Artificial Intelligence, Competition, and Welfare

Susan Athey and Fiona Scott Morton

19 September 2025

Draft – Comments Welcome

Abstract

We study how market power in artificial intelligence (AI) shapes wages and welfare in open-economy general equilibrium by treating AI as a *priced, imported factor*. Across three models, we separate technical efficiency from the impact of upstream price setting. In a two-traded-goods benchmark, the incidence of AI price changes depends on how sectoral skill intensity changes with AI prices; non-monotone intensity can generate "double harm" for unskilled workers (lower real wage after a large *decrease* in the price of AI, and real wage decreases further when the AI price *rises* as a result of market power). With one non-traded sector, we observe that the classic "Dutch disease" effect here would arise when one sector gets more productive and draws labor away from other sectors, creating scarcity and raising prices; but this is not what we expect from the introduction of labor-substituting AI. In contrast, our last model considers two non-traded sectors and CES/free entry, and the opportunity for *discrete adoption* of technology that replaces unskilled labor from the AI-using sector. When AI reduces unit costs and increases variety, it will not pull U from non-tradables, instead it will displace workers from the AI-using sector and lower wage due to diminishing returns in alternative sectors. Strategic upstream pricing of AI then harms welfare through *unit-cost* (usage fees) and *variety* (access fees) channels, with *income leakage* abroad. We derive an adoption frontier tying feasible usage prices to displaced workers' outside options and show a monopolist typically prices on this boundary; capping one instrument shifts rents to the other. Broad gains for the adopting country relies on pressure (or regulation) on both usage and access fees and as well as policy that supports productive absorption of displaced labor. The framework clarifies when AI can lower real wages and aggregate welfare despite efficiency gains.

Introduction

Economists have observed that artificial intelligence (AI) may have a variety of downstream benefits for the economy: it may improve productivity, enable innovation and the entry of new firms, and it may be used to reduce switching costs, eroding the market power of existing firms. While recognizing the potential benefits of AI, regulators also grapple with its risks. In addition to concerns about war and democracy, there are economic risks such as worker displacement and industry reorganization that affect distribution within and across countries. In this paper, we focus on the relationship between these risks and what we view as a critical factor affecting the balance of benefits and risks: market power in AI. Competition authorities worldwide have recognized that AI may create new bottlenecks, leading to higher prices, reduced innovation, and lower benefits for consumers. The UK's Competition and Markets Authority, the US FTC, and other agencies have issued reports and are monitoring commercial relationships within the AI stack.[1]

Scrutiny is warranted because different layers of the AI stack—from chips to foundation models to applications and distribution channels—exhibit concentration. Examples include NVIDIA's dominance in advanced chips, concentration among foundation model providers (OpenAI, Gemini, Claude), and the dependence of downstream applications (e.g. search engines, language tutors, writing assistants, customer service platforms, document search tools) on a few foundation models. Access to data that is crucial for AI performance may be gated by incumbents in both the consumer and enterprise software markets. Distribution channels, such as mobile operating systems, productivity software, and online platforms, are especially concerning given their entrenched power. When these channels are controlled by firms that also supply AI or search, entrants face foreclosure risk, and even highly efficient AI applications can struggle to gain distribution (as was publicly argued by Perplexity in 2025 in the context of the Google search remedies discussion, e.g. Vasant [2025]).

Competition enforcers are particularly wary of underenforcement in AI, recalling slow responses to earlier digital platform monopolization (Google in search and ads, Meta in social networking, Amazon in e-commerce, Apple and Google in mobile ecosystems). In those cases, delayed enforcement allowed monopolies to entrench themselves. In the case of AI, by contrast, agencies are proactively investigating potential harms while the market remains in flux. This approach may be especially important because AI is a general-purpose technology, meaning any monopoly harm will extend well beyond a single sector. Having said that, the enforcement tools available to antitrust authorities are limited, and enforcement may be misdirected or ineffective. The general equilibrium framework introduced in this

[1] [Competition and Markets Authority, 2023, 2024, Portuguese Competition Authority, 2023]

paper formalizes the economy-wide consequences of unchecked market power.

In this paper, we argue that the extent to which the artificial intelligence (AI) industry provides services to traditional industries at competitive, low prices has critical implications for the impact of AI on income distribution and welfare. To do so, we consider extensions and variants of standard general equilibrium and trade models, highlighting how different assumptions about the nature of technology, the productive alternative uses of labor, and the aggregate market power in the artificial intelligence stack leads to different conclusions about whether *real* wages for different categories of workers rise or fall. Unlike much of the literature that focuses primarily on distribution, we consider assumptions that are rich enough such that productivity improvements from AI do not necessarily result in improvements in economy-wide welfare, let alone welfare for specific groups.

By focusing on a limited set of extensions to existing models, we are able to highlight novel forces that arise with different assumptions in familiar modeling frameworks, but our view is that future research may productively expand the set of models to better match possible impacts of the introduction of AI. For example, AI may make it difficult for some groups of workers to transition from shrinking sectors to growing ones, for countries to increase exports at fixed prices, and for citizens to maintain influence over government policies. Future research might also provide a deep analysis of incentives for innovation as well as industrial policies and policies around the procurement and provision of government services that strategically position countries to benefit as much as possible, and mitigate harm, from AI. But in this paper our goal is primarily to set up some of the key forces and highlight the tradeoffs that we believe advance a critical research agenda.

Specifically, this paper develops formal models to study how the arrival and pricing of artificial intelligence (AI) services affect wages, output, and welfare. To do so, we observe that the arrival of a new technology can be modeled similarly to the availability of a new input factor available through trade. Thus, we analyze the impact of AI using models of a small economy, considering cases of open and closed economies. However, we include a competition element in the model by treating AI as an intermediate input X supplied either by a foreign monopolist, supplied competitively, or somewhere in between. The price p_X of AI is therefore not a world price taken as given, but a strategic variable chosen by the upstream supplier, where the supplier may set the price to maximize revenue, but may also be constrained by policy. The goal is to understand how the introduction of AI (moving p_X from ∞ to a finite level) and subsequent changes in its price redistribute income across factors, shift output, and affect aggregate welfare. We are particularly interested in what happens to real wages for workers, and whether high prices for AI lead to reductions in the real wage despite efficiency benefits in production. In extensions, we highlight the particularly challenging but realistic

scenario where AI providers can use nonlinear pricing schemes, influencing entry, innovation and industry structure in downstream industries.

In our models the supplier of AI may be able to extract the surplus created by AI technology, while, at the same time, the exercise of market power inflicts additional loss on the workers displaced by the AI. This contrasts with a standard model where, if a technology is introduced that disproportionately displaces a group of workers, raising prices for the input helps that group by diminishing the substitution.

We also highlight the key question of whether other sectors, either non-traded sectors or export sectors, can productively absorb displaced workers, an assumption that may be questionable in the face of a major technological shift. This focus underscores the critical role played by government policies that change how productive workers are in sectors *less* affected by AI. For example, the government may increase spending on services that benefit society and then help displaced workers transition into those industries. In many countries education, health care, policing, or child and elder care are consumed at levels too low for citizens to achieve their potential in production. Government policy can affect reskilling in these areas, and AI can be used to help workers become more productive in new roles.

The popular interest in AI has stimulated a macroeconomic literature modeling the possible faster growth in several ways. For example, Aghion et al. [2019] embed AI into standard growth frameworks to explore implications for the labor–capital ratio and the share of expenditure on automated tasks. Outcomes depend on parameters such as substitution elasticities between automated and non-automated goods, the speed of cost reductions for automated goods, and the proportion of tasks that can be automated. Nordhaus [2021] provides a clear treatment of the channels through which AI can increase economic growth. He presents a model in which AI induces capital deepening, which in turn accelerates growth. Nordhaus emphasizes that returns accrue primarily to capital: *"Capital eventually gets virtually all the cake, but the crumbs left for labor—which are really small pieces of the increasingly huge mountains of cake—are still growing at a phenomenal rate"* (p. 14).

Studies of task–based and automation frameworks (e.g., Autor et al., 2003, Acemoglu and Restrepo, 2018, 2020, Autor and Salomons, 2018, Autor and Thompson, 2025) study how technology displaces or complements labor when technology costs and output prices are taken as parametric. Growth papers embed AI as a driver of factor usage and total factor productivity (e.g., Aghion et al., 2019, Nordhaus, 2021); open-economy work on directed technical change treats final goods as traded and incidence of impacts as disciplined by terms-of-trade movements (e.g., Korinek and Stiglitz, 2018, 2021). Our Model 1 nests that benchmark. The new forces we emphasize—AI as a *priced, imported factor* with upstream market power and endogenous adoption—require letting p_X be strategic and allowing non-

tradables to affect the CPI, so that aggregate income gains and real-wage losses can coexist.

A second strand of research connects to Dutch-disease and non-tradables [Corden and Neary, 1982, Corden, 1984]: here the "boom" in B is not a free endowment but requires paying p_X, so income can leak abroad.

Korinek and Stiglitz [2018, 2021] propose a series of models that focus on distributional consequences and policy responses from exogenous technological shocks or changes in the returns to the resources of small countries. A complementary line of work, going back to Stiglitz [1976] and further developed in Delli Gatti et al. [2012], Gatti et al. [2012], studies dual-economy settings in which a constant-returns agricultural sector coexists with an urban sector featuring wage rigidity (e.g., due to efficiency-wage considerations). In their baseline environment, an agricultural productivity improvement is *unambiguously* welfare-reducing: higher rural productivity contracts urban employment when nominal wages cannot adjust, and with flexible wages an induced wage decline can further depress demand and raise unemployment. Although entrepreneurs may gain from lower wages, inequality-averse social welfare functions imply aggregate welfare losses. Our models connect to these results but identify distinct channels that do not rely on nominal rigidities.

Overall, these literatures do not consider jointly the overall and distributional impacts of the technology together with how the technology is priced. Additionally, in a typical growth model both the technology and the workers are located in the same economy.

Our contribution is to bring a trade/general equilibrium approach to analyzing AI, treating AI as a *priced, imported factor* whose upstream market power (and tariff–like instruments) is central. This allows us to separate (i) pure technological efficiency from (ii) the impact of the *price path of AI services* and the corresponding changes in factor markets and input markets, and to ask how welfare depends jointly on whether AI prices fall with productivity and on whether displaced workers are sufficiently productive in sectors less affected by AI. We show that conclusions about wages and aggregate welfare turn on a small set of modeling choices—whether goods are traded versus non-traded (and in turn, whether there are diminishing marginal utility in the output markets), whether automation reduces demand for some or all groups of labor to (near) zero so that output expansions in newly productive industries do not lead to increases in labor demand, and whether the upstream supplier can use nonlinear pricing—that are especially salient for AI. A recurring theme is that some familiar GE mechanisms are *not* first-order when automation in B removes unskilled labor (so $u_B \to 0$). In that case, "Dutch-disease" style labor-pull out of the non-traded sector is mechanically weaker, and the key margins instead become the *price index in* B (via m_B and variety) and *income leakage* to a foreign AI provider. Our third model is designed to surface precisely these forces.

We proceed by developing a sequence of three models, highlighting mechanisms in each model and noting where they do or do not match what we might expect in the case of AI. We then discuss policy conclusions and open research questions, taking the perspective that the simple models we propose are only a start at answering key policy questions.

Throughout, the *price of AI*, denoted p_X, is a strategic object, and firms and market outcomes respond to it. Even holding the technology set fixed, changing p_X reallocates income between factors and shifts national income. This "price versus technology" distinction underpins our richer models when output and variety are endogenous. However, we can think of large changes in price as very similar to a change in technology, since at different prices, different values of the factor will be chosen, and the cost-minimizing choice of factors and equilibrium wages can shift qualitatively. We consider both continuous and discrete production functions, where a large fall in the price of an AI input may fundamentally change which factor of production (e.g. skilled or unskilled labor) is dominant, and where labor demand may be insensitive to input prices for technologies that fully automate. We also pay special attention to the choices firms make about adopting AI in the face of a strategic monopolist setting prices for the AI input to extract surplus.

