixed proportion to inference volume. Moreover, a levy strictly dominates a compute*

threshold: for any threshold $\bar{K}$ that induces the same expected risk, the levy attains it at strictly lower resource cost whenever providers are heterogeneous in the value of scale.

The compute thresholds already written into regulation are a crude version of this instrument. The model says that a price would dominate a threshold, for the usual reason that a threshold is a tax rate of zero followed by a tax rate of infinity, and that the base should be training compute rather than tokens served. Two caveats belong on the record.[B.1]

There is a second and more prosaic reason to separate training from inference, and it belongs with the environmental argument rather than the risk argument. Training consumes electricity and emits carbon, but it produces *no tokens*. The externality calculation of Section 6 is computed per query and therefore covers inference only; the energy burned training the model is nowhere in the base of a token tax. Its implicit tax rate is not low but exactly zero. Estimates of the inference share of AI compute energy vary widely, historical figures ran from about a third to sixty percent, and recent industry estimates put inference at eighty to ninety percent of a deployed model's lifetime energy, so somewhere between one tenth and two fifths of the sector's electricity is invisible to a levy on tokens. A charge at the meter reaches it automatically, because a kilowatt-hour drawn to train a model and a kilowatt-hour drawn to serve a query are billed identically. This is a distinct point from the mis-scaling of Table 1: there the token tax applies the wrong rate to the harm, here it applies no rate at all to a large and growing part of it. Given that training is recurring rather than sunk, as Section B.2 argues, that blind spot is permanent rather than transitional.

B.5 Dynamic efficiency: why the gap grows

Proposition 3 compares the two instruments at a fixed efficiency technology. But the efficiency of AI provision is not fixed; it is the outcome of investment, and the striking fall in energy per token over recent years was a response to the cost of energy, not a windfall. Once efficiency is a durable investment rather than a static choice, the two instruments diverge by more, and the divergence compounds over time. This is the one place where the static formula understates its own conclusion.

Nothing here requires a long horizon. The essential structure is two periods (an ex ante date at which the durable efficiency investment is chosen and an ex post date at

[B.1] Convexity of R is an assumption, not a finding; and if the marginal risk is not convex but has a threshold shape, the ranking can reverse, which is presumably the implicit theory behind the current rules.

which output is served) and the reader who prefers that reading can set $T = 2$ throughout and lose only the comparative static in part (ii) below. We write the model with T periods because the horizon is itself the object of interest: it is what makes the instrument gap grow rather than merely exist. Extend the model to T periods, then. In period one, providers make a durable investment in efficiency ξ at cost $\varphi(\xi)$, which lowers energy per unit $e(\xi)$ in every subsequent period. Energy is purchased period by period at price p_e, and the social cost of a kilowatt-hour is $p_e + d$. A provider facing an instrument that raises its perceived energy price to $p_e + t$ chooses

$$\varphi'(\xi) \;=\; -\left(\textstyle\sum_{s=1}^{T} \rho^s\right)(p_e + t)\, e'(\xi)\, Q, \tag{15}$$

where ρ is the discount factor and Q is per-period output. The bracketed horizon term multiplies the energy price, so a longer horizon raises the investment in efficiency for any given t.

Proposition B.3 (The dynamic instrument gap)**.** *Let $\xi^E(T)$ be the efficiency chosen under an emissions charge $t_E = d$ and $\xi^U(T)$ that chosen under a revenue-equivalent usage tax, which leaves the perceived energy price at p_e. Then:*

(i) $\xi^E(T) > \xi^U(T)$ for all $T \geq 1$: the emissions charge always induces more efficiency investment, because only it raises the perceived price of energy.

(ii) Both are increasing in the horizon, and the gap $\xi^E(T) - \xi^U(T)$ is increasing in T: the longer the assets live, the more the emissions charge pulls ahead.

(iii) The cumulative welfare gap between the instruments, $\Delta(T)$, is increasing and convex in T. A usage tax indexed to match the emissions charge on the scale margin in period one falls progressively further behind on the efficiency margin as the horizon lengthens.

The proof is in Appendix A; the mechanism is entirely in (15). A usage tax is a wedge on output, so it never enters the energy price and never appears in the efficiency condition, at any horizon. An emissions charge enters (15) directly and scales up with the life of the asset. The efficiency margin is where the observed order-of-magnitude gains in energy per token came from, so an instrument that is neutral toward it is not neutral toward the thing that has actually reduced the footprint of AI.

The dynamic result sharpens the policy reading in a way worth stating plainly. A common defense of the token tax is that it is simple and the emissions data are noisy, so the token tax is a reasonable approximation to be tightened later. Proposition B.3 says the

approximation gets worse, not better, with time, because it forgoes an investment response that accumulates. The right sequence is the opposite of the one usually proposed: get the base right first, even crudely, because the base is what steers the durable investment, and refine the rate later.

B.6 Business stealing and the market-power term

Proof of Proposition 7. Consider the symmetric Cournot free-entry equilibrium with inverse demand $P(Q)$, constant marginal cost c, fixed cost F, n firms and per-firm output $x = Q/n$. Let $s \geq 0$ be a marginal output subsidy, equivalently a marginal reduction in t_U, and write $m \equiv P - c + s$ for the margin over subsidized marginal cost. The Cournot first-order condition is $m = -P'(Q)\,x$ and free entry requires $m\,x = F$. Eliminating x between the two,

$$n \;=\; Q\sqrt{\frac{-P'(Q)}{F}}, \qquad m \;=\; \sqrt{-P'(Q)\,F}, \tag{16}$$

so along the free-entry locus both the number of firms and the equilibrium margin are functions of Q alone.

Welfare is $W = \int_0^Q P(q)\,dq - cQ - nF$; the subsidy is a transfer and does not enter directly. Hence

$$\frac{dW}{ds} \;=\; (P-c)\,\frac{dQ}{ds} \;-\; F\,\frac{dn}{ds}.$$

Free entry gives $nF = m\,Q$, so $F\,dn = Q\,dm + m\,dQ$. Substituting, and using $P - c = m - s$,

$$\frac{dW}{ds} \;=\; (m-s)\frac{dQ}{ds} - Q\frac{dm}{ds} - m\frac{dQ}{ds} \;=\; -\,s\,\frac{dQ}{ds} \;-\; Q\,\frac{dm}{ds},$$

which at $s = 0$ is $-Q\,dm/ds$. The naive allocative gain, which is what (11) prices, is $m\,dQ/ds$. Defining χ as one minus their ratio,

$$\chi \;=\; 1 - \frac{-Q\,dm/ds}{m\,dQ/ds} \;=\; 1 + \frac{Q\,dm}{m\,dQ} \;=\; 1 + \frac{\mathrm{d}\ln m}{\mathrm{d}\ln Q},$$

evaluated along (16), which is the statement of the proposition. The economics is that the direct gain from expanding output is exactly cancelled by the fixed cost of the entry it induces, term for term, leaving only the second-order effect that operates through the movement in the equilibrium margin itself.

Part (i), linear demand. With $P = a - bQ$ we have $-P' = b$ constant, so (16) gives $m = \sqrt{bF}$, independent of Q and hence of s. Then $\mathrm{d}\ln m/\mathrm{d}\ln Q = 0$ and $\chi = 1$: a marginal output subsidy has zero first-order welfare effect, every dollar of allocative gain

being consumed by the fixed cost of the marginal entrant. This is Mankiw and Whinston (1986) at the subsidy margin rather than at the entry margin, and it is the sharpest available statement of business stealing.

Part (ii), constant-elasticity demand. With demand elasticity σ_q, the Cournot condition is $m/P = 1/(n\sigma_q)$ and free entry gives $mQ/n = F$, so $n = \sqrt{PQ/(\sigma_q F)}$ and $m = P/(n\sigma_q)$. Therefore $\mathrm{d}\ln m = \mathrm{d}\ln P - \mathrm{d}\ln n = \mathrm{d}\ln P - \frac{1}{2}(\mathrm{d}\ln P + \mathrm{d}\ln Q) = \frac{1}{2}(\mathrm{d}\ln P - \mathrm{d}\ln Q)$. Demand gives $\mathrm{d}\ln P = -\mathrm{d}\ln Q/\sigma_q$, so

$$\frac{\mathrm{d}\ln m}{\mathrm{d}\ln Q} \;=\; -\frac{1}{2}\Big(1+\frac{1}{\sigma_q}\Big) \;=\; -\frac{\sigma_q+1}{2\sigma_q}, \qquad \chi \;=\; 1-\frac{\sigma_q+1}{2\sigma_q},$$

which is increasing in σ_q, equals zero at $\sigma_q = 1$ and tends to $\frac{1}{2}$ as $\sigma_q \to \infty$. The value at $\sigma_q = 1$ is a consistency check rather than a coincidence: with unit-elastic demand revenue is invariant, so by (16) the equilibrium number of firms does not respond and there is no entry to dissipate the gain. At the paper's baseline $\sigma_q = 1.965$, $\chi = 0.246$. □

Proof of Proposition 8. At the margin the buyer's willingness to pay equals the price, so marginal consumer surplus is zero and the entire wedge m accrues to the provider as profit. The social value of a marginal unit of provider profit is $(1-f)g_S$: the share f accruing abroad enters domestic welfare at zero, and the domestic remainder is valued at the planner's weight on shareholder income relative to the income-weighted average. Scaling the allocative gain of Proposition 7 by this factor and expressing the wedge as the rent share r of the excess markup, as (5) requires, gives the stated expression. The four zero cases follow by inspection. □

Remark. The same force is what makes Proposition 6 an ambiguous welfare statement in one direction and an unambiguous positive statement in the other. A usage tax raises the equilibrium markup and reduces n; whether the reduction in n is a welfare loss depends on χ exactly as the subsidy case does. What does not depend on χ is the incidence result of Proposition E.1: incumbents gain from the reduction in n regardless of whether the planner should want it.

B.7 The mobility–identification frontier

Proposition 5 says the pair $(\Delta_\ell, \mathcal{E})$ is not free. The frontier is drawn in the left panel of Figure 1; it is the object a calibration should be disciplined against, and it makes clear which combinations of beliefs are actually available. Three readings follow.

Eliminating κ between $\Delta_\ell = \tanh(\kappa\varrho/4)$ and $\mathcal{E} = \kappa(1-\Delta_\ell^2)$ gives

$$\mathcal{E} \;=\; \frac{4}{\varrho}\left(1-\Delta_\ell^2\right)\operatorname{artanh}(\Delta_\ell),$$

which is single-peaked in Δ_ℓ, equals zero at both $\Delta_\ell = 0$ and $\Delta_\ell = 1$, and shifts down uniformly as the comparative-advantage gap ϱ widens.

First, the feasible set rules out the configuration on which the robot-tax case implicitly rests. That case wants a large exposure gap (a sharply identified group of losers, so that $\hat{g}_1$ has something to attach to) together with slow reallocation, so that the wage effect is not competed away. The frontier says these are compatible only in the high-Δ_ℓ corner, and in that corner the model is the fixed-assignment model, whose empirical premise is that AI exposure is perfectly segregated across ex ante groups. It is not.

Second, the horizon enters through ϱ rather than through κ. Over a quarter, comparative advantage is whatever accumulated human capital made it, and ϱ is large. Over a decade, cohorts choose skills knowing what AI can already do, which compresses ϱ toward zero and takes Δ_ℓ with it. This is a cleaner statement of the transitional argument than the one in terms of a rising mobility parameter: what decays is not the friction but the identifiability of the losing group, and the predistribution term decays with it. The right sunset is the horizon over which ϱ compresses, which depends on cohort turnover and on how fast the exposure frontier itself moves.

Third, two caveats cut in opposite directions and neither is resolved here. Reallocation is not free even when it happens: switching costs, foregone tenure, and human capital specific to the abandoned task are real resource losses that a relative-wage statistic does not capture, so a positive $\mathcal{E}$ understates the welfare loss from displacement even as it damps the measured wage effect. Against that, the exposure gap is measured on *ex ante* groups, and if the planner's actual normative concern is with individuals rather than groups, with whoever happens to hold strong comparative advantage in a routine task, then the two-group structure still overstates how cleanly the burden falls on an identifiable bloc, which is the limitation Thümmel (2023) identifies. What survives both is that the redistributive case for taxing AI usage is weaker, more transitional, and more dependent on an unmeasured sorting pattern than the two-type model implies. We carry Δ_ℓ explicitly from here on, report it as a range, and never rely on the $\Delta_\ell = 1$ corner alone.

B.8 The zeros of Proposition 4 in practice

Neutrality is a property of an idealized cash flow base and no enacted statute has had all of its features. Four channels put a positive wedge back in, and each has a known drafting remedy.

- *Imperfect loss offset.* If losses are carried forward without interest, or expire, the state takes a full share of the upside and less than a full share of the downside. This is a tax on risk-taking and it falls hardest on exactly the frontier investments whose returns are most dispersed. Remedy: carry-forward with interest at the government's borrowing rate, or refundability.
- *An incomplete definition of the base.* Anything not expensed (capitalized software, acquired intangibles, capitalized research under the current U.S. treatment) leaves part of the normal return inside the base, and the tax then reaches marginal as well as inframarginal returns.
- *Entity mobility.* A cash flow tax on a legally resident entity is neutral with respect to real decisions but not with respect to where the entity is resident. The base is stickier than a residence-based profits tax only to the extent that the underlying compute is physically located, which is one more argument for the meter.
- *Time inconsistency.* A tax that is neutral because the capital is sunk is one a rational investor expects to be repeated, and the anticipation raises the required return on new investment. This is the sharpest objection to the sunk-stock reading of the base, and it is why §B.2 prefers the stationary reading.

B.9 Diffusion spillovers

If adoption generates learning spillovers, so that the social return to AI use exceeds the private return by ψ per unit, the corrective term becomes $(D - \psi)/p$. We do not attempt to sign ψ: knowledge spillovers push it up, business stealing pushes it down, and the empirical literature is young. What can be said is the direction of its effect on policy. Any positive ψ lowers the optimal usage tax, and it lowers it against a Pigouvian term that Section 6 shows to be around one percent. It does not take a large spillover to cancel the environmental case for a token tax.

Collecting the extensions, the general expression is

B.10 Recurring investment: derivations

The rent tax raises less than the sunk-stock reading suggests. At $\mu = 2.5$ the operating surplus is sixty percent of revenue, but the taxable residual is $60 - 100\phi$ percent. If recurring investment runs at thirty percent of revenue, the base is halved, and a fifty percent cash flow tax raises about 0.23 rather than 0.45 percent of GDP. Against the usage tax's static yield this is a multiple of roughly one and a half rather than three; against its behavior-adjusted yield, roughly two and a half rather than five. We report the sunk-stock and stationary figures side by side in Table 5 rather than choosing between them, because which is right depends on a fact about the industry (whether frontier training is a phase or a permanent condition) that is not settled. What does *not* change is the ranking: the cash flow tax still raises more than the usage tax with zero first-order distortion, and (14) is precisely the base that preserves neutrality, since expensing means the marginal training run is never taxed.

Recurring investment pins down r, and pins it low. The same parameter disciplines the market-power term, and this is the more important consequence. In a stationary equilibrium the fixed cost recovered per unit is not a legacy item to be amortized away but a permanent claim on the margin, so (5) evaluated in steady state gives

$$r \;=\; 1 - \frac{\phi}{(\mu - \bar{\mu})/\mu},$$

and at $\mu = 2.5$, $\bar{\mu} = 1.25$ a recurring investment share of $\phi = 0.2$, 0.3, 0.4 implies $r = 0.6$, 0.4, 0.2 respectively. The calibrated range $r \in [0.2, 0.4]$ that Section 6 obtains from an accounting decomposition is therefore reproduced, independently, by the observation that training recurs. That is a useful convergence: the parameter on which the sign of the optimal tax turns is disciplined twice over, and both routes say the frontier markup is mostly the price of staying at the frontier rather than rent. The limiting case $r \to 0$ is not the artifact it might appear; it is what a stationary free-entry industry with permanent retraining requirements looks like.

The general lesson is that treating the enormous fixed costs of this sector as sunk once is an ex post convenience that flatters both the rent tax's revenue and the market-power case against a usage tax. Read ex ante and in steady state, the rent is smaller, the revenue prize is smaller, and the case for a usage tax is weaker still, because a larger share of what looks like extractable surplus is in fact the recurring cost of producing the good at all.

C Derivations and interpretation

This appendix develops three results stated compactly in the body.

C.1 The negative wedge and the role of ownership

Does the formula call for a subsidy? It does, and we should say so plainly rather than hedge. Set $\hat{g}_1 = 0$ or $\sigma = 1$, set $D = 0$, and (12) returns $t_U^\star = -(1-f)\, g_S\, r(1 - \bar{\mu}/\mu)(1-\chi) < 0$ whenever $f < 1$, $g_S > 0$, $r > 0$ and $\chi < 1$. This is not an artifact. It is the oldest result in the economics of monopoly: a firm pricing above marginal cost produces too little, and a per-unit subsidy that closes the gap restores the competitive quantity. Cournot (1838) understood this, and any model with a markup and no offsetting distortion will reproduce it. A reader who finds a negative optimal wedge counterintuitive should be reassured that the model is behaving exactly as the textbook says it should.