Model 1: AI as a new factor in a small economy with two traded goods.

We begin with the standard two–good Heckscher–Ohlin model. Sector A produces using skilled and unskilled labor (L_S, L_U) alone; sector B combines (L_S, L_U) with AI, purchased at price p_X from the foreign supplier. We write the quantity of labor type $i \in \{S, U\}$ in sector $j \in \{A, B\}$ as i_j. Both goods A and B are traded at exogenous world prices $(\bar{p}_A, \bar{p}_B)$. When $p_X = \infty$, AI is unavailable and the economy collapses to the benchmark two–sector HO model. When p_X is finite, sector B's costs c_B depend on (w_S, w_U, p_X), where (w_S, w_U) are the equilibrium wages of skilled and unskilled labor. Throughout, we measure the AI input price p_X in units of the traded good B (normalizing $\bar{p}_B = 1$), so payments to the foreign supplier are valued at world prices.

In this all–traded setting, zero–profit conditions determine (w_S, w_U) as a function of p_X. The Stolper–Samuelson mapping implies that one wage rises while the other falls when p_X changes; which factor gains depends on the relative intensity of A and B. Formally, for each sector $j \in \{A, B\}$, let $\theta_i^j = w_i a_i^j / c_j$ denote the cost share of factor $i \in \{S, U\}$, and let $R_j = \theta_S^j / \theta_U^j$ denote the sectoral intensity ratio. B is skill–intensive relative to A if $R_B > R_A$, and unskilled–intensive if $R_B < R_A$.

We assume that for sufficiently high p_X, sector B's skill intensity $R_B(p_X) := \theta_S^B/\theta_U^B$ is *decreasing* in p_X (equivalently, *increasing* as AI gets cheaper). A stronger condition that we refer to in some results is *strong substitutability* (SS) of X for unskilled labor in B, under which the Hicksian unit demand a_U^B falls as p_X falls and, for low enough p_X, $a_U^B \to 0$; SS implies $R_B'(p_X) < 0$ for all p_X and ensures there is at most one crossing of R_B with R_A as we vary p_X.

General-equilibrium mapping and reallocation. Totally differentiating zero-profit in A and B at fixed $(\bar{p}_A, \bar{p}_B)$ yields

$$\theta_S^A\, dw_S + \theta_U^A\, dw_U = 0, \qquad \theta_S^B\, dw_S + \theta_U^B\, dw_U + \theta_X^B\, dp_X = 0.$$

Eliminating dw_U,

$$\begin{aligned} &\left(\theta_S^B - \theta_U^B \tfrac{\theta_S^A}{\theta_U^A}\right) dw_S = -\,\theta_X^B\, dp_X, \\ &\implies \quad \operatorname{sign}\left(\tfrac{dw_U}{dp_X}\right) = \operatorname{sign}(R_B - R_A), \qquad \operatorname{sign}\left(\tfrac{dw_S}{dp_X}\right) = -\operatorname{sign}(R_B - R_A). \end{aligned} \tag{GE}$$

This highlights that a change in p_X affects R_B through both substitution across sectors as well as across factors. If B is *unskilled-intensive* ($R_B < R_A$) and p_X increases, then w_S rises and w_U falls; activity tilts from B toward A (which is relatively more skill-intensive). Applying the envelope theorem to national income (NI), we have $\frac{dY}{dp_X} = -M_X \le 0$: increasing p_X decreases income.

How do the introduction of AI and the exercise of market power affect workers? *Case A: R_B decreasing in p_X everywhere and B skill-intensive everywhere ($R_B > R_A$).* From (GE), $\frac{dw_U}{dp_X} > 0$ at all finite p_X: raising p_X helps unskilled workers. Thus the *introduction of AI* (moving from autarky $p_X = \infty$ to any finite p_X) is bad for unskilled workers, while starting from a low p_X, an increase in p_X benefits them locally. (Skilled wages move in the opposite direction: $\frac{dw_S}{dp_X} < 0$.)

Case B: R_B decreasing in p_X and crossing R_A once as p_X rises (from skill- to unskilled-intensive). Now consider the case where in autarky, Sector B is unskilled-intensive, but as p_X drops, X substitutes for unskilled workers and eventually at low p_X the sector is skilled-intensive. Then, there exists p_X^* with $R_B > R_A$ for $p_X < p_X^*$ and $R_B < R_A$ for $p_X > p_X^*$.

- For *low* p_X (where B is skill-intensive), $\frac{dw_U}{dp_X} > 0$: At low prices, the introduction/expansion of AI can be bad for unskilled workers, if the area of skill-intensity is large enough; but

anywhere this area is large enough occurs at prices where $R_B > R_A$ and so increasing p_X helps them.

- For *high* p_X (where B is unskilled-intensive), $\frac{dw_U}{dp_X} < 0$: unskilled workers are helped by the introduction of AI, but locally *hurt* by an increase in p_X.

In summary, if as $p_X \downarrow 0$ the unit demand $a_U^B \to 0$—Sector B becomes more skill intensive as AI gets cheaper. In this monotone case there can be at most a *single* intensity reversal at some p_X^* where $R_B(p_X) = R_A$. Around such a single reversal, the *local* effect of a small change in p_X and the *global* level effect relative to autarky need not align: near p_X^* the level effect for unskilled workers is positive on both sides (opening to moderately cheap AI benefits them overall), while the local derivative flips sign at p_X^* (so a price increase in AI can either help or hurt them depending on the side). This is the familiar trade intuition that the losers from ongoing technology adoption in the skill-intensive region may be locally helped by a higher p_X (e.g., a tariff or a markup reversal). However, if unskilled workers are hurt by the introduction of AI, the prices must be low enough that Sector B is skill-intensive, so that higher p_X helps them.

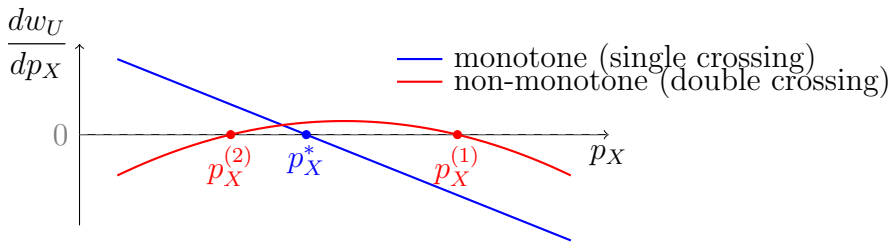


Figure 1: Local incidence of AI input prices on the unskilled wage: $\frac{dw_U}{dp_X}$ vs. p_X. Above the zero line, unskilled wages rise with p_X; below, they fall. A monotone path yields a single crossing (blue), while a non-monotone path yields two (red).

Double vs. single reversals. $R_B(p_X)$ need not be monotone. For example, sector B may be able to choose among multiple techniques with different optimal unit skill requirements (a_S^k, a_U^k, a_X^k). A natural pattern is: in autarky B is relatively unskilled–intensive; as p_X falls it becomes skill–intensive (AI substitutes for unskilled tasks); and at very low p_X AI substitutes strongly for both types of labor so B again looks unskilled–intensive:

$$R_B < R_A \quad (p_X \text{ very high}), \qquad R_B > R_A \quad (p_X \text{ intermediate}), \qquad R_B < R_A \quad (p_X \text{ very low}).$$

With two crossings of R_A, the introduction of AI supplied by a firm with market power

can deliver a "double-harm" to unskilled workers: for sufficiently low p_X, the unskilled wage can lie *below* its autarky value *and* be locally *decreasing* in p_X. The local decrease arises because at those very low prices B is again unskilled–intensive relative to A, so a higher p_X reduces B's output and differentially reduces demand for unskilled labor.

This highlights that the *shape* of $R_B(p_X)$ is critical for who is helped and hurt by input price changes. At the aggregate level, the country gains from cheaper AI (by the envelope result $dY/dp_X = -M_X \leq 0$), and one group's loss is the other's gain via the Stolper–Samuelson mapping.

Model 2: One traded, one non-traded good.

We then extend Model 1 by making A non-traded and closing the economy with Cobb–Douglas preferences over (A, B): $U = C_A^{1-\alpha} C_B^{\alpha}$, where C_j is consumption of sector j. As A is non-traded, we must have C_A equal to the output of A, Q_A.

This extension provides several additional mechanisms to analyze. First, this change endogenizes the non-traded price p_A through market clearing and introduces a novel *income–demand–price channel* on top of the Stolper–Samuelson logic, the classic "Dutch-disease" mechanism: as B scales, resources shift out of A, so Q_A decreases and $p_A = \frac{(1-\alpha)Y}{Q_A}$, increases. This raises the CPI and can lower both groups' real wages.

To see this in more detail, first recall that the fact that B is exported and the country is small induces the country to shift into exports of B when that industry gets more productive. Because household utility is Cobb–Douglas, aggregate expenditure equals domestic absorption, that is, domestic production net of AI imports:

$$Y \;=\; p_A Q_A + \bar{p}_B Q_B - p_X M_X,$$

where Y is national income (the wage bill plus rebates), Q_B is output of B, and M_X are AI imports. The consumption price index is

$$P = p_A^{1-\alpha}\, \bar{p}_B^{\alpha}, \qquad p_A = \frac{(1-\alpha)Y}{Q_A},$$

so both factors' real wages share a common cost-of-living term because both types of workers purchase more expensive A goods as a share $1-\alpha$ of expenditure. If p_A rises sharply when resources reallocate to increase production of B and decrease production in A, both real wages can fall simultaneously—an outcome that is impossible when all goods are traded. We emphasize that this mechanism depends on B actually *pulling* labor from A—an as-

sumption that may be inappropriate when AI removes unskilled workers from B's marginal cost (motivating modifications to the setup in Model 3).

Second, allowing a non-traded good (so that the CPI is endogenous) creates a channel that can generate a *double-harm* for unskilled labor. In the two-traded-goods benchmark, if $R_B(p_X)$ is monotone, then whenever a fall in p_X (the "cheap-AI" move) leaves unskilled workers worse off in levels at the new equilibrium, a marginal increase in p_X at that point must raise their (nominal and real) wage—the local and global effects necessarily have opposite signs; only with a non-monotone R_B path can both effects be adverse.

With a non-traded closure, the CPI adds a wedge: $P = p_A^{1-\alpha}\bar{p}_B^{\alpha}$ (with $\bar{p}_B$ tied to world prices), so the local real-wage response is

$$d\ln(w_U/P) = d\ln w_U - (1-\alpha)\, d\ln p_A = d\ln w_U - (1-\alpha)\,(d\ln Y - d\ln Q_A).$$

Even if the nominal Stolper–Samuelson effect points one way (e.g., B is unskilled-intensive so $d\ln w_U/d\ln p_X < 0$), a sufficiently strong co-movement of the non-traded price with p_X can overturn the sign of the real response:

$$\frac{d\ln(w_U/P)}{d\ln p_X} < 0 \quad \Leftrightarrow \quad (1-\alpha)\,\frac{d\ln p_A}{d\ln p_X} > \frac{d\ln w_U}{d\ln p_X}.$$

Hence it is possible that (i) unskilled workers are worse off in levels at the low-p_X ("cheap-AI") equilibrium relative to a high-p_X benchmark, and (ii) the local derivative of their real wage with respect to p_X at that point is negative—i.e., a marginal increase in p_X would make them even worse off. In the Appendix we give sufficient conditions—in terms of $(R_A, R_B(p_X))$ and the non-traded expenditure share $1-\alpha$—under which this alignment of global and local losses occurs.