Three of the four scalings are ways the negative term can be small without being zero, and Section 5.6 argues that at defensible values their product is about one half, so what the formula returns at the baseline is a wedge of about minus five percent rather than the minus fifteen the unscaled expression suggests. A theoretical case for a subsidy of that size is not a policy proposal. The reasons for not recommending one are practical as well as small, and they separate cleanly. Theoretically, the wedge should be negative. Practically, implementing it requires the government to know r (the object Section 6 shows is least observable and on which the sign turns) and to hand public money to the most concentrated firms in the economy on the strength of that estimate, with an obvious incentive for those firms to overstate the rent. Moreover, the subsidy is dominated: competition policy, interoperability, and public compute attack the markup itself rather than paying to offset it, and they do not require the planner to know the rent to a first approximation. So the operational content of the negative term is the modest one, a usage tax should not be *added* on top of a genuine rent in the name of correcting some other margin, but the modesty is a judgment about implementation, not a claim that the formula says something else.

Why ownership matters when the issue is allocative. The scaling by $(1-f)$ invites an obvious objection: if the problem is that a sector with market power produces too little, why should it matter who owns it? The answer is that the term is not measuring the allocative distortion as a global planner would see it. It is measuring the gap between the buyer's price and the *domestic* social marginal cost, and rent enters that gap differently

depending on where it lands. When a domestic firm collects a rent, the payment leaves one domestic pocket and enters another; it is a transfer that cancels in domestic welfare, so the resource cost of a unit of AI is below its price and the good is underprovided from the domestic point of view. When a foreign firm collects the same rent, the payment leaves the country. It is then a genuine domestic resource cost, the domestic social marginal cost rises to meet the price, and the apparent underprovision disappears, not because the world allocation improved, but because the surplus the domestic economy forgoes by buying less is surplus it was sending abroad anyway. A global planner, who counts all rents as transfers wherever they land, faces the $f = 0$ case by construction and always sees underprovision. This is why we insist that the f-dependence be labelled an optimal tariff rather than an efficiency correction: it is a national accounting statement, not a statement about the world allocation.

Two things follow that are easy to conflate. First, none of this argues against *taxing the rent*: a cash flow tax reaches the rent regardless of where it accrues, and Section 5.2 recommends it for domestic and foreign rents alike. The f-scaling applies to the usage tax, which is a bad instrument for reaching rent, and not to the rent tax, which is a good one. Second, the instruments for the markup itself are competition policy and the profits tax, not a transaction wedge in either direction.

C.2 Interpreting the markup: rent, fixed-cost recovery, and timing

Defining r against $(\mu - 1)m$ instead would leave (5) inconsistent with the social marginal cost derived in Appendix A, which nets out $\bar{\mu}$; the two definitions coincide only when $\bar{\mu} = 1$. When $r = 1$ the entire excess markup is rent, the textbook monopoly case. When $r = 0$ the excess markup exactly covers fixed cost and economic profit is zero, which is the free-entry outcome and also the Ramsey case in which the firm marks up only to break even. This distinction is central to everything that follows. A markup that recovers fixed cost is not a distortion the government should want to offset; it is the price of getting the good produced at all, and taxing against it destroys the surplus the fixed investment created.[C.1] Only the rent component $r(\mu - \bar{\mu})$ signals genuine underprovision relative to the social optimum. We therefore carry r explicitly rather than treating the observed

[C.1] The argument is ex ante rather than ex post: once the investment is sunk, taxing its quasi-rents distorts nothing static, and our claim is the forward-looking one about the marginal model or entrant that is then not built. Appendix C.4 sets out both horns of the resulting time-consistency problem (Kydland and Prescott, 1977).

markup as if it were all rent. Section 6 shows that plausible values of r for frontier inference are well below one, which materially changes the sign of the market power term.

C.3 Pass-through, entry, and the evidence on grid cost shifting

This appendix collects material referenced in Sections 2 and 3: the fuller statement of the pass-through discipline, the caveats to the entry-amplification result, and the detail of the ratepayer evidence.

Pass-through. Because the same imperfectly competitive sector must deliver both the targeting result and the entry result, we fix the pass-through of taxes into prices once, from (4), and use it consistently. With constant-elasticity demand and the conduct in (4), the pass-through of a cost shock (such as the emissions charge, which raises m) and of the ad valorem usage tax are, respectively,

$$\rho_E \;\equiv\; \left.\frac{\mathrm{d}\ln p}{\mathrm{d}\ln m}\right|_{n \text{ fixed}} \;=\; 1, \qquad \rho_U \;\equiv\; \left.\frac{\mathrm{d}\ln q}{\mathrm{d}\ln(1+t_U)}\right|_{n \text{ fixed}} \;=\; 1, \tag{17}$$

at fixed n, since a constant-elasticity markup is multiplicative and passes both a proportional cost change and a proportional tax through one for one. The amplification in Proposition 6 comes *entirely* from the entry margin, that is from $\mathrm{d}\ln n < 0$, and not from any difference in short-run pass-through between the two instruments. This is the disciplined version of the claim: at fixed market structure the two taxes pass through identically, so the case against the usage tax rests on entry and on targeting, never on an assumed pass-through asymmetry. We prove (17) in Appendix A and use it in both propositions.

The amplification is large. At $\mu = 2.5$ and $\sigma_q = 2$, $\Phi = 2.5$: a usage tax of one percent raises the buyer price of AI by two and a half percent, because it also drives out entrants and lets the survivors raise their margins. A usage tax on a concentrated sector with heavy fixed costs is not a tax on rent. It is a barrier to entry that the government has erected on the incumbents' behalf, and the incumbents capture part of it.

What survives the obvious caveats is the direction.[C.2] Any model in which fixed costs are large and entry responds to market size delivers $\Phi > 1$, and the static formula therefore *understates* the distortion of a usage tax.

[C.2] Two are worth recording. Free entry can be socially excessive in the sense of Mankiw and Whinston (1986), so a reduction in n is not automatically a welfare loss, and the business-stealing effect works against the amplification result. And the Cournot structure is doing work: with contestable markets or a competitive fringe, $\Phi \to 1$.

The Union of Concerned Scientists identified 130 transmission projects across seven states built only to connect data centers, and found that more than 95 percent of the connection costs, some $4.3 billion in 2024 alone, were spread across all ratepayers rather than charged to the facilities that caused them (Jacobs, 2025). In the capacity market of the largest U.S. grid operator, the independent market monitor attributed 63 percent of a record price increase to data center load, translating to roughly $9.3 billion in a single delivery year that will be recovered from 65 million customers (Monitoring Analytics, 2025). The pass-through is visible on household bills: residential customers in parts of the mid-Atlantic have seen monthly increases of $15 to $25 attributed to capacity prices, and consumer advocates project cumulative increases of over $100 billion across the region by the early 2030s (Natural Resources Defense Council, 2025).

C.4 Modelling choices and caveats

This appendix collects three qualifications to the model of Section 3 that would otherwise sit in long footnotes.

The grid-cost wedge and the pricing benchmark. The wedge is a statement about marginal cost causation, not about the level at which the utility prices. Regulated utilities recover fixed and capacity costs through average-cost (break-even) tariffs rather than at marginal cost, so the relevant benchmark here is not whether price sits above or below marginal cost, which under a break-even constraint it must on average. What γ isolates is that the *incremental* capacity cost caused by the marginal large load is allocated to other customer classes rather than to the load that causes it; that misallocation survives whether one benchmarks against marginal or average cost, and it is the object Borenstein and Bushnell (2022) study when they separate cost recovery from efficiency. In a frictionless long-run competitive equilibrium marginal and average cost coincide and the question would not arise, but cost-of-service ratemaking with large point loads is exactly the case in which they do not. Separately, a large new load can also *lower* average rates when it improves utilization of fixed assets that are already in the rate base, since the fixed costs are then divided among more kilowatt-hours (Borenstein, 2025); a detailed study of the largest data center market found no evidence of a historical cross-subsidy and attributed only about half of a recent capacity price spike to load growth (Riu et al., 2025). Both can be true at once.

Fixed-cost recovery, ex ante and ex post. The argument is ex ante, not ex post. Once the fixed investment is sunk, a tax that drives accounting profit negative is indeed non-distortionary on any static margin, and a planner facing an already-built model could tax its rents, and even its fixed-cost recovery, without discouraging an investment that has already happened. Our claim is the forward-looking one: if providers anticipate that fixed-cost-recovery margins will be taxed away, the marginal frontier model, chip cluster, or entrant is not built, which is the margin Section 5.5 makes endogenous. The two readings are the two horns of the standard time-consistency problem (Kydland and Prescott, 1977): taxing sunk capital is tempting precisely because it looks free, and looks free precisely because the commitment not to do so is what elicited the investment. We take the side of the problem in which the government cannot separately promise not to expropriate fixed-cost recovery, so the distinction between rent and fixed-cost recovery has bite; a reader who assumes full commitment and a one-shot, post-investment game will read the fixed-cost-recovery portion as just another taxable rent, and for that reader the market-power term is correspondingly larger.

Scope of the externality. The externality attaches to the electricity drawn by the AI sector, not to final-good production, and not to labor. We are not assuming AI is the only emitting activity in the economy; we are isolating the emissions associated with compute, which is the object of the policy debate this paper addresses. Economy-wide emissions from other sectors are outside the model and are, by construction, the province of the general carbon price τ_c, whose incompleteness ($\delta > \tau_c$) is what leaves a residual wedge on AI's electricity in the first place. This matters for the instrument comparison below: because the wedge lives on the kilowatt-hour delivered to the data center, the corrective charge t_E of Proposition 3 falls on that node alone and does not touch routine or non-routine labor, whereas a broad carbon price would of course fall on all fuel use in the economy. The two are complements, not substitutes.

D Additional calibration detail

This appendix reports the calibration material summarized in Section 6.

D.1 How much rides on σ?

Every magnitude above depends on the elasticity of substitution between AI and routine labor, and σ is the parameter we know least well, so we report how much of the argument

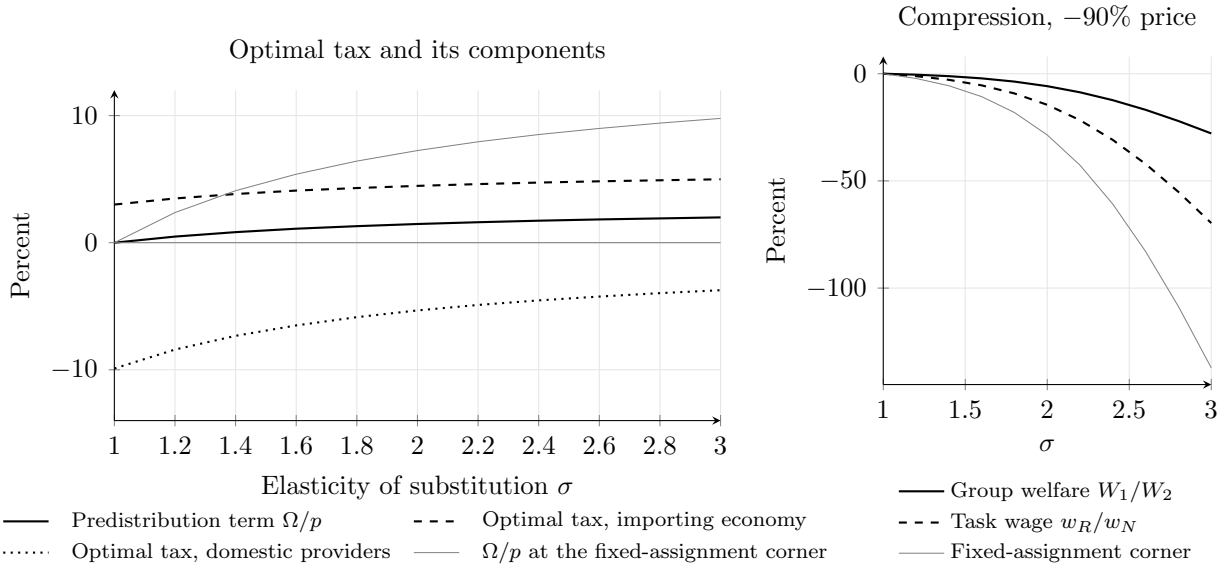


Figure D.1: Sensitivity to the substitution elasticity, at the calibrated sorting parameters. Both panels are drawn at $\Delta_\ell = 0.4$ and $\kappa = 1.5$, hence $\mathcal{E} = 1.26$; the thin grey lines reproduce the fixed-assignment corner $\Delta_\ell = 1$, $\mathcal{E} = 0$ in which the robot-tax literature works, and the gap between the black and grey lines is the whole content of Proposition 5. Left: the predistribution term and the optimal usage tax, for an importing economy ($f = 1$) and for one that owns its providers ($f = 0$), across $\sigma \in [1, 3]$. The sign of the optimal tax is governed by ownership, not by σ, across the entire range, and the predistribution term now stays below two percent everywhere. Right: the compression following a ninety percent fall in the price of AI, the shock of calculation 3, shown for both the relative *task* wage and the relative *ex ante group* welfare index, which differ by the factor Δ_ℓ. The distributional stakes remain highly convex in σ even after both corrections. Other parameters as in Table 4.

survives across its plausible range rather than resting on a single value. The early evidence bounds the range loosely. Controlled deployments find large productivity gains from generative AI concentrated among lower-skilled workers in writing and support tasks (Noy and Zhang, 2023; Brynjolfsson et al., 2025), which is the signature of substitution rather than complementarity and rules out σ near or below one for the exposed tasks, but the studies do not pin down a point estimate, and the tasks AI reaches are still expanding. We therefore report every result across $\sigma \in [1.2, 3]$ and let the reader locate their own prior within it.

Figure D.1 carries three messages, and all three are robust to where one stands on σ. First, the predistribution term is bounded, and tightly. It rises with σ but flattens quickly, peaking below two percent even at $\sigma = 3$ with our generous welfare weight; the grey line shows that even at the fixed-assignment corner it never leaves single digits. The robot tax logic is real but it is not a large number for any defensible elasticity, and it is an order of magnitude smaller than the market-power term across the whole range. Second, the sign of the optimal usage tax is set by ownership rather than by technology, but both curves

are now close to zero across the whole range. For an importing economy the optimal tax runs from three to five percent, positive and mildly increasing in σ; for an economy that owns its providers it is negative throughout but never by more than ten percentage points, and it *rises* toward zero as σ grows, because a higher substitution elasticity raises both the predistribution term and, through σ_q, the share of the market-power correction that entry dissipates. The figure draws the domestic case at the all-rent baseline $r = 1$; at the lower end of our preferred range for r the domestic optimum is pushed close to zero rather than deeply negative, as the corresponding row of Table 4 shows, so the robust statement is that domestic ownership removes the case for a positive usage tax rather than that it always calls for a large negative one. No value of σ reconciles the importing and domestic cases, which is the formal statement of the tariff-versus-efficiency tension. Third, and this is the uncomfortable one, the distributional *stakes* are far more sensitive to σ than the optimal *tax* is. The right panel shows that the compression from cheap AI runs from well under one percent at $\sigma = 1.2$ to twenty-eight percent at $\sigma = 3$ measured on group welfare, and roughly two and a half times that measured on the relative task wage. The tax that the model recommends barely moves across that range, while the problem the tax is meant to address explodes. That divergence is a property of the problem rather than of the formula: if σ is high, the appropriate response is a large adjustment to the income tax and transfer system, not a marginal usage tax that works through the same weak channel as the price movement it is trying to offset. The usage tax is close to irrelevant to the distributional question at exactly the elasticities where that question becomes urgent.

The grey reference lines in both panels are worth dwelling on, because they are the whole of Proposition 5 in one picture. Moving from the fixed-assignment corner to the calibrated sorting parameters multiplies every curve by $\Delta_\ell(1+\varepsilon_L)/(1+\varepsilon_L+\mathcal{E}) \approx 0.20$, and roughly two thirds of that reduction is the exposure gap rather than reallocation. Neither of the three messages changes, the term stays bounded, ownership still sets the sign, and the stakes remain far more convex in σ than the tax is, but every magnitude is smaller, and the third message is the stronger for it. If σ is high, so that displacement is severe, the mobility margin is precisely what a planner should be trying to strengthen, and doing so mechanically shrinks the case for the usage tax.

D.2 Calculation 3: cheap AI and the wage distribution

3. What cheap AI does to the wage distribution. Now run the model in the other direction, which is where the anxiety actually comes from. Suppose the price of AI falls by ninety percent at constant capability, roughly what has happened over the past two

years. With CES and $\sigma = 2$, the odds ratio of the cost shares satisfies $s/(1-s) \propto q^{-1}$, so a tenfold fall in q multiplies the odds by ten. Starting from $s = 0.05$, the AI share within automatable tasks rises to

$$s' \;=\; \frac{10 \times 0.0526}{1 + 10 \times 0.0526} \;\approx\; 0.35.$$

Exposure rises from 1.5 to 10.5 percent of GDP, and by Corollary 1 the labor share falls from 98.5 to 89.5 percent. The routine wage falls relative to the non-routine wage by $\ln\big[(1-s')/(1-s)\big]/(1+\varepsilon_L+\mathcal{E})$, which is -29% at $\sigma = 2$ at the fixed-assignment corner but about -15% at $\mathcal{E} = 1.26$, and the gap in ex ante group welfare is Δ_ℓ times that, about -6%, though as Figure D.1 showed this figure is highly convex in σ and runs from a mild 2% at $\sigma = 1.2$ to well over 100% at $\sigma = 3$.