Third, price increases by the AI provider ($p_X \uparrow$) are now *partially* cushioned or amplified through movements in p_A, depending on which sector is relatively factor-intensive. When a fall in p_X expands B and contracts A, the non-traded price rises by the closure $p_A = (1-\alpha)Y/Q_A$: higher income and lower Q_A both bid up p_A. We can formally analyze the welfare ratio between an AI and no-AI regime as follows. By the envelope theorem, $\frac{dY}{dp_X} = -M_X$, and further:

$$\frac{U^{*AI}}{U^{*0}} = \left(\frac{Y^{*AI}}{Y^{*0}}\right)^{\alpha}\left(\frac{Q_A^{*AI}}{Q_A^{*0}}\right)^{1-\alpha}.$$

In applications, the two sufficient statistics, M_X/Y and $\frac{1}{Q_A}\frac{dQ_A}{dp_X}$, provide a diagnostic: if the latter is small because B's expansion does not reduce Q_A, the income gain from cheaper AI

dominates; if it is large and positive, CPI pressures can overturn income gains.

These mechanisms connect our model to the classic Dutch disease and real-exchange-rate literatures [e.g. Corden and Neary, 1982, Corden, 1984, Burstein et al., 2003, Corsetti et al., 2008].[2], which emphasize how a resource boom or terms-of-trade improvement raises national income but simultaneously appreciates non-tradables, reallocating factors to e.g. exports and raising the cost of living. The novelty in our analysis is to embed these price-index channels in a small-open-economy model that captures the introduction of a new technology, AI, sold by a foreign supplier with market power. Here the comparative static is with respect to AI pricing, not a resource windfall. This shift allows us to characterize welfare and distributional incidence in transparent sufficient statistics: import exposure M_X/Y and non-traded scarcity elasticities dQ_A/dp_X.

Two features of Model 2 can overstate CPI pressures in Sector A as a result of AI. First, if automation in B makes $u_B \approx 0$, scaling B need not pull unskilled labor out of A; the mechanical shift of resources out of Sector A may not occur, and p_A may rise less (or not at all). Second, treating the good from B as traded at an exogenous world price abstracts from the possibility that AI adoption is global, limiting a small country's ability to expand B at fixed $\bar{p}_B$. Next, we introduce Model 3 that addresses both issues by (i) making B non-traded with CES/entry (so the price index moves through unit-cost and variety) and (ii) modeling discrete adoption that eliminates unskilled workers from B ($u_B \to 0$).

Model 3: Discrete Technology Adoption with Two Non-traded Sectors

As in Model 2, in Model 3 households have Cobb–Douglas preferences over two goods, so both types of workers face a common CPI. Here, however, two new features are introduced. First, *both* sectors are non–traded, so there is no world-price anchor: both P_A and P_B are determined domestically. Second, sector B consists of many differentiated varieties under CES demand (elasticity $\sigma > 1$) with constant markup $\mu = \sigma/(\sigma - 1) > 1$ and free entry subject to a per-firm domestic license fee F (rebated lump-sum to households) and a fixed access charge, ϕ, that is collected by the foreign monopolist.

Primitives and price indices. Households spend a constant share $\alpha \in (0, 1)$ on B and $1 - \alpha$ on A:

$$U = C_A^{1-\alpha} C_B^{\alpha}, \qquad E_B = \alpha Y, \quad E_A = (1 - \alpha)Y. \tag{1}$$

[2]See also **?**Chari et al. [2002], De Gregorio and Wolf [1994].

Within B, symmetry across the N active firms implies

$$Q_B = \Big(\sum_{i=1}^{N} q_i^{\frac{\sigma-1}{\sigma}} \Big)^{\frac{\sigma}{\sigma-1}}, \qquad P_B = \Big(\sum_{i=1}^{N} p_i^{1-\sigma} \Big)^{\frac{1}{1-\sigma}} = N^{\frac{1}{1-\sigma}} \, p_B, \tag{2}$$

and the CPI is

$$P = p_A^{1-\alpha} p_B^{\alpha}. \tag{3}$$

Without loss of generality, we normalize $p_A = 1$ so that all prices are relative to the price index in Sector A; although at times we discuss the price effects in Sector A, these should be interpreted as relative prices.

Pricing and free entry in B. Let m_B be the unit marginal cost and $p_B = \mu \, m_B$ the symmetric price. With per-firm outlays $F + \phi$ and free entry,

$$(p_B - m_B)q = (F + \phi) \quad \Rightarrow \quad p_B \, q = \tfrac{\mu}{\mu-1}(F + \phi). \tag{4}$$

Because the expenditure in Sector B is $E_B = \alpha Y = N p_B q$, the number of active firms is

$$N = \underbrace{\tfrac{\mu-1}{\mu} \, \alpha}_{\kappa} \, \frac{Y}{F + \phi}, \qquad P_B = \mu \, m_B \left(\kappa \, \frac{Y}{F + \phi} \right)^{\frac{1}{1-\sigma}}. \tag{5}$$

Thus, holding Y and $(F + \phi)$ fixed, changes in m_B move p_B but do *not* move N directly. In general equilibrium, however, p_X and ϕ move Y and hence N through free entry. This yields a clear welfare decomposition: unit-cost (m_B) and variety (N) effects are distinct levers, both sensitive to upstream market power.

It is natural to study AI adoption in a model with fixed costs because digital technologies and AI services often require substantial up-front investment (training, deployment, access) but then exhibit low marginal cost of use. A CES/free-entry structure captures this well: the number of firms N depends on income and fixed costs, not directly on marginal cost. This makes it possible to analyze how AI pricing reshapes industry structure (the equilibrium number of firms, variety, and the quality-adjusted price index). Moreover, because AI is a general-purpose technology that enables a wide range of applications, entry and variety are themselves first-order welfare channels. A CES environment highlights how monopolist pricing can distort not only wages and aggregate income but also the diversity of products consumers access. This goes beyond the simpler competitive frameworks and justifies the more complex setup.

AI technology and marginal cost in B. A foreign upstream supplier charges a per-firm access fee ϕ and a per-unit usage price p_X (we start by setting $\phi = 0$ to facilitate comparisons to Model 2, and then generalize the model). Adoption in B is discrete and takes one of two forms:

$$\textbf{Partial automation: } m_B = s_B w_S + p_X, \qquad \textbf{Full automation: } m_B = p_X. \tag{6}$$

In both cases, adoption implies $u_B = 0$, so displaced unskilled labor reallocates to A. In contrast to Model 2, because $u_B = 0$, a marginal expansion of B need not *pull* U from A; instead, U is already in A and faces diminishing returns there. The CPI channel now runs primarily through p_B (via m_B and N), not through p_A (via Q_A).

Sector A and factor markets. Sector A is produced competitively with both skilled and unskilled labor. With fixed endowments L_S, L_U,

$$L_S^A + L_S^B = L_S, \qquad L_U^A + L_U^B = L_U, \qquad L_U^B = 0 \ \text{ under adoption.} \tag{7}$$

With diminishing returns in A, the crowding of unskilled labor into A lowers w_U, while w_S reflects the allocation of L_S across sectors and—under partial automation—the demand for L_S in B.

Income and external payments. Domestic income includes wages and the rebated F (as in Model 2) but excludes foreign AI payments:

$$Y = w_S L_S + w_U L_U + FN, \qquad \text{Foreign outflow} = \phi N + p_X Q_B. \tag{8}$$

In contrast to the case of the fixed cost F, if the foreign AI supplier charges a per–firm access fee ϕ (as we consider below), those payments do not recycle domestically. Unlike F, ϕ reduces both entry and *domestic absorption*, since the transfers are lost to the foreign monopolist.

Exogenous AI Prices

Adoption frontier defined. AI is adopted only if it beats the baseline when rivals adopt. As we discuss in more detail below, this yields a downward-sloping frontier of the set of incentive-compatible (ϕ, p_X) pairs, where the frontier can be written as $p_X \leq p_{X,\max}(\phi)$. Importantly, adoption constraints are more challenging when other firms adopt, since adoption frees up labor and pushes down wages. In this section, we first consider comparative statics

on usage and access fees within the frontier, and then return to consider the frontier in more detail when we consider the problem faced by a monopolist AI provider.

Equilibrium. Let $Z-(\alpha, \sigma, F, s_B;\ L_S, L_U)$ denote the vector of exogenous primitives (preferences, technologies, and endowments) for Model 3. Given (Z, ϕ, p_X), a (competitive) equilibrium consists of prices, quantities, and allocations

$$\Big(p_A,\ P_B,\ p_B,\ m_B,\ N,\ q,\ Q_B;\ w_S, w_U;\ L_S^A, L_S^B, L_U^A, L_U^B;\ Y\Big)^{*r},$$

where the superscript ^{*r} denotes equilibrium values in Model 3 for regime $r \in \{AI, 0\}$ ($r = 0$ being no AI), that jointly satisfy equations (1)-(8). When analyzing the monopolist's problem, the feasible (ϕ, p_X) must also satisfy the adoption frontier $p_X \leq p_{X,\max}(\phi)$. For any equilibrium variable X in the list above, we use the superscript notation $X^{*r}(Z, \phi, p_X)$ to denote its equilibrium value as a function of (Z, ϕ, p_X).

Welfare Ratio In the Appendix, we show that substituting in equilibrium conditions and comparing the AI equilibrium to the no–AI baseline yields

$$\frac{W^{*AI}}{W^{*0}} = \underbrace{\left(\frac{m_B^{*0}}{m_B^{*AI}}\right)^{\alpha}}_{\text{unit cost / } B\text{-side price index}} \underbrace{\left(\frac{F}{F+\phi}\right)^{\frac{\alpha}{\sigma-1}}}_{\text{variety / entry}} \underbrace{\left(\frac{Y^{*AI}}{Y^{*0}}\right)^{\alpha+\frac{\alpha}{\sigma-1}}}_{\text{aggregate income}} \tag{9}$$

That is, the welfare ratio can be decomposed into the product of three ratios. The first is proportional to the ratio of marginal costs across regimes, which affects the welfare ratio through the price index; the second is the ratio of fixed costs, which affects the ratio of the number of firms and thus variety; the third is proportional to the ratio in aggregate income; and the expenditure ratio in Sector A, which captures scarcity that arises in Sector A when production shifts to Sector B. (Recall that p_A is normalized to 1 so that income is considered relative to p_A). Each of these terms depends on the fees (ϕ, p_X). At low/zero prices, unit costs fall, the variety/entry effect is equal to one, and equilibrium income rises. On the other hand, at high access or usage prices (but low enough to induce adoption), the unit cost/price index effect is dampened, the variety/entry effect reduces the ratio, and income is reduced.

The same decomposition can also be applied separately for skilled and unskilled workers, where the first two terms are the same across worker groups. However, wage shifts due to the adoption of AI drive differences in the income ratio changes. Displaced workers may have lower labor income, so that unskilled workers may be harmed even by cheap AI, depending on parameter values, similar to the analysis of Models 1 and 2.

Linear pricing (usage–only, $\phi = 0$). Under linear pricing the foreign supplier sets the per–unit usage price p_X with no access fee ($\phi = 0$).

In Model 3, AI enters sector B as an upstream input at unit price p_X. Adoption eliminates all labor from B's production (full automation) or alternatively allows only L_S to remain (partial automation). The adoption condition is discrete: a firm uses AI only if its unit cost with AI is below the baseline technology that employs both L_S and L_U. Importantly, adoption is *harder to sustain as it becomes widespread*. Once many firms adopt, displaced L_U shifts into sector A, driving down w_U and making the outside option, the baseline production technology, more attractive, which tightens the adoption constraint. Despite this tightening, the monopolist can still set a strictly positive p_X even though AI reduces costs relative to autarky.