This number should be read as an illustrative upper bound, not a forecast, and the reason is a limitation of the two-type structure. Thümmel (2023) shows that two-group automation models systematically mislead about distributional effects, because the real action is within groups and in the reallocation of workers across a continuum of tasks, which a routine/non-routine split cannot represent. Our two-type model collapses that continuum and so delivers the compression as a single relative-wage number, which overstates how cleanly the burden falls on an identifiable group. Lemma 3 makes the same point from the tax side: a shock this large reshapes the earnings distribution in ways a linear income tax cannot neutralize by a proportional adjustment, so the clean separation of common and relative components on which the predistribution term rests is itself strained at these magnitudes. Both caveats point the same way. The calculation establishes that the distributional stakes *can* be large, not that they fall as a tidy transfer between two blocs, and certainly not that a usage tax is the instrument to address them. If a shock of this size arrives, the instrument that meets it is a genuinely nonlinear income tax and the transfer system, not the price of a token, which works through the same weak channel as the price movement it is trying to resist and is an order of magnitude too small to offset it.

D.3 Calculation 5: the welfare cost of the wrong instrument

5. The welfare cost of choosing the wrong instrument. It is worth putting one number on the central recommendation, since the paper otherwise trades in ratios. Suppose a government wishes to raise revenue equal to 0.05% of GDP from the AI sector, a modest target. Levied as a usage tax, the required rate is about 3.3% on a base of 1.5%

of GDP, and by Proposition 6 the quantity distortion is amplified by $\Phi = 2.5$ through the entry margin. The resulting deadweight loss, of Harberger form $\frac{1}{2}\,\Phi\,\sigma_q\,pa\,t_U^2$, is about 0.004% of GDP, or roughly eight cents of deadweight per dollar raised. Levied instead as a cash flow tax on rents, Proposition 4 makes the deadweight loss zero to first order: the same revenue is raised with essentially no distortion, or on the pessimistic residual wedge of §7.3 about two cents on the dollar rather than eight. And levied as an emissions charge, the "cost" is negative, because the instrument corrects a pre-existing externality worth far more per unit than the revenue it raises; the first \$0.05 per kWh of the charge buys a social gain, not a loss. The ranking is therefore not close. Against the rent tax the usage tax loses eight cents on the dollar; against the emissions charge it forgoes a first-order welfare gain to incur a deadweight loss. These are small numbers in levels only because the AI sector is still small; they scale with exposure θ, which every forecast has rising several-fold within the decade, so the cost of entrenching the wrong instrument now grows with the sector.

Sensitivity to σ. The substitution elasticity is the parameter we know least well, and Appendix D.1 reports every result across $\sigma \in [1.2, 3]$. Three findings there are worth carrying forward. The predistribution term is bounded: it rises with σ but never leaves single digits anywhere in the range. The sign of the optimal usage tax is set by ownership rather than by technology, at every point in the range. And the distributional stakes are far more sensitive to σ than the optimal tax is, the wage compression from cheap AI runs from a tolerable two percent at $\sigma = 1.2$ to well over one hundred percent at $\sigma = 3$, while the tax the model recommends barely moves. That last point is the uncomfortable one: the usage tax is close to irrelevant to the distributional question at exactly the elasticities where that question becomes urgent, which is the model telling us that a large σ calls for the income tax and transfer system rather than a marginal wedge on compute.

D.4 The probabilistic calibration in full

Independence across parameters is itself an assumption, and not an innocent one: a world with high markups is plausibly also a world in which more of the markup is genuine rent, which would fatten the lower tail. The right panel of Table D.1 therefore reimposes the same marginals through a Gaussian copula with $\rho(\mu, \phi) = -0.4$, $\rho(\sigma, \hat{g}_1) = 0.3$, $\rho(d, \text{energy share}) = 0.3$ and $\rho(\kappa, \Delta_\ell) = -0.4$. The effect is exactly what the objection predicts and smaller than one might fear. The fifth percentile of the domestic optimum falls from -21.1% to -22.4%, so the tail does fatten; the median moves by half a per-

Table D.1: Probabilistic calibration: the distribution of $t_U^\star$

	Independent priors			Correlated priors		
	Median	5th	95th	Median	5th	95th
Predistribution term	+1.3%	+0.4%	+3.9%	+1.3%	+0.3%	+4.6%
Net externality term	+1.4%	−1.1%	+4.6%	+1.4%	−1.2%	+5.0%
Market power term	−9.2%	−17.6%	−2.4%	−9.0%	−18.4%	−2.1%
$t_U^\star$, domestic (f=0)	**−6.2**%	−15.2%	+1.6%	**−5.9**%	−15.9%	+2.2%
$t_U^\star$, importing (f=1)	**+2.9**%	+0.1%	+6.8%	**+3.1**%	−0.1%	+7.5%
$\Pr(t_U^\star < 0 \mid f{=}0)$		0.900			0.888	
$\Pr(\lvert t_U^\star\rvert < 5\% \mid f{=}0)$		0.403			0.398	
$\Pr(t_U^\star > 10\% \mid f{=}0)$		0.000			0.000	
$\Pr(t_U^\star < 0 \mid f{=}1)$		0.046			0.056	
$\Pr\big(0 < (1) + (2) < 10\%\big)$		0.951			0.936	

Notes. Uniform marginals over the ranges in Table D.3; 200,000 draws; $\tau = 0.35$, $\bar{\mu} = 1.25$, $s = 0.05$, $\theta = 0.015$ fixed. Every draw sets $\mathcal{E} = \kappa(1 - \Delta_\ell^2)$, so the mobility elasticity and the exposure gap are never combined in a configuration Proposition 5 rules out. The correlated columns impose a Gaussian copula with $\rho(\mu, r) = 0.6$, $\rho(\sigma, \hat{g}_1) = 0.3$, $\rho(d, \text{energy share}) = 0.3$ and $\rho(\kappa, \Delta_\ell) = -0.4$, leaving the marginals unchanged so the two panels are directly comparable. Replication code is available from the authors.

centage point; and $\Pr(t_U^\star < 0)$ falls slightly, from 0.78 to 0.75. The two forces largely offset, which is a mild but genuine robustness result: the conclusion is not an artifact of the independence assumption.

Four probabilities summarize the exercise, and they are the numbers we would ask a skeptical reader to engage with rather than any single column of Table 4. All are reported for the independent case, with the correlated value in parentheses where it differs.

- $\Pr(t_U^\star < 0) = 0.90$ (0.89) for an economy that owns its providers, and 0.05 (0.06) for one that imports. Ownership still determines the sign, though not overwhelmingly, because the market-power term is weighted by the social value of shareholder income and netted of entry dissipation. The sharper statement is about magnitude rather than sign: $\Pr(t_U^\star > 5\%) = 0.006$ and $\Pr(t_U^\star > 10\%)$ is numerically zero. *No draw supports a usage tax at the rates proposed*, and that holds whatever one concludes about the sign. *Ownership, not technology, determines the sign*, and it does so across essentially the whole parameter space rather than at a particular calibration.
- Pr(market power term exceeds the net externality term) = 0.96 (0.95). The horse race that motivates the paper is not close: on these priors the environmental case for a usage tax is dominated by the market-power case against it in nineteen draws

out of twenty.

- Pr(predistribution term $< 5\%$) $= 0.98$ (0.96), with a median of 1.3 percent once the exposure gap and the reallocation elasticity are allowed to vary jointly over their full range. The robot-tax logic is real, bounded, and roughly a fifth of what the fixed-assignment corner suggests.
- Pr($d/p_e > 50\%$) $= 0.75$. The claim that the wedge on the *kilowatt-hour* is large survives the full range of the contested grid-shifting parameter, for the reason Figure D.2 gives: carbon, not cost shifting, dominates the level.

The exercise also disciplines our own claims. The domestic optimum is negative in most of the space but by no means all of it, and three of this paper's own corrections are what opened it up. Three restrictions do most of the work in locating the distribution: the shareholder weight of Lemma 4, the entry dissipation of Proposition 7, and the tie between the rent share and recurring investment in Corollary 5. Each follows from taking the model's own assumptions seriously rather than from new evidence, and they do not all push the same way: the first raises the market-power term while the second lowers it. What no draw supports remains the same thing, and it is now the paper's sole quantitative claim about the level: a usage tax at the rates actually proposed. Pr($t_U^\star > 5\% \mid f{=}0$) $=$ 0.01 and Pr($t_U^\star > 10\%$) is numerically zero. What no draw supports is a usage tax at the rates actually proposed. That is the robust finding, and it does not depend on where one stands on σ.