We consider the introduction of the technology, modeled as a shift from autarky ($p_X = \infty$) to finite p_X. Two new forces arise relative to Model 2. First, a *variety channel*: lower p_X reduces unit costs in B, raises output and income, and—through free entry—supports more firms and more varieties. More varieties lower the quality–adjusted price index P_B, a channel absent when B was traded at a fixed world price. Second, a *displacement channel*: once AI is adopted, all L_U is pushed into sector A. With Cobb–Douglas in A, this crowding reduces w_U through diminishing returns, a sharper incidence than the partial intensity drift in Model 2.

In equilibrium, a lower p_X reduces unit costs in B and raises variety. Under partial automation, the additional production in B pulls L_S out of A, which pushes p_A upward [3]. Under mild conditions, welfare rises relative to autarky, as the unit–cost and variety gains outweigh the rise in p_A. But unskilled workers' real wage typically falls, since the unskilled nominal wage is depressed by crowding in A while the CPI rises with p_A. Skilled workers may gain under partial automation (more demand for L_S in B), but under full automation, they too may lose as all labor is absorbed into A.

A small increase $dp_X > 0$ when AI adoption constraint is slack. Suppose the adoption constraint is slack. Then, raising p_X increases unit costs in B and contracts its output. Because B has CES demand and free–entry, this also reduces the number of firms, raising P_B via the variety channel. At the same time, fewer firms in B release some L_S back to A, lowering p_A.

Unlike Model 2, these local changes do not induce substitution back to L_U in B: adoption is discrete, so a marginally higher p_X worsens P_B without undoing displacement. When the

[3] We occasionally say "p_A rises" for economic intuition. With $p_A \equiv 1$ as numeraire, this should be read as "the relative price of A would rise under an external numeraire," which in our units appears as a fall in the relative price of B.

Cobb-Douglas parameter α determining expenditure share in B is moderate, the worsening of P_B and the fall in income dominate the relief in p_A. In this case, the CPI rises, welfare falls, and both w_S/P and w_U/P typically decline. Only in special cases (tiny α, extreme scarcity in A) can the fall in p_A offset these forces. Thus, local and global (shift from autarky) incidence can align negatively for a particular worker group, the "double harm" discussed in Models 1 and 2. Here the intuition is similar at a high level, but the effect arises from both discrete technology adoption and through prices.

Two–part tariffs (ϕ, p_X)**.** Allowing the foreign AI supplier to set both a per–unit price p_X and an access fee ϕ adds a second lever absent in Model 2. A higher ϕ directly reduces entry in B, raising P_B, and, unlike F, ϕ's proceeds leak abroad, lowering domestic income.

Relative to Model 2, *two CPI channels* (higher unit costs and lower variety) and a new *income–leakage channel* (foreign ϕ) arise when the AI provider raises prices. Both effects raise P_B and reduce welfare. Real wages of both L_S and L_U typically fall, since all workers share the CPI, while the loss of the domestic license fee F from the reduction in firms increases relative differences. Regulating one lever alone simply induces rent shifting to the other; restoring efficiency requires disciplining both.

Overall, the combination of discrete displacement, endogenous variety, and two–part tariffs makes it far easier for market power in AI supply to depress welfare, and it is more likely that the introduction of AI harms unskilled workers, and that local exercise of market power furthers the harm.

AI Monopolist Choice of Usage and Access Fees

Sector B Firm Profits. We can also examine the impact of access and usage fees on Sector B production and gross profits. Define the gross firm–side aggregate operating surplus before fees as follows (where we substitute in equilibrium conditions):

$$\mathcal{E}^{\text{gross}}(Z,\phi,p_X) \;\equiv\; (\mu-1)\, m_B^{*AI}(Z,\phi,p_X)\, Q_B^{*AI}(Z,\phi,p_X) = \frac{(\mu-1)\alpha}{\mu}\, Y^{*AI}(Z,\phi,p_X). \quad (10)$$

Note that AI access and usage fees impact this *only* through Y^{*AI}. Its ϕ–slope is

$$\frac{\partial \mathcal{E}^{\text{gross}}}{\partial \phi}(Z,\phi,p_X) = \frac{(\mu-1)\alpha}{\mu}\, \frac{dY^{*AI}}{d\phi}(Z,\phi,p_X) \;<\; 0, \quad (11)$$

where $dY^{*AI}/d\phi$ incorporates a *direct* income effect, where ϕ affects the share of consumer expenditure retained by firms, and an *indirect* effect via the induced change in Q_B^{*AI}. In contrast, p_X has only an indirect effect on $\mathcal{E}^{\text{gross}}$, through wages and output that affect

income. If we think of $\mathcal{E}^{\text{gross}}(Z,\phi,p_X) - N^{*AI}F$ as the "size of the pie" to be extracted by the monopolist through access and usage fees, we can see that both ϕ and p_X reduce the size of the pie by distorting income. This contrasts with the typical nonlinear pricing problem from industrial organization, where the number of firms is fixed and there are no general equilibrium effects, so that access fees do not distort production while usage fees do. In the latter case, it is optimal for a monopolist to keep usage fees as low as possible and extract surplus using access fees.

Adoption frontier in equilibrium. Our next step will be to analyze strategic pricing by the monopolist, which in general, at the profit maximum, lies on the adoption frontier. Here, we develop the frontier in more detail. The frontier $p_X \leq p_{X,\max}(\phi)$ captures the idea that when rivals adopt, wages (w_S^{*AI}, w_U^{*AI}) fall for the displaced factor, so the baseline alternative improves; sustaining adoption therefore requires *lower* p_X or *lower* ϕ. We argue below that in general, at the monopolist optimum the frontier binds, so upstream prices track the outside option of the displaced workers, a feature absent in standard industrial organization models of licensing without GE feedback.

Consider one firm that *deviates* to the baseline (non-adopting) technology while all rivals keep adopting and charging $p_B^{*AI} = \mu\, m_B^{*AI}$. Let the deviator set the usual markup price $p_{\text{dev}} = \mu\, m_{\text{dev}}$, where its baseline marginal cost is evaluated at the AI equilibrium wages:

$$m_{\text{dev}}(Z,\phi,p_X) \;\equiv\; s_B\, w_S^{*AI}(Z,\phi,p_X) \;+\; u_B\, w_U^{*AI}(Z,\phi,p_X). \tag{12}$$

Under CES demand, the deviator's quantity at (Z,ϕ,p_X) is

$$q_{\text{dev}}(Z,\phi,p_X) = \alpha\, Y^{*AI}\,(\mu\, m_{\text{dev}})^{-\sigma}\,(P_B^{*AI})^{\sigma-1} = \frac{\alpha}{\mu}\,\frac{Y^{*AI}}{m_{\text{dev}}}\left(\frac{m_B^{*AI}}{m_{\text{dev}}}\right)^{\sigma-1}\frac{1}{N^{*AI}}, \tag{13}$$

where we used $P_B^{*AI} = p_B^{*AI}(N^{*AI})^{\frac{1}{1-\sigma}} = \mu\, m_B^{*AI}(N^{*AI})^{\frac{1}{1-\sigma}}$.

The deviator does not pay the AI access fee (it does not adopt), so its profit is

$$\pi_{\text{dev}}(Z,\phi,p_X) = (\mu-1)\, m_{\text{dev}}\, q_{\text{dev}} - F = \frac{(\mu-1)\alpha}{\mu}\, Y^{*AI}\left(\frac{m_B^{*AI}}{m_{\text{dev}}}\right)^{\sigma-1}\frac{1}{N^{*AI}} - F. \tag{14}$$

Using $N^{*AI} = \kappa\, Y^{*AI}/(F+\phi)$ with $\kappa = \frac{\mu-1}{\mu}\alpha$ (from (5)), the no-deviation condition $\pi_{\text{dev}} \leq 0$ is equivalent to

$$\left(\frac{m_B^{*AI}}{m_{\text{dev}}}\right)^{\sigma-1} \leq \tfrac{F}{F+\phi} \iff m_B^{*AI} \leq m_{\text{dev}}\left(\tfrac{F}{F+\phi}\right)^{\frac{1}{\sigma-1}} \Bigg|_{(Z,\phi,p_X)}. \tag{15}$$

Writing out m_B^{*AI} for the two adoption modes in (6) yields explicit upper bounds on p_X:

$$\begin{aligned} &\textbf{Partial automation:} \quad m_B^{*AI} = s_B w_S^{*AI} + p_X \quad \Rightarrow \quad p_X \le m_{\text{dev}} \left(\tfrac{F}{F+\phi}\right)^{\frac{1}{\sigma-1}} - s_B w_S^{*AI}, \\ &\textbf{Full automation:} \quad m_B^{*AI} = p_X \qquad\qquad\qquad \Rightarrow \quad p_X \le m_{\text{dev}} \left(\tfrac{F}{F+\phi}\right)^{\frac{1}{\sigma-1}}. \end{aligned} \tag{16}$$

Appendix Proposition 9 provides conditions under which any profit-maximizing (ϕ, p_X) lies on the adoption frontier $p_X = p_{X,\max}(\phi)$.

Monopolist's Optimal Usage Fee Consider the case where $\phi = 0$, and only usage fees are assessed. Given (Z, p_X), the supplier's revenue is

$$\Pi^{\text{lin}}(Z, p_X) \;\equiv\; p_X \, Q_B^{*AI}(Z, 0, p_X). \tag{17}$$

Using properties of the CES/free–entry model,

$$Q_B^{*AI}(Z, \phi, p_X) \;=\; \frac{\alpha}{\mu} \frac{Y^{*AI}(Z, \phi, p_X)}{m_B^{*AI}(Z, \phi, p_X)}, \tag{18}$$

so

$$\Pi^{\text{lin}}(Z, p_X) \;=\; \frac{\alpha}{\mu} \, Y^{*AI}(Z, 0, p_X) \, \frac{p_X}{m_B^{*AI}(Z, 0, p_X)}. \tag{19}$$

The AI marginal cost in B is given in (6).

Adoption constraint under $\phi = 0$. With $\phi = 0$ and all terms evaluated at $(Z, 0, p_X)$, the no–deviation frontier (15) reduces to

$$\text{Partial automation:} \qquad p_X \;\le\; p_{X,\max}^{\text{part}}(Z) \;\equiv\; m_{\text{dev}} - s_B \, w_S^{*AI} \;=\; u_B \, w_U^{*AI}, \tag{20}$$

$$\text{Full automation:} \qquad p_X \;\le\; p_{X,\max}^{\text{full}}(Z) \;\equiv\; m_{\text{dev}} \;=\; s_B \, w_S^{*AI} + u_B \, w_U^{*AI}. \tag{21}$$

Problem and characterization. The linear–pricing problem is

$$\max_{p_X \ge 0} \;\; \Pi^{\text{lin}}(Z, p_X) \qquad \text{s.t.} \quad p_X \;\le\; p_{X,\max}^{\text{mode}}(Z), \tag{22}$$

where "mode" is either "part" or "full" according to (20)–(21). When the adoption constraint is slack, an interior candidate satisfies

$$\frac{d}{dp_X} \Pi^{\text{lin}}(Z, p_X) = \underbrace{Q_B^{*AI}(Z, 0, p_X)}_{\text{usage fee gain}} + \underbrace{p_X \, \frac{dQ_B^{*AI}}{dp_X}(Z, 0, p_X)}_{\text{scale loss}} = 0, \tag{23}$$

with

$$\frac{dQ_B^{*AI}}{dp_X}(Z,0,p_X) = \frac{\alpha}{\mu}\,\frac{1}{\left(m_B^{*AI}\right)^2}\Big(m_B^{*AI}\,\frac{dY^{*AI}}{dp_X} - Y^{*AI}\,\frac{dm_B^{*AI}}{dp_X}\Big)\Big|_{(Z,0,p_X)}. \tag{24}$$

Otherwise, the optimum is on the *boundary*. In the Appendix, we show that under relatively mild conditions, the optimum is at the boundary, so that the AI provider charges a usage fee equal to per-unit wage for unskilled workers. Note that at $p_X = 0$, the scale loss in (23) (the inframarginal effect of losing sales) is zero, so that it is optimal to charge a positive fee.

Access pricing. Adoption requires satisfying a downward–sloping frontier: a higher ϕ must be offset by a lower p_X to keep firms willing to adopt. In the Appendix, we show that the optimum is on the boundary and that in general, the monopolist splits extraction across ϕ (per–firm) and p_X (per–unit) *along* a downward–sloping cap $p_{X,\max}(\phi)$. : ϕ (per–firm) and p_X (per–unit).

Fix (Z, p_X) and let

$$\Pi^{\text{acc}}(Z,\phi \mid p_X) \;\equiv\; \underbrace{\phi\, N^{*AI}(Z,\phi,p_X)}_{\text{access revenue}} \;+\; \underbrace{p_X\, Q_B^{*AI}(Z,\phi,p_X)}_{\text{usage revenue}} \tag{25}$$

denote the foreign supplier's revenue as a function of ϕ.

Define the elasticities and weights

$$\varepsilon_{Y,\phi}(Z,\phi,p_X) \;\equiv\; -\frac{\phi}{Y^{*AI}}\,\frac{dY^{*AI}}{d\phi}, \qquad \eta_{Q_B,\phi}(Z,\phi,p_X) \;\equiv\; -\frac{\phi}{Q_B^{*AI}}\,\frac{dQ_B^{*AI}}{d\phi}, \tag{26}$$

$$\kappa^{\text{lin}}(Z,\phi,p_X) \;\equiv\; \frac{p_X}{(\mu-1)\,m_B^{*AI}(Z,\phi,p_X)}\cdot\frac{F+\phi}{\phi}. \tag{27}$$

Using the identity $N^{*AI}/Q_B^{*AI} = (\mu-1)\,m_B^{*AI}/(F+\phi)$ and other equilibrium identities, we have

$$\frac{\partial \Pi^{\text{acc}}}{\partial \phi} = N^{*AI}\left[\frac{F}{F+\phi} - \varepsilon_{Y,\phi}(Z,\phi^\star,p_X) - \kappa^{\text{lin}}\,\eta_{Q_B,\phi}(Z,\phi^\star,p_X)\right]. \tag{28}$$

If the adoption frontier binds at (Z,ϕ,p_X), the solution is on the boundary at the largest feasible ϕ consistent with adoption; if the optimum is interior, (28) determines the optimum. In general, the monopolist will use a mix of access and usage fees to extract as much revenue as possible without hurting worker income and the number of firms too much. Capping one fee in isolation generally induces rent shifting to the other instrument: a tighter cap on ϕ (resp. on p_X) raises the chosen p_X (resp. ϕ) along $p_{X,\max}(\phi)$ (See Appendix Proposition 11).

Summary of Model 3

All three of our models demonstrate that monopoly power in AI depresses welfare and can harm both skilled and unskilled workers. Relative to Model 2, the distinctive mechanisms in Model 3 are: (i) *no unskilled labor-pull* from A when $u_B = 0$, so p_A plays a smaller CPI role; (ii) a *variety channel* in B that is directly distorted by fixed access charge ϕ; and (iii) *boundary pricing* where p_X is tied to displaced workers' outside options via the adoption frontier. These differences alter which mechanism delivers the harm. In our first two models, welfare losses come from the cost of buying the AI which shows up in the import bill (M_X) and, when there is a non-traded sector, from non-traded scarcity (dQ_A/dp_X). Distributional harm may occur as well. In Model 3, welfare losses operate through three distinct margins: (i) higher unit costs from the per output fee, p_X, (ii) fewer varieties from the fixed fee, ϕ, and (iii) direct income leakage abroad as before. Distributional impacts arise for reasons similar to Model 2: displacement of unskilled workers depresses real wages, while changes in the industry structure and output in Sector A govern the skill premium.

The features in Model 3 better align with AI deployments where automation removes unskilled labor from the marginal cost of scalable services and the adoption of the technology is a discrete decision. The richer structure allows us to analyze how a monopolist splits extraction between fixed and variable fees, how policy interventions that cap one lever shift extraction into the other, and why regulating both fees together is essential. Model 3 shows how monopoly power in a general-purpose technology affects not only wages and aggregate income but also the range of products and the quality-adjusted consumption basket. It also introduces the adoption frontier, more realistically modeling discrete changes in the organization of production in response to AI availability.

Conclusions: A Policy-Relevant Research Agenda

The models developed above are deliberately stylized, but the models and results can be used to identify and prioritize open questions for research and considerations for policy-makers. Two high-level insights for policy follow from our GE/trade framing. First, broad welfare gains require that *AI prices fall with productivity* (competition or policy imposed on p_X and ϕ) *and* that displaced labor be *absorbed productively* in sectors less affected by AI. Second, policies that only cap one instrument (say ϕ) risk rent-shifting into the other (higher p_X); disciplining both levers is necessary when the upstream supplier prices on the adoption frontier. Beyond these high-level insights, further work should attend to the specific circumstances of the country or worker type under study.

Consider the perspective of a policy-maker in a particular country. Our models highlight that policy impacts depend on whether the value added from AI accrues domestically or abroad. If profits from AI are earned by domestic firms, they directly contribute to national income and help sustain demand for non–traded goods, which supports employment across the economy. If instead AI services are imported, their cost may be high, and rents may accrue to a foreign monopolist or strategic supplier. Additionally, if the AI is sold by a monopolist owned by a small number of natural persons – perhaps persons with access to skilled tax lawyers – the surplus is less likely to contribute to national demand for non-traded goods. In that case, domestic absorption is lower and welfare and inequality may be harmed. Research might consider for different groups of countries, how likely are they to participate in the production of AI, in particular the parts that give rise to high markups and profits? If a country is likely an importer of AI technology, the concerns highlighted in this paper are particularly relevant, although even the countries that produce AI may find that the concentration of profits with a domestic firm does not lead to desirable efficiency and distributional consequences.

For workers, much depends on whether displaced labor can be reallocated productively. Improving the productivity of workers is itself an enormous research agenda for labor economists given the many choices governments make concerning education, training, and subsidies for socially-valuable activities. In non–traded industries such as teaching, nursing, or care, the relevant question is whether policy can stimulate demand and raise productivity fast enough to absorb displaced workers. Governments typically play a role in funding and procuring these services, and AI assistants might help with both transitions and productivity of the displaced workers in the services they now provide. The political economy of these skills investment choices will also require modeling because the scale of the problem is so large that it will be a significant spending decision for government. Another alternative to training and re-organizing the entire economy in response to AI is to provide a universal basic income to the displaced workers or to all workers. Part of the political economy research agenda is modeling the source of the revenue for protecting workers, what entity or activities can be taxed most efficiently, and, critically, whether the owners of the AI are able to affect the government's decision. If the AI monopolist can capture the government, the monopolist may be able to escape its share of the costs it imposes on broader society.

If AI–enabled industries are export industries, then increased productivity can soften the impact on demand for labor (by increasing production in the newly efficient sector) as long as world prices remain high. But if other countries adopt AI at the same time, world prices may fall, reducing the scope for export–led mitigation of labor displacement. Some export industries that a country might have previously participated in, such as outsourcing

of call centers, may shrink substantially. The impact of declines in outsourcing on developing countries and changes in industry structure, international competitiveness, and global distribution of production deserve further study.

Another concern is the possibility of permanent loss of productive capabilities. If AI adoption displaces domestic production capacity, and those capabilities are difficult to rebuild, a country may become more vulnerable to market power in the future. We have in mind familiar issues from the industrial policy literature such as learning by doing, agglomeration efficiencies, organizational forgetting, supply chain resilience, etc. Permanent loss of an industry raises new questions about industrial policy towards AI: how should countries weigh the short–term efficiency gains from AI adoption against the longer–term risks of dependency and erosion of local expertise? A darker concern along these lines is the possibility that the monopoly AI provider becomes controlled by a hostile state. Such an actor might not want to maximize profit but rather execute on different goals such as lowering domestic capabilities, creating vulnerabilities, impoverishing workers, and so forth.

The interaction between AI and innovation itself also opens a rich research agenda. On the one hand, AI can erode market power in existing industries by lowering costs and enabling entry. On the other hand, our analysis shows that if fixed costs rise, industry structure may become more concentrated, reducing the variety available to end consumers and slowing innovation. Moreover, if monopoly rents are the primary reward for investing in frontier AI, restricting those rents too aggressively might dull incentives for innovation at the top of the AI stack. Understanding how different forms of pricing, competition, and regulation shape innovation incentives in AI, and how this interacts with broader patterns of technological progress, remains an open and pressing challenge.

Another open area for research is whether and how market power arises or can be constrained, and whether open source models have a role to play in constraining market power. Recently policy issues have concerned whether open models (such as Meta's Llama) should be regulated for national security reasons, though the arrival of DeepSeek highlighted that it may be difficult to keep countries from fast-following others' innovations. More broadly, research might consider how the presence of open models, competitive entrants, or fast followers might discipline incumbents and reduce the scope for sustained market power, and whether industrial policy could be crafted that improves welfare for a country.

A related set of questions concerns the industrial structure of the AI industry itself. Because the AI "stack" involves multiple layers—chips, training, models, applications—any single bottleneck can sustain high prices for downstream users. Mapping which layers are prone to bottlenecks, and how policy can monitor and respond to them quickly enough to be relevant, is an urgent task. The structure of the industry may affect the extent to which

the final AI pricing is similar to that of a monopolist, and how regulation on pricing (or other trade instruments) best respond to the incentives of AI providers. From a modeling perspective, the simple structure of CES demand and a monolithic AI industry could both be generalized, yielding more realistic insights.

Overall, the payoff from preserving competition in AI may be underappreciated in macroeconomic discussions. Growth models and popular narratives often assume that AI will deliver cheap goods and services in abundance, where the main drawbacks are distributional. Our observation is that this outcome is unlikely in the absence of competition, particularly if the profits from the technology accrue outside a country or more broadly are not shared throughout the economy. Instead, we show how monopoly in AI allows the provider to extract rents and prevent prices from falling to match declines in labor. Because AI is a general-purpose technology, monopoly harms extend well beyond a single sector, justifying a general equilibrium framework. Our models demonstrate that the welfare impact of AI depends critically on market structure; they further highlight the important role of worker transitions and their productivity in alternative sectors.

References

Daron Acemoglu and Pascual Restrepo. The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, 108(6):1488–1542, 2018. doi: 10.1257/aer.20160696.

Daron Acemoglu and Pascual Restrepo. Robots and jobs: Evidence from us labor markets. *Journal of Political Economy*, 128(6):2188–2244, 2020. doi: 10.1086/705716.

Philippe Aghion, Benjamin F. Jones, and Charles I. Jones. Artificial intelligence and economic growth. In Ajay Agrawal, Joshua Gans, and Avi Goldfarb, editors, *The Economics of Artificial Intelligence: An Agenda*, pages 237–282. University of Chicago Press, 2019.

Mark Armstrong. Competition in two-sided markets. *The RAND Journal of Economics*, 37 (3):668–691, 2006. doi: 10.1111/j.1756-2171.2006.tb00037.x.

David Autor and Neil Thompson. Expertise. *Journal of the European Economic Association*, page jvaf023, 2025.

David H. Autor and Anna Salomons. Is automation labor-displacing? productivity growth, employment, and the labor share. *Brookings Papers on Economic Activity*, 2018(1):1–63, 2018. doi: 10.1353/eca.2018.0000.

David H. Autor, Frank Levy, and Richard J. Murnane. The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, 118(4): 1279–1333, 2003. doi: 10.1162/003355303322552801.

Ariel Burstein, João C. Neves, and Sergio Rebelo. Distribution costs and real exchange rate dynamics during exchange-rate-based stabilizations. *Journal of Monetary Economics*, 50 (6):1189–1214, 2003. doi: 10.1016/S0304-3932(03)00075-8.