D.5 The marginal value of public funds

Revenue and deadweight loss are not directly comparable across instruments whose burdens fall on different people. The marginal value of public funds of Hendren and Sprung-Keyser (2020) provides a common denominator: the willingness to pay of those who bear a reform, divided by its net cost to the government. For revenue-raising instruments a *lower* MVPF means a cheaper way to raise a dollar. Table D.2 reports it for each instrument in our framework.

The ordering is stable and it is the paper's recommendation restated in a metric a finance ministry already uses. The emissions charge is the cheapest way to raise a dollar from this sector under either ownership structure, because part of what looks like a burden is a correction. The cash flow tax is next, and it is free to an importing economy, which is the terms-of-trade point of Corollary 2 in MVPF form. The usage tax is the

Table D.2: Marginal value of public funds, by instrument

Instrument	Domestic providers (f=0)	Importing economy (f=1)
Data center emissions charge, long run ($\eta_E \approx 1$)	0.38–0.69	0.38–0.69
Data center emissions charge, short run ($\eta_E = 0.1$)	0.94–0.97	0.94–0.97
Cash flow tax on rents	1.00	0.00
Usage tax at 3.3%	1.18	2.93
Usage tax at 10%	1.79	4.43

Notes. Lower is better for a revenue-raiser. The cash flow tax is treated as a pure transfer with no behavioral response, so its MVPF is the domestic ownership share $1 - f$: it costs domestic households a dollar per dollar raised when providers are domestic, and nothing when they are foreign. That entry inherits the idealization of Proposition 4; on the pessimistic residual wedge of §7.3 it rises to about 1.02 for $f = 0$, which changes no ordering in the table. The usage tax MVPF is $[\Phi - (\Phi - 1)(1 - f)]/[1 - \Phi\sigma_q t/(1 + t)]$, where the numerator nets out the share of the buyer-price increase that accrues to domestic incumbents through the entry channel and the denominator is the fiscal externality. The emissions charge MVPF is $1 - (d/p_e)\,\eta_E\,(1 - \ell)$: part of the burden is offset by a corrective benefit, but *only* for emissions actually avoided rather than relocated, so the relocating share ℓ is netted out. Ranges span $\ell \in [0, 0.5]$; η_E from Burke and Abayasekara (2018). All entries at the baseline calibration.

most expensive under both structures, and its cost rises with the rate because the entry channel compounds.

Two qualifications keep the comparison honest, and both cut against our preferred instrument. First, the emissions charge's advantage is a long-run advantage. At the short-run elasticity its MVPF is 0.94–0.97, barely below the rent tax and far from the headline number; the gap opens only as the efficiency and siting margins operate, which is the same horizon logic as Proposition B.3 and one more reason the dynamic result matters. Second, the corrective benefit must be netted of relocation, and the evidence is pointed here: Burke and Abayasekara (2018) find the long-run industrial elasticity is the largest of any sector precisely because electricity-intensive activity clusters in low-price states. Read through our model, a substantial part of what looks like abatement in the elasticity literature is the leakage ℓ of Proposition B.1, and it delivers no global emissions benefit at all. This is a genuine limitation of a unilateral charge rather than a modelling convenience, and it is precisely why the destination-based backstop is in the paper. One asymmetry is worth recording.[D.1]

[D.1] The usage tax looks *worse* for an importing economy on this metric, not better, because the buyer price rises by Φ while the incumbent margin gain leaves the country. The tariff argument of Corollary 2 operates through the social marginal cost of the rent, not through the MVPF of the levy, and the two should not be conflated.

D.6 Parameter provenance and the electricity wedge

This appendix collects the provenance table referenced in Section D.6 and the robustness of the electricity wedge to the contested grid-shifting parameter γ.

Table D.3: Parameter provenance

Par.	Value / range	Evidence	Design	External-validity caveat
σ	2; $[1.2, 3]$	Gains concentrated among lower-skilled workers in writing, support (Noy and Zhang, 2023; Brynjolfsson et al., 2025)	RCT; firm rollout	Task gains are not a production elasticity; rules out $\sigma \leq 1$ but pins no point estimate.
$\hat{g}_1$	0.3; $[0.1, 0.5]$	Normative	—	A social preference, not an estimate; reported across its range.
ε_L	0.3; $[0.1, 0.5]$	Standard range in public finance	Quasi-experimental	Intensive-margin estimates; the margin here may be occupational switching.
κ	1.5; $[0, 3]$	Roy–Fréchet dispersion of comparative advantage (Hsieh et al., 2019; Galle et al., 2023)	Structural Roy model	Estimated for occupational and sectoral sorting, not for AI exposure. Not a mobility parameter: it sets Δ_ℓ and $\mathcal{E}$ jointly through Proposition 5.
Δ_ℓ	0.4; measured $[-0.50, +0.45]$	*Computed here* (Table 2, Appendix D.7) from O*NET 30.3 education distributions (National Center for O*NET Development, 2026) and seven public exposure measures (Webb, 2020; Felten et al., 2021; Eloundou et al., 2024; Eisfeldt et al., 2023), employment-weighted with Dorn (2009) labor-supply weights	Descriptive; occupation-level	Set to one by assumption throughout the robot-tax literature; that corner is empirically false and quintuples the term. The carried 0.4 is the value on the single most favourable measure; all four direct AI measures give a *negative* gap (Remark 2). Occupation-level education distributions substitute for ACS microdata, so the within-occupation grouping is imputed.
χ	0.25	*Derived*, not calibrated: $1 - (\sigma_q + 1)/2\sigma_q$ under this paper's CES demand (Mankiw and Whinston, 1986)	Theory	Demand-system dependent. Equals 1 under linear demand, which is reported as a bound. Not a free parameter.
g_S	0.65; $[0.45, 0.85]$	Implied by $\hat{g}_1 = 0.3$ under isoelastic weights, given the earnings position of routine occupations and the concentration of equity ownership (Lemma 4)	Calibration	Depends on the isoelastic form and on log-normality of earnings; range spans $s \in [0.7, 1.0]$, $z_R \in [-0.5, -0.25]$, $z_S \in [2.0, 3.0]$. Scales the market-power term.

Table D.3 (continued)

Par.	Value / range	Evidence	Design	External-validity caveat
ϕ	$[0.2, 0.4]$	Recurring training and capacity investment as a share of revenue	Accounting	Not separately observed; disciplines r in steady state (§B.2) and shrinks the cash-flow base.
μ	2.5; $[2, 3]$	Reported gross margins on frontier inference	Company disclosure	Gross margin over *short-run* marginal cost, not an estimated markup.
$\bar{\mu}$	1.25	Economy-wide markup (De Loecker et al., 2020)	Production-function estimation	Contested; sensitive to cost-share assumptions.
r	$[0.2, 0.4]$	Constructed from fixed-cost recovery, §6	Engineering, accounting	Not observed. The parameter the sign turns on, hence a range.
ζ	0.4 kg/kWh	Grid intensity serving U.S. data centers (International Energy Agency, 2025)	Measurement	Varies by an order of magnitude across sites; a fleet average hides the siting margin.
δ	\$190/t	Social cost of carbon (U.S. Environmental Protection Agency, 2023)	Integrated assessment	Generous: above most market prices and above Nordhaus (2017).
γ	$[0, 0.03]$ /kWh	Socialized capacity and transmission cost (Jacobs, 2025; Monitoring Analytics, 2025) vs. (Energy and Environmental Economics (E3), 2026; Riu et al., 2025)	Cost allocation	*Genuinely disputed.* Figure D.2: the level conclusion survives the range; §6.5 is where it bites.
e	0.3–150 Wh	Energy per query by workload (International Energy Agency, 2026; Luccioni et al., 2024; de Vries, 2023)	Benchmark	Workload mix is not publicly observed; see notes to Table 1.
θ	1.5%	AI spending as a share of output	National accounts	Generous; rising fast, so a statutory rate dates quickly.
η_E	0.1 SR; ≈ 1 LR	Price elasticity of U.S. electricity demand (Burke and Abayasekara, 2018); meta-analysis (Labandeira et al., 2017)	IV panel, state $\times$ sector $\times$ year	Not estimated for data centers. Much of the large long-run industrial elasticity reflects clustering in low-price states, leakage ℓ here, not abatement.
ℓ	**unmeasured**	—	—	**No credible estimate.** Proposition B.1, the one surviving case for a usage tax, rests on it.
ψ	**unsigned**	—	—	**Not signed by existing work.** Spillovers push up, business stealing down; any $\psi > 0$ lowers $t^{\star}_{U}$.

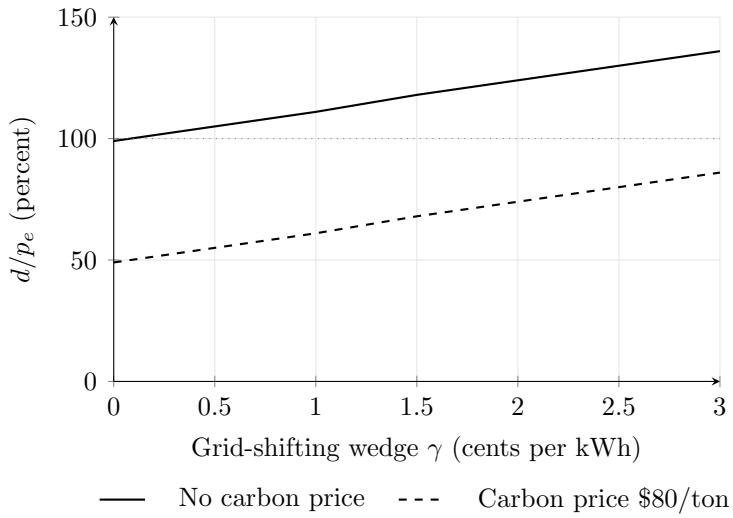


Figure D.2: The electricity wedge is large for any γ. The residual external cost per kilowatt-hour, as a share of the price a data center pays, plotted against the socialized grid-cost wedge γ, with and without a pre-existing carbon price. Even at $\gamma = 0$ the wedge is roughly half to all of the electricity price, because unpriced carbon dominates. The contested magnitude of γ moves the wedge only modestly; its real work is distributional, not in the level. Other parameters as in Table 4.

D.7 The exposure gap: estimand and construction

Because Δ_ℓ is new to this literature and is the input on which the predistribution magnitude chiefly depends, we state it precisely enough to be computed and disputed.

Estimand. Fix a partition of the employed workforce into two ex ante groups and a binary classification of occupations into exposed and unexposed. Then

$$\Delta_\ell \;=\; \ell_{1R} - \ell_{2R}, \qquad \ell_{iR} \;=\; \frac{\sum_{n \in i} \omega_n \, \mathbf{1}\{o(n) \in R\}}{\sum_{n \in i} \omega_n},$$

with ω_n the survey weight and $o(n)$ the occupation of worker n. Three features are worth stating because each is a place the literature's convention hides an assumption. First, $\Delta_\ell \in [-1, 1]$: it is signed, and Remark 2 explains why the sign is not guaranteed. Second, it is a difference of shares rather than a regression coefficient, so it has no residual variance and needs no identifying assumption beyond the classification itself; the only sampling uncertainty is that of two proportions. Third, it is defined on *ex ante* groups, which by Assumption 1 are unobservable, so any implementation replaces them with an observable proxy and Δ_ℓ is therefore proxy-dependent. That is a feature: the planner's instrument set is also proxy-dependent, and the relevant Δ_ℓ is the one attached to whatever grouping the planner's welfare weights actually track.

Construction. One year of American Community Survey microdata, restricted to employed workers aged 25 to 64, merged by SOC code to an occupation-level exposure score, with "exposed" defined as a score above an employment-weighted quantile. Three choices have to be reported rather than fixed:

- *Exposure measure.* At minimum the routine-task intensity of Autor and Dorn (2013), with the conventional cut at the 66th percentile, and the LLM exposure score of Eloundou et al. (2024), cut at the median. The first matches the object σ describes; the second matches what the technology currently does. They disagree in sign, which is the substantive finding rather than a nuisance.
- *Ex ante grouping.* Less than a bachelor's degree versus the rest; high school or less versus the rest; below versus above median earnings; and a young non-college interaction. The paper's argument requires only that the maximum of $|\Delta_\ell|$ across these be well below one.
- *Threshold.* The share of employment classified as exposed. Reporting Δ_ℓ across a range of thresholds is cheap and is the natural robustness check, since a measure that generates a large gap only at an extreme cut is not describing segregation.

Implementation. Table 2 implements the occupation-level counterpart of the estimand above. Absent worker microdata, group membership enters as a probability rather than an indicator: writing E_o for employment in occupation o and $p_o = \Pr(\text{group } 1 \mid o)$,

$$\ell_{1R} = \frac{\sum_o E_o\, p_o\, \mathbf{1}\{o \in R\}}{\sum_o E_o\, p_o}, \qquad \ell_{2R} = \frac{\sum_o E_o\, (1 - p_o)\, \mathbf{1}\{o \in R\}}{\sum_o E_o\, (1 - p_o)}.$$

This coincides with the microdata estimand whenever exposure is assigned at the occupation level, which it is for every published measure, and it differs from it only through the accuracy of p_o. We take p_o from the incumbent-reported Required Level of Education distribution in O*NET 30.3 (National Center for O*NET Development, 2026), averaged from eight-digit to six-digit SOC and renormalized to sum to one across the twelve education categories; "less than BA" is categories one to five, "HS or less" categories one and two, "less than associate" categories one to four. The fourth grouping uses the O*NET job zone, a preparation rather than an education measure, and is included because it is the closest available substitute for the earnings grouping that microdata would permit. Employment weights are the Dorn (2009) `occ1990dd` labor-supply weights from the replication file of Webb (2020), allocated to detailed SOC occupations: exactly at the level of

the SOC major group, and within a major group in proportion to a title match between the harmonized 1990 census categories and SOC titles. The resulting panel covers 765 detailed occupations and 152 million workers, with a major-group employment distribution close to the published national one. The Felten–Raj–Seamans and Eisfeldt et al. scores are carried from SOC 2010 to SOC 2018 with the BLS crosswalk (U.S. Bureau of Labor Statistics, 2018); the Eloundou et al. scores are averaged from O*NET-SOC to six-digit SOC.

Robustness. Three checks. Dropping the employment weights entirely and treating each detailed occupation equally leaves the robot gap at +0.44 against +0.41 and moves every AI measure further negative, to between −0.49 and −0.61. The weighting is not what produces the sign pattern. We should say why we weight at all, since the choice matters elsewhere. Detailed occupations vary enormously in size and the boundaries between them are partly conventional, so an unweighted average answers a question about the classification and not about the workforce. Tomlinson et al. (2025) find the same choice mattering in the other direction. Their measure of what AI is actually used for correlates with occupational wages at +0.13 weighted and +0.17 to +0.32 unweighted, and they attribute part of the strong positive wage gradient in earlier work to the absence of weighting. Our sign pattern survives either way, which is why we report both. Varying the exposure threshold compresses every gap toward zero except the already-small Webb software gap, which is flat; Figure 5 plots the full sweep. And the four groupings never disagree in sign on any measure, with a maximum spread of 0.09 within a row, so the result is not an artifact of where the education line is drawn.

Limitations. We see three, and each is a reason to treat the numbers as an upper bound on $|\Delta_\ell|$ and not a point estimate. First, p_o is the education distribution that O*NET incumbents report as *required*, which is not the distribution actually held. It is probably more concentrated than the truth, and concentration makes groups look better sorted across occupations than they are. Second, the allocation of employment within a SOC major group rests on a title match and not on a validated crosswalk, though the unweighted check above shows that the sign pattern does not depend on it. Third, we report no standard errors. The sampling variance of the estimand is that of two proportions in microdata we do not have. The dispersion across the seven measures is an order of magnitude larger than any plausible sampling error, and that is the uncertainty that matters here.

Exposure is not adoption. A fourth issue is conceptual, and it applies to every measure in Table 2 and not to our construction of them. The object σ describes is AI substituting for labor in production. The seven measures describe the scope for it to do so. Bick et al. (2026) compare the two using a nationally representative survey that asks workers which of their own tasks they use generative AI for, and the gap is large. Exposure scores correlate positively with realized adoption, but they account for at most about half of the variation across workers, and occupation-level adoption rates more than double the explanatory power of the leading exposure measures in predicting whether a given worker uses the technology. Much of the residual is individual and not occupational: within narrowly defined tasks adoption is far from universal, and worker fixed effects absorb more variation than task characteristics do.

We expect this to reinforce the pattern in Table 2 and not to overturn it. Their high-adoption occupations are computing, business, finance and management, and their low-adoption occupations are personal service, transport and manual work. This is the gradient that produces our negative gaps, so recomputing Δ_ℓ on adoption should give something at least as negative as the direct AI measures give. We say expect because the occupation-level adoption index is not yet public. It is promised in the paper's replication package, and we would run it against this table when it appears.