V. V. Chari, Patrick J. Kehoe, and Ellen R. McGrattan. Can sticky price models generate volatile and persistent real exchange rates? *Review of Economic Studies*, 69(3):533–563, 2002. doi: 10.1111/1467-937x.00216.

Competition and Markets Authority. Ai foundation models technical update report, 2023. URL `https://assets.publishing.service.gov.uk/media/65081d3aa41cc300145612c0/Full_report_.pdf`.

Competition and Markets Authority. Ai foundation models update report, 2024. URL `https://assets.publishing.service.gov.uk/media/661e5a4c7469198185bd3d62/AI_Foundation_Models_technical_update_report.pdf`.

W. Max Corden. Booming sector and dutch disease economics: Survey and consolidation. Oxford Economic Papers on Energy 3, Oxford Institute for Energy Studies, 1984.

W. Max Corden and J. Peter Neary. Booming sector and de-industrialisation in a small open economy. *The Economic Journal*, 92(368):825–848, 1982. doi: 10.2307/2232670.

Giancarlo Corsetti, Luca Dedola, and Sylvain Leduc. International risk sharing and the transmission of productivity shocks. *Review of Economic Studies*, 75(2):443–473, 2008. doi: 10.1111/j.1467-937x.2008.00475.x.

Jos'e De Gregorio and Holger C. Wolf. Terms of trade, productivity, and the real exchange rate. NBER Working Paper 4807, National Bureau of Economic Research, 1994.

Domenico Delli Gatti, Mauro Gallegati, Bruce C Greenwald, Alberto Russo, and Joseph E Stiglitz. Sectoral imbalances and long-run crises. In *The global macro economy and finance*, pages 61–97. Springer, 2012.

Wilfred J. Ethier. National and international returns to scale in the modern theory of international trade. *American Economic Review*, 72(3):389–405, 1982.

Maryam Farboodi, Roxana Mihet, Thomas Philippon, and Laura Veldkamp. Big data and firm dynamics. *American Economic Review: Papers & Proceedings*, 109:38–42, 2019. doi: 10.1257/pandp.20191001.

Domenico Delli Gatti, Mauro Gallegati, Bruce C Greenwald, Alberto Russo, and Joseph E Stiglitz. Mobility constraints, productivity trends, and extended crises. *Journal of Economic Behavior & Organization*, 83(3):375–393, 2012.

Gene M. Grossman and Esteban Rossi-Hansberg. Trading tasks: A simple theory of offshoring. *The American Economic Review*, 98(5):1978–1997, 2008.

Morton I. Kamien and Yair Tauman. Fees versus royalties and the private value of a patent. *The Quarterly Journal of Economics*, 101(3):471–491, 1986. doi: 10.2307/1885693.

Michael L. Katz and Carl Shapiro. Network externalities, competition, and compatibility. *The American Economic Review*, 75(3):424–440, 1985. URL `http://www.jstor.org/stable/1814809`.

Michael L. Katz and Carl Shapiro. Technology adoption in the presence of network externalities. *The Journal of Political Economy*, 94(4):822–841, 1986. doi: 10.1086/261409.

Anton Korinek and Joseph E Stiglitz. Artificial intelligence and its implications for income distribution and unemployment. In *The economics of artificial intelligence: An agenda*, pages 349–390. University of Chicago Press, 2018.

Anton Korinek and Joseph E Stiglitz. Artificial intelligence, globalization, and strategies for economic development. Technical report, National Bureau of Economic Research, 2021.

Jean-Jacques Laffont and Jean Tirole. *Competition in Telecommunications*. MIT Press, 2000.

William D. Nordhaus. Are we approaching an economic singularity? information technology and the future of economic growth. *American Economic Journal: Macroeconomics*, 13(1): 299–332, 2021. doi: 10.1257/mac.20170105.

Portuguese Competition Authority. Competition and generative artificial intelligence, 2023.

Kalyan K. Sanyal and Ronald W. Jones. The theory of trade in middle products. *American Economic Review*, 72(1):16–31, 1982.

Carl Shapiro. Antitrust in a time of populism. *International Journal of Industrial Organization*, 61:714–748, 2018. doi: 10.1016/j.ijindorg.2018.01.001.

Joseph E Stiglitz. Monopoly and the rate of extraction of exhaustible resources. *The American Economic Review*, 66(4):655–661, 1976.

Wolfgang F. Stolper and Paul A. Samuelson. Protection and real wages. *The Review of Economic Studies*, 9(1):58–73, November 1941. doi: 10.2307/2967638. URL https://doi.org/10.2307/2967638.

Khushita Vasant. Duckduckgo, perplexity keen to buy google chrome if divested, us judge hears, April 2025. URL https://www.mlex.com/mlex/articles/2329582/duckduckgo-perplexity-keen-to-buy-google-chrome-if-divested-us-judge-hears. Accessed: 2025-09-18.

ONLINE APPENDIX

1 Longer Literature Review

Our framework connects to multiple strands of research about automation and offshoring, industrial organization, and several other literatures in macroeconomics and trade.

Automation and Technology. Autor et al. [2003] provided an early theoretical task-based framework: computers substitute for routine tasks but complement non-routine ones, implying wage polarization. Acemoglu and Restrepo [2018] modeled automation as the removal of tasks; automation displaces labor in those tasks but may create new ones, so net labor demand is ambiguous. Their later empirical study (Acemoglu and Restrepo [2020]) showed that exposure to industrial robots reduced employment and wages in U.S. commuting zones. Autor and Salomons [2018] used cross-country data to argue that productivity gains from technology often expand output enough to offset displacement, sustaining aggregate labor demand. Autor and Thompson [2025] combined theory and empirics to show that outcomes depend on which tasks are automated: removing inexpert tasks raises expertise thresholds, increasing wages but lowering employment, while removing expert tasks reduces expertise requirements, lowering wages but expanding employment. In contrast to these frameworks, our models emphasize endogenous input pricing and adoption in general equilibrium, which can generate more complex welfare effects (such as reduced welfare across workers groups) than when technology costs and goods prices are fixed. Korinek and Stiglitz [2018] emphasize distribution under *incomplete insurance* and *costly redistribution*, treating innovation (including AI) as a shock that private insurance cannot address, so that some groups lose despite efficiency gains.

Offshoring. Grossman and Rossi-Hansberg [2008] developed a model where firms perform a continuum of tasks, and some can be moved abroad at lower cost. Offshoring acts like a productivity improvement: it reduces production costs, expands output, and has Stolper–Samuelson-type (Stolper and Samuelson [1941]) wage effects as tasks intensive in one factor shift abroad. Because the framework has only traded goods and takes offshoring costs as parametric, it cannot capture channels that arise when non-tradables constrain the CPI or when input costs are set strategically. Subsequent empirical work confirms cost reductions and distributional impacts, but similarly treats offshore costs as exogenous. Our

work differs by incorporating strategic pricing by the input provider and discrete technology adoption decisions by domestic firms.

Korinek and Stiglitz [2021] analyze directed technical change in competitive open-economy GE with *traded* final goods and domestically immobile primary factors (labor, capital, and sometimes a non-accumulable resource). With no non-traded CPI block, no imported intermediate factor of production, and no variety/entry margin, incidence is organized by *terms-of-trade* (ToT) movements and factor-price discipline. Labor-saving technology depresses the world's relative demand for labor; resource-saving technology compresses the return to the scarce resource; and the scarce factor can appropriate gains. Our Model 1 is closest in spirit (all traded, competitive pricing). By contrast, Models 2–3 add the non-traded CPI channel and an *imported, priced AI input* that creates leakage $p_X Q_B$, as well as a differentiated B sector with free entry—margins central for AI that are absent in their setup. The p_X is analogous to ToT movements, but we add strategic, potentially non-linear pricing in a setting with multiple channels for restructuring industries and affecting income distribution.

Dutch disease and non-tradables. Corden and Neary [1982] and Corden [1984] showed how a positive resource shock or favorable terms-of-trade shift reallocates resources toward the booming sector and raises the price of non-tradables. The main consequence is distributional: some factors gain, while others lose from higher non-traded prices. In our setting, the mechanism is structurally similar—the key channel is again the pressure that traded-sector shocks place on non-tradables—but the nature of the shock is different. Rather than a free endowment gain, we study a foreign intermediate that must be purchased, often at a monopoly markup. This distinction matters: the Dutch disease models predict aggregate income rises even if some groups are hurt, while in our framework national income can fall because part of the gains leak abroad as payments to the input supplier.

Market power and factor prices. A separate line of work studies how imperfect competition shapes the distribution of income. Farboodi et al. [2019] and Shapiro [2018] emphasize how market power in digital technologies and AI adoption can distort innovation and increase inequality. Classic trade models with intermediates typically assume input prices are set competitively on world markets, so factor prices move only through technology or endowment shocks. Our framework differs in making the cost of the new input endogenous to a foreign supplier with market power. The result is that wage and welfare effects depend not only on the direction of technological change, as in the prior literature, but also on how rents are extracted by the input supplier.

Industrial organization and licensing. The IO literature studies how technology owners design tariffs when licensing innovations. Katz and Shapiro [1985, 1986] and Kamien and Tauman [1986] show that, under standard conditions, a monopolist sets the per-unit royalty at zero and extracts rents through a fixed fee: this maximizes adoption and total output while transferring surplus. Similarly, in network and access pricing models Armstrong [2006], Laffont and Tirole [2000], two-part tariffs are efficient on the margin, with rent extraction handled through the fixed component. In our framework, this benchmark breaks down due to the setup of the model, where with free entry, CES demand and general equilibrium, marginal costs do not directly affect the overall profits in the productive sector, and because access and usage fees feed back into monopoly profits through income. Our departure from the standard IO insight (zero per-unit royalty with extraction via a fixed fee) stems from GE and free-entry feedback: both p_X and ϕ distort the size of the pie by lowering Y and N, while the adoption frontier ties feasible usage pricing to displaced workers' outside options. This makes *both* instruments behaviorally and welfare-relevant in our setting.

Other trade models with intermediates. Sanyal and Jones [1982] modeled trade as the exchange of intermediate "middle products," showing how this reallocates factors across sectors. Ethier [1982] modeled intermediates with love-of-variety, where more imported inputs raise productivity and welfare through scale effects. Both assume competitive world markets where input prices are exogenous. In that setting, more varieties or lower costs always raise national income, and factor price movements follow directly from comparative advantage. By contrast, in our framework the key input is supplied by a foreign monopolist that sets its own price. This shifts the analysis from a pure comparative-advantage story to one where input prices and adoption levels are strategic, and part of the gains from new intermediates can leak abroad or fail to materialize.

2 Formal Model 1: Open Economy, Two Traded Goods

This is a baseline model with two traded goods, a small open economy where consumption is not modeled.

2.1 Setup, accounting, and objects

Goods, prices, and factors. The economy consists of two traded final goods A, B sold at exogenous world prices $(\bar{p}_A, \bar{p}_B)$ (small open economy) and produced with two primary factors: skilled labor S and unskilled labor U, with endowments (L_S, L_U) and factor prices

(w_S, w_U). A foreign input X is used *only* by sector B at price p_X set abroad. Technologies are CRS; markets are competitive.

Unit costs, unit input coefficients, and zero profit. Let $c_A(w_S, w_U)$ and $c_B(w_S, w_U, p_X)$ be unit-cost functions. Zero-profit (unit-cost) equalities:

$$\bar{p}_A = c_A(w_S, w_U), \qquad \bar{p}_B = c_B(w_S, w_U, p_X). \tag{ZP}$$

Define *unit input coefficients* (Hicksian, i.e., per unit of output):

$$a_S^A := \frac{\partial c_A}{\partial w_S}, \quad a_U^A := \frac{\partial c_A}{\partial w_U}; \qquad a_S^B := \frac{\partial c_B}{\partial w_S}, \quad a_U^B := \frac{\partial c_B}{\partial w_U}, \quad a_X^B := \frac{\partial c_B}{\partial p_X}.$$

Define cost shares: $\theta_S^j = \frac{w_S a_S^j}{\bar{p}_j}$, $\theta_U^j = \frac{w_U a_U^j}{\bar{p}_j}$, and in B also $\theta_X^B = \frac{p_X a_X^B}{\bar{p}_B}$, with $\theta_S^A + \theta_U^A = 1$ and $\theta_S^B + \theta_U^B + \theta_X^B = 1$.

Full employment.

$$a_S^A Q_A + a_S^B Q_B = L_S, \qquad a_U^A Q_A + a_U^B Q_B = L_U. \tag{FE}$$

Imports and national income (GNP at world prices). The quantity of X imported is $M_X = a_X^B Q_B$; the import bill is $p_X M_X$.

$$Y \;=\; \bar{p}_A Q_A + \bar{p}_B Q_B \;-\; p_X M_X \;=\; w_S L_S + w_U L_U. \tag{NI}$$

The first equality is value-added (final output value minus imported intermediates); the second uses zero profits in A, B. *Envelope in p_X:*

$$\frac{dY}{dp_X} \;=\; -\,M_X \;\leq 0. \tag{Y–p_X}$$

Hence a fall in p_X raises national income (aggregate welfare at traded-goods prices), while a rise lowers it.

Trade and consumption. We do not model home demand; production and factor prices are determined by (ZP)–(FE). The trade-balance identity is $\bar{p}_A(C_A - Q_A) + \bar{p}_B(C_B - Q_B) + p_X M_X = 0$; the country can export B ($Q_B > C_B$). Welfare statements use Y (equivalently, any homothetic utility at prices $(\bar{p}_A, \bar{p}_B)$).

2.2 Comparative statics at a given p_X

Define the *relative factor-intensity indices*:

$$R_A := \frac{\theta_S^A}{\theta_U^A}, \qquad R_B(p_X) := \frac{\theta_S^B}{\theta_U^B}.$$

We say B is *skill-intensive (relative to A)* if $R_B(p_X) > R_A$, and *unskilled-intensive* if $R_B(p_X) < R_A$.

Totally differentiating (ZP) at fixed $(\bar{p}_A, \bar{p}_B)$ yields:

$$\left(\theta_S^B - \theta_U^B \, \tfrac{\theta_S^A}{\theta_U^A}\right) dw_S = -\,\theta_X^B \, dp_X, \qquad dw_U = -\,\tfrac{\theta_S^A}{\theta_U^A} \, dw_S. \tag{L}$$

Implications. (i) Exactly one worker type wins locally (the two wages move in opposite directions). (ii) Who wins locally depends *only* on the current intensity ordering:

$$\begin{cases} R_B(p_X) > R_A \ (B \text{ skill-intensive}) \ \Rightarrow \ \dfrac{dw_S}{dp_X} < 0, \ \dfrac{dw_U}{dp_X} > 0, \\ R_B(p_X) < R_A \ (B \text{ unskilled-intensive}) \Rightarrow \ \dfrac{dw_S}{dp_X} > 0, \ \dfrac{dw_U}{dp_X} < 0. \end{cases} \tag{Signs}$$

(iii) Aggregate income moves monotonically by (Y–p_X): $dp_X \downarrow$ raises Y, $dp_X \uparrow$ lowers Y.

2.3 Strong substitution and intensity reversals

Strong substitution (SS) condition. We say X is a *strong substitute* for U in B if the Hicksian unit demand satisfies

$$\begin{gathered} a_U^B(w_S, w_U, p_X) \text{ is strictly decreasing in } p_X^{-1}, \\ \text{and for sufficiently low } p_X \text{ we have } a_U^B = 0. \end{gathered} \tag{SS}$$

SS is *not* imposed for the general comparative statics; it is used to describe how $R_B(p_X)$ *can* evolve as p_X falls, by pushing θ_U^B down and thus $R_B = \theta_S^B/\theta_U^B$ up.

Single intensity reversal. A *single intensity reversal* occurs at some $p_X^* \in (0, \infty)$ if

$$R_B(p_X) \begin{cases} < R_A, & p_X > p_X^*, \\ = R_A, & p_X = p_X^*, \\ > R_A, & p_X < p_X^*. \end{cases} \tag{Rev}$$

(SS) is a *sufficient* force that can generate (Rev) (by making B increasingly skill-intensive as X gets cheaper), but (Rev) does not require SS per se.

2.4 Autarky $\to$ low p_X and divergence near a reversal

Fix a low input price p_X^ℓ and compare to autarky in X ($p_X = \infty$). The *level* change in the unskilled wage is

$$w_U(p_X^\ell) - w_U(\infty) \;=\; -\int_{p_X^\ell}^{\infty} \frac{dw_U}{dp_X}\, dp_X,$$

where the integrand's sign at each p_X is given by (Signs). The *local* effect at p_X^ℓ is sign $\left(\frac{dw_U}{dp_X}\big|_{p_X^\ell}\right)$ and depends only on the current ordering $R_B(p_X^\ell)$ vs. R_A.

Proposition 1 (Single intensity reversal)**.** *Suppose p_X^* satisfies* (Rev)*. Then in neighborhoods of p_X^*:*

- *If $p_X^\ell \downarrow p_X^*$ from above (so $R_B(p_X^\ell) < R_A$), then*

$$w_U(p_X^\ell) - w_U(\infty) \;>\; 0 \quad \textit{and} \quad \frac{dw_U}{dp_X}\Big|_{p_X^\ell} \;<\; 0.$$

  Thus introducing a cheap X is good *for unskilled, and a marginal increase in p_X at that point is* bad *for unskilled.*

- *If $p_X^\ell \uparrow p_X^*$ from below (so $R_B(p_X^\ell) > R_A$), then $w_U(p_X^\ell) - w_U(\infty) > 0$ while $\frac{dw_U}{dp_X}\big|_{p_X^\ell} > 0$: opening to cheap X is* good *for unskilled, and a marginal increase in p_X is* good *for them. That implies that the opening is* bad *for skilled, and a marginal increase in p_X is also bad for skilled workers.*

In all cases, aggregate income moves with p_X by (Y–p_X).

Proposition 2 (Double-hurt in the all-traded model with non-monotone skill intensities)**.** *Suppose there exist thresholds $p_X^{(1)} > p_X^{(2)} > 0$ such that a double-reversal holds*

$$R_B(p_X) \begin{cases} < \; R_A, & p_X > p_X^{(1)} \qquad \textit{(very high } p_X\textit{: } B \textit{ unskilled-intensive)}, \\ > \; R_A, & p_X^{(2)} < p_X < p_X^{(1)} \qquad \textit{(intermediate } p_X\textit{: } B \textit{ skill-intensive)}, \\ < \; R_A, & p_X < p_X^{(2)} \qquad \textit{(very low } p_X\textit{: } B \textit{ unskilled-intensive again)}. \end{cases} \tag{TwoX}$$

Then it is possible that for $p_X^\ell \in (0, p_X^{(2)})$ sufficiently close to $p_X^{(2)}$,

$$\underbrace{w_U(p_X^\ell) - w_U(\infty)}_{\text{level vs. autarky}} < 0 \qquad \text{and} \qquad \underbrace{\frac{dw_U}{dp_X}\Big|_{p_X^\ell}}_{\text{local at } p_X^\ell} < 0.$$

Hence introducing (very) cheap X makes unskilled worse off in levels and *a marginal increase in p_X at p_X^ℓ also makes them worse off locally.*

Sketch. By (Signs), sign $\left(\frac{dw_U}{dp_X}\right)$ equals the sign of $R_B(p_X) - R_A$. Under (TwoX), the derivative is positive on the intermediate band $(p_X^{(2)}, p_X^{(1)})$ and negative on the tails $(p_X^{(1)}, \infty)$ and $(0, p_X^{(2)})$. The level change is the path integral $w_U(p_X^\ell) - w_U(\infty) = \int_\infty^{p_X^\ell} \frac{dw_U}{dp_X}\, dp_X$, whose sign is the net of those regions. For p_X^ℓ close to $p_X^{(2)}$, the (large) positive-derivative region in $(p_X^{(2)}, p_X^{(1)})$ dominates, making the integral negative; at $p_X^\ell < p_X^{(2)}$, we have $R_B < R_A$ so the local derivative is negative. □

Example of Double Reversal

Technologies. We specify Sector A technology as a single CRS technique T_A (Leontief or CES) and B as a *menu of techniques* T_k, $k \in \{1, 2, 3\}$, with CRS unit input vectors (a_S^k, a_U^k, a_X^k). The operative input vector in B at (w, p_X) is $k^\star \in \arg\min_k \{w_S a_S^k + w_U a_U^k + p_X a_X^k\}$, hence

$$R_B(p_X; \omega) = \omega \cdot \frac{a_S^{k^\star}}{a_U^{k^\star}} \quad \text{and} \quad \theta_X^B = \frac{p_X a_X^{k^\star}}{c^{k^\star}(w, p_X)}.$$

Lemma 1 (Cost functions inherit CRS technology properties)**.** *Let Sector A technology be CRS, either Leontief (a_S^A, a_U^A) or CES with elasticity $\sigma \in (0, \infty)$. Let Sector B technology be CRS with a finite menu of techniques $\{T_k\}_{k=1}^3$, each specified by nonnegative coefficients (a_S^k, a_U^k, a_X^k). Then:*

The unit–cost functions

$$c_A(w_S, w_U) = \begin{cases} a_S^A w_S + a_U^A w_U & \textit{(Leontief)}, \\ \left[\alpha w_S^{1-\sigma} + (1-\alpha) w_U^{1-\sigma}\right]^{\frac{1}{1-\sigma}} & \textit{(CES)}, \end{cases}$$

$$c_B(w_S, w_U, p_X) = \min_k \Big\{ w_S a_S^k + w_U a_U^k + p_X a_X^k \Big\}.$$

are linearly homogeneous, concave, and strictly increasing in each input price.

Lemma 2 (Compact wage-ratio band)**.** *Suppose c_A, c_B are strictly increasing, concave and linearly homogeneous in (w_S, w_U) for each fixed $p_X > 0$, as in our working example. Fix*

$(\bar{p}_A, \bar{p}_B)$ *and restrict* p_X *to a compact interval* $[p^{\min}, p^{\max}]$ *with* $0 < p^{\min} < p^{\max} < \infty$*. Then the equilibrium wage ratio* $\omega(p_X)$ *solving* $\bar{p}_A = c_A$ *and* $\bar{p}_B = c_B(\cdot, p_X)$ *lies in a compact interval* $\Omega := [\underline{\omega}, \overline{\omega}] \subset (0, \infty)$*, and* $\omega(p_X)$ *is continuous on* $[p^{\min}, p^{\max}]$*.*

Sketch. By duality and homogeneity, each zero-profit locus is a continuous, strictly monotone curve in (w_S, w_U); intersecting them for each $p_X \in [p^{\min}, p^{\max}]$ yields a unique (w_S, w_U) and hence a continuous $\omega(p_X)$. Monotonicity and positive prices bound the locus away from axes, giving $\underline{\omega} > 0$ and $\overline{\omega} < \infty$. □

Assumption 1 (Sector-A intensity band)**.** *There exist CRS* T_A *and world prices* $(\bar{p}_A, \bar{p}_B)$ *such that along the wage-ratio band* Ω *the (endogenous) sector-A intensity ratio* $R_A(\omega)$ *lies in an interval* $I_A := [\underline{R}_A, \overline{R}_A] \subset (0, \infty)$*.*

[Three techniques in B] Choose three CRS techniques for B with unit inputs

$$T_1 : (a_S, a_U, a_X) = (s_1, u_1, 0), \qquad T_2 : (s_2, u_2, x_2), \qquad T_3 : (s_3, u_3, x_3),$$

satisfying, for all $\omega \in \Omega$,

$$\boxed{\omega \frac{s_1}{u_1} \;<\; \underline{R}_A \;\le\; R_A(\omega) \;\le\; \overline{R}_A \;<\; \omega \frac{s_2}{u_2} \quad \text{and} \quad \omega \frac{s_3}{u_3} \;<\; \underline{R}_A \,.} \tag{B-order}$$

(Thus T_1 and T_3 are *unskilled*-intensive relative to A, while T_2 is *skill*-intensive.) Pick (x_2, x_3) so that the $T_1 \leftrightarrow T_2$ and $T_2 \leftrightarrow T_3$ switch prices

$$p_{12}(\omega) = \frac{w_S(s_1 - s_2) + w_U(u_1 - u_2)}{x_2}, \qquad p_{23}(\omega) = \frac{w_S(s_2 - s_3) + w_U(u_2 - u_3)}{x_3 - x_2},$$

satisfy $p_{12}(\omega) > p_{23}(\omega)$ for all $\omega \in \Omega$, with a uniform gap $p_{12} - p_{23} \ge \delta > 0$.

Lemma 3 (Technique switch prices and their ordering)**.** *Fix three CRS techniques in Sector B with unit inputs* $T_1 : (s_1, u_1, 0)$*,* $T_2 : (s_2, u_2, x_2)$*, and* $T_3 : (s_3, u_3, x_3)$*, where all coefficients are nonnegative and* $x_3 > x_2 > 0$*. For any wage vector* $w = (w_S, w_U) \in (0, \infty)^2$*, define the* switch prices $p_{12}(w)$ *and* $p_{23}(w)$ *as the values of* p_X *that solve*

$$c^{(1)}(w) = c^{(2)}(w, p_X) \quad \textit{and} \quad c^{(2)}(w, p_X) = c^{(3)}(w, p_X),$$

respectively, where $c^{(k)}(w, p_X) = w_S s_k + w_U u_k + p_X x_k$ *(with* $x_1 \equiv 0$*). Then:*

1. *The switch prices exist uniquely and are given by*

$$p_{12}(w) \;=\; \frac{w_S(s_1 - s_2) + w_U(u_1 - u_2)}{x_2}, \qquad p_{23}(w) \;=\; \frac{w_S(s_2 - s_3) + w_U(u_2 - u_3)}{x_3 - x_2}. \tag{29}$$

2. *Writing the wage ratio $\omega := w_S/w_U$ and factoring out w_U, we have*

$$p_{12}(w) \;=\; w_U \cdot \underbrace{\frac{\omega(s_1 - s_2) + (u_1 - u_2)}{x_2}}_{=:\ \pi_{12}(\omega)}, \qquad p_{23}(w) \;=\; w_U \cdot \underbrace{\frac{\omega(s_2 - s_3) + (u_2 - u_3)}{x_3 - x_2}}_{=:\ \pi_{23}(\omega)}. \tag{30}$$

Consequently, the ordering $p_{12}(w) > p_{23}(w)$ *depends only on ω (not on the absolute scale of wages): it is equivalent to $\pi_{12}(\omega) > \pi_{23}(\omega)$.*

3. *(Uniform gap over a wage-ratio band.) Let $\Omega \subset (0, \infty)$ be a compact interval that contains the equilibrium wage ratio $\omega(p_X)$ for all p_X in the policy-relevant range (cf. Lemma 2). If there exists $\delta_\pi > 0$ such that*

$$\inf_{\omega \in \Omega} \big\{ \pi_{12}(\omega) - \pi_{23}(\omega) \big\} \;\geq\; \delta_\pi, \tag{31}$$

then $p_{12}(w) - p_{23}(w) \;\geq\; w_U\, \delta_\pi$ for all wage pairs with $\omega \in \Omega$. In particular, if w_U is bounded below on the equilibrium locus, $w_U \geq \underline{w}_U > 0$, then the absolute *switch-price gap is uniformly positive:*

$$p_{12}(w) - p_{23}(w) \;\geq\; \underline{w}_U\, \delta_\pi \;=:\; \delta \;>\; 0. \tag{32}$$

Proof. (1) Solving $c^{(1)}(w) = c^{(2)}(w, p_X)$ gives $w_S s_1 + w_U u_1 = w_S s_2 + w_U u_2 + p_X x_2$, hence (29) for p_{12} since $x_2 > 0$. Likewise, $c^{(2)}(w, p_X) = c^{(3)}(w, p_X)$ yields p_{23} with denominator $x_3 - x_2 \neq 0$ by assumption. (2) Substitute $w_S = \omega w_U$ and factor out w_U to obtain (30); both p_{12} and p_{23} are proportional to w_U, so their ordering is governed by $\pi_{12}(\omega)$ vs. $\pi_{23}(\omega)$. (3) Under (31), for any $\omega \in \Omega$ we have $\pi_{12}(\omega) - \pi_{23}(\omega) \geq \delta_\pi$. Multiplying by w_U gives $p_{12} - p_{23} \geq w_U \delta_\pi$, and if $w_U \geq \underline{w}_U$ along the equilibrium locus, then (32) follows. □

Remark 1. *The numerators in* (30) *are linear in ω; the denominators are strictly positive by construction ($x_2 > 0$ and $x_3 > x_2$). Condition* (31) *is a simple linear-inequality requirement over the band Ω:*

$$\frac{\omega(s_1 - s_2) + (u_1 - u_2)}{x_2} \;-\; \frac{\omega(s_2 - s_3) + (u_2 - u_3)}{x_3 - x_2} \;\geq\; \delta_\pi \quad \textit{for all } \omega \in \Omega.$$

Given (s_k, u_k), this can be ensured by picking x_2, x_3 (with $x_3 > x_2 > 0$) so that the right-hand side stays uniformly positive on Ω; geometrically, the $T_1 \leftrightarrow T_2$ threshold is to the right *of the $T_2 \leftrightarrow T_3$ threshold throughout the wage-ratio band.*

Lemma 4 (Two crossings are robust)**.** *Under Lemma 2, Assumption 1, and Construction 2.4, the induced path $R_B(p_X; \omega(p_X))$ crosses $R_A(\omega(p_X))$ exactly twice as p_X falls from*

$p^{\max}$ *to* $p^{\min}$*:*

high p_X : T_1 $(R_B < R_A)$ $\rightarrow$ *intermediate* p_X : T_2 $(R_B > R_A)$ $\rightarrow$ *very low* p_X : T_3 $(R_B < R_A)$.

Moreover, the ordering $p_{12}(\omega) > p_{23}(\omega)$ *and the inequalities in* (B-order) *persist for all* $\omega \in \Omega$*; hence the two crossings are preserved by all wage movements along the GE path.*

Proposition 3 (Endogenous double-harm: existence and openness)**.** *There exists a nonempty open set of primitives (world prices* $(\bar{p}_A, \bar{p}_B)$*, endowments* (L_S, L_U)*, sector-A technology* T_A*, and sector-B menu* $\{T_1, T_2, T_3\}$*) such that:*

1. $R_A(\omega(p_X)) \in I_A$ *for all* $p_X \in [p^{\min}, p^{\max}]$ *and* $R_B(p_X; \omega(p_X))$ *follows the low→high→low pattern in Lemma 4.*
2. *For any* $p_X^\ell \in (p_{23}^{\max} - \varepsilon, p_{23}^{\max})$ *(sufficiently close to the lower switch) we have*

$$w_U(p_X^\ell) - w_U(\infty) \;<\; 0 \qquad \textit{and} \qquad \left.\frac{dw_U}{dp_X}\right|_{p_X^\ell} \;<\; 0,$$

i.e. the "double-harm" (global and local losses) in the all-traded, nominal *model.*

Sketch. (1) follows from Lemma 2, (B-order), and continuity: the inequalities hold for all $\omega \in \Omega$, hence for $\omega(p_X)$. The uniform gap δ keeps the threshold order under small wage movements; openness follows by continuity of all objects. (2) By (Signs), $\frac{dw_U}{dp_X} < 0$ when $R_B < R_A$; this holds at $p_X^\ell < p_{23}$. The level change equals $-\int_{p_X^\ell}^{\infty} \frac{dw_U}{dp_X} dp_X$, which is dominated by the intermediate band where $R_B > R_A$ and the integrand is positive; hence the effect is negative near the lower switch. □

Example with specific parameters Pick $\Omega = [0.8, 1.5]$. Let T_A be Leontief with $(a_S^A, a_U^A) = (0.50, 0.45)$ so that $R_A(\omega) = \omega \cdot (0.50/0.45) \in [0.89, 1.67]$. For B choose

$$T_1 : (0.40, 1.20, 0), \quad T_2 : (1.00, 0.40, 0.60), \quad T_3 : (0.15, 0.45, 3.50).$$

Then, for all $\omega \in [0.8, 1.5]$,

$$R_B^{(1)}(\omega) = 0.333\omega < 0.89 \le R_A(\omega) \le 1.67 < 2.5\omega = R_B^{(2)}(\omega), \qquad R_B^{(3)}(\omega) = 0.333\omega < 0.89,$$

so (B-order) holds on Ω. With a reference wage pair $w_S = w_U = 1$, the switch prices are $p_{12} = 0.333$ and $p_{23} = 0.276$; they remain ordered with a positive gap when (w_S, w_U) move within a neighborhood implied by Lemma 2. Proposition 3 then applies.

2.5 Output responses in the general comparative statics of p_X

Stack the full-employment system (FE) as $A\,Y = \bar{F}$ with

$$A := \begin{bmatrix} a_S^A & a_S^B \\ a_U^A & a_U^B \end{bmatrix}, \qquad Y := \begin{bmatrix} Q_A \\ Q_B \end{bmatrix}, \qquad \bar{F} := \begin{bmatrix} L_S \\ L_U \end{bmatrix}.$$

Totally differentiate at fixed endowments:

$$A\,dY \;=\; -\,(dA)\,Y \quad\Longrightarrow\quad dY \;=\; -\,A^{-1}(dA)\,Y. \tag{Mix}$$

With $A^{-1} = \frac{1}{\Delta}\begin{bmatrix} a_U^B & -a_S^B \\ -a_U^A & a_S^A \end{bmatrix}$ where $\Delta := a_S^A a_U^B - a_S^B a_U^A > 0$ (distinct factor intensities), the B-output response is

$$dQ_B \;=\; \frac{1}{\Delta}\Big\{\, Q_A\big(a_U^A\, da_S^A - a_S^A\, da_U^A\big) \;+\; Q_B\big(a_U^A\, da_S^B - a_S^A\, da_U^B\big) \,\Big\}. \tag{$\star$}$$

Here each da_i^j is a Hicksian coefficient response induced by the endogenous wage changes from (L) and the direct p_X-effect in B.

What can be signed in general. Let $R_B \lessgtr R_A$ denote the local intensity ordering. Under standard Hicksian regularity (negative own-price effects, symmetric substitution) and mild cross-substitution assumptions, we obtain:

Proposition 4 (Output of B as p_X changes)**.** *Fix a point p_X and suppose: (i) B is unskilled-intensive locally, $R_B < R_A$; (ii) X is a (weak) Hicksian substitute for U in B, so $\partial a_U^B/\partial p_X > 0$ (equivalently, lowering p_X reduces a_U^B); (iii) own-price effects are negative for A and B (e.g., $\partial a_U^A/\partial w_U < 0$, $\partial a_S^A/\partial w_S < 0$). Then, for a small change in p_X,*

$$\frac{dQ_B}{dp_X} \;<\; 0.$$

If instead $R_B > R_A$ (skill-intensive B), the sign of $\frac{dQ_B}{dp_X}$ is a priori ambiguous *without further structure (e.g., on $\partial a_S^B/\partial p_X$ and cross-substitution magnitudes).*

Proof