The inputs warrant one further caution. Yin et al. (2026) replicate the Eloundou et al. (2024) rating procedure across models and find the resulting exposure scores less stable than their widespread use assumes. Two of our seven rows come from that procedure. We therefore read Table 2 as a statement about the spread across measures, not about any single one.

What would change our conclusions. A value of Δ_ℓ above roughly 0.8 on the grouping a planner would plausibly weight would restore the fixed-assignment magnitudes and, with a low rent share, restore a small positive optimal usage tax for a domestic economy. Table 2 contains no value above 0.45, and none above 0.51 in absolute value, on any measure, grouping or threshold we examined. The falsification is attempted and fails. A value near zero or below removes the predistribution term entirely or reverses it, and that is what the four direct AI measures deliver. The case in the text is the intermediate one, and it is the one favourable to the instrument.

D.8 The CES correction to the exposure cancellation

Lemma D.1 (CES correction to the exposure cancellation)**.** *Replace the Cobb–Douglas aggregator in* (1) *with a constant-elasticity-of-substitution aggregator,* $Y = \Lambda\big[\lambda_0 X_A^{(\eta-1)/\eta} + (1-\lambda_0) X_M^{(\eta-1)/\eta}\big]^{\eta/(\eta-1)}$*, in which* η *is the elasticity of substitution between the automatable and non-automatable bundles and* $\eta = 1$ *recovers the Cobb–Douglas case of the text. Then the predistribution term of Proposition 2 is multiplied by the factor* $1 + (\eta - 1)\,\phi$*, where* ϕ *is the equilibrium response of the automatable value share to the AI price and satisfies* $\phi = O(\theta)$*. At the observed exposure* $\theta < 0.02$ *the correction is below two percent of the term for any* $\eta \in [0.5, 2]$*, so the Cobb–Douglas cancellation of* θ *is exact to first order and accurate to second.*

Lemma D.1 disarms the obvious objection that the clean cancellation of exposure is an artifact of the unit elasticity. It is not: the cancellation survives to first order for any outer elasticity, and the correction is bounded by exposure, which is small. We use Cobb–Douglas in the body for readability and rely on the lemma for generality.

E Implementation and the policy debate

This appendix collects the implementation detail behind Section 7.

E.1 Administrability of each instrument

Optimal tax formulas assume the base is observable. Each instrument here faces a different administrative constraint, and in one case the constraint is severe enough to qualify a result we have otherwise defended. We follow Keen and Slemrod (2017) in treating administrability as part of instrument choice rather than as an afterthought.

The emissions charge is the easy case. The base is metered by the utility, at the point of delivery, for billing purposes that already exist. Tons of CO_2 follow from kilowatt-hours and a grid intensity factor that system operators already publish. There is no new reporting entity, no new valuation question, and no transfer-pricing margin, because electricity delivered to a physical site cannot be booked in another jurisdiction. This is close to the ideal case for a Pigouvian levy and it is one more reason to prefer the meter.

The rent tax is administratively demanding but familiar. Neutrality requires a cash flow base with expensing and full loss offset, which means the tax authority must

distinguish economic rent from the normal return, the same problem faced by every cash flow tax proposal (Auerbach and Devereux, 2018). The known failure mode is political rather than technical: what gets legislated is usually a levy on gross revenue rather than on cash flow, with the consequences §5.2 describes. The administrative recommendation is therefore specific and checkable. Expensing, loss carry-forward with interest, and a base defined on cash flow rather than turnover.

The usage tax is easy to administer and that is part of the problem. Providers already meter tokens and issue invoices, so the base is available at near-zero cost. Administrative convenience is precisely why the instrument is attractive to legislatures and precisely what the rest of this paper argues against; the point is worth making explicitly, because in a Keen–Slemrod framework a low enforcement cost is a genuine advantage that must be weighed, not dismissed. Our claim is that it does not come close to offsetting a base that is wrong by an order of magnitude on the harm-to-value ratio.

The border adjustment is the hard case, and we do not want to overclaim it. Proposition B.1 is the one surviving argument for a usage tax, and it is also the least administrable thing in the paper. Implementing it requires attributing embodied energy and grid intensity to an imported AI service, which means knowing which data center served a given query and on what grid, information that is not on the invoice, that providers route dynamically for reasons of latency and capacity, and that no customs authority currently has any means of verifying. Three fallbacks exist, in descending order of fidelity: provider-level attestation of fleet-average energy and carbon intensity, audited as financial statements are; a default benchmark rate applied absent attestation, set conservatively so that attestation is the preferred option; or a rebuttable presumption keyed to the provider's jurisdiction of primary compute. Each is cruder than the emissions charge it substitutes for. This is a genuine limitation of the border instrument, and it is a further argument for the central recommendation: reaching the externality at source, where it is metered, is administratively as well as allocatively superior to reaching it at the border, where it must be inferred.

E.2 The political economy of instrument choice

If the usage tax is so poorly suited to every stated goal, why is it the instrument most often proposed? The model contains the beginning of an answer, and it comes from the entry result rather than from any appeal to confusion among policymakers.

Proposition E.1 (Incumbents prefer the usage tax)**.** *Consider the incumbent providers active before a tax is imposed, and compare a usage tax with a revenue-equivalent cash flow tax on rents. The cash flow tax lowers incumbent profit one for one with the revenue it raises. The usage tax, by Proposition 6, raises the equilibrium markup through the entry channel; incumbents therefore recover a share* $\Phi - 1$ *of the statutory burden through higher margins on retained sales, and for a sufficiently high fixed cost the surviving incumbents can be left* better off *than under no tax at all. Incumbents rank the instruments in the opposite order to the planner.*

The proof is immediate from the pass-through and entry results and is given in Appendix A. The content is a positive prediction, not a normative claim: the instrument that is worst for welfare is the one that the politically most concentrated interest, the incumbent frontier labs, has least reason to oppose and some reason to favour, because it raises rivals' entry costs and lifts margins on retained sales. A cash flow tax falls squarely on those same incumbents with no offset. Standard theories of capture then predict the observed pattern, energetic resistance to profit taxes and competition policy, comparative equanimity toward usage levies.[E.1]

This yields a testable implication that separates the account from mere pessimism. If incumbents prefer the usage tax for the entry-barrier reason, their revealed positions should track firm size and sunk cost: the largest labs least opposed to usage levies and most opposed to rent taxes and interoperability mandates, smaller entrants holding the reverse ranking. That is a pattern one can look for in lobbying disclosures and public comments. The normative implication is not defeatism but sequence: because the worst instrument is the path of least resistance, the reforms most worth defending are the ones the political system will fight hardest.

E.3 Reading the debate

Section 7 scores the four instruments against each other on revenue, distortion and incidence, and we do not repeat that ranking here. Two features of the debate are not captured by any scoring table, and this subsection records them.

[E.1] Two further forces push the same way. Revenue salience favours the usage tax, since a fast-growing visible transaction base is easier to legislate and score than economic rent, whose base is contested and whose yield is uncertain. And the incidence of the emissions charge, though efficient, is geographically concentrated on districts hosting data centers, whose representatives internalize the local cost while the climate benefit is diffuse, the familiar configuration that blocks Pigouvian taxes generally.

Electricity and the grid. This is where the model puts nearly all of the weight, and where the instruments partly exist already. The reforms that matter are not new taxes: they are cost allocation rules that charge large loads for the capacity and transmission they cause, tariffs with meaningful minimum take and exit provisions to protect ratepayers against stranded investment, and a carbon price that applies to data center load like any other. The evidence on whether cost shifting has occurred is genuinely mixed (Energy and Environmental Economics (E3), 2026; Wiser et al., 2026; Joint Legislative Audit and Review Commission, 2024), but the direction of reform does not depend on resolving it. Charging a load for the cost it causes is right whether or not it has so far failed to be charged.

Neutrality. Acemoglu et al. (2020) show that the tax code already tilts the user cost toward capital and away from labor. In the language of (12), that is an existing negative t_U that nobody legislated. The first-best response is to remove it rather than to add an offsetting tax on AI, since a symmetric distortion is not a substitute for no distortion.

Remark E.1 (The formula is not stable). The ratio D/p moves over time and in both directions. The price of a token has been falling by roughly an order of magnitude per year at fixed capability, which pushes D/p up. Energy per token has also been falling, which pushes it down. Reasoning and agentic workloads consume far more energy per unit of delivered value, which pushes it up again. A rate written into statute today will be wrong within two years. This is an argument for indexing any levy to a measured physical quantity, kilowatt-hours or tons of CO_2, rather than to the value of a transaction. It is the targeting principle, arriving from a different direction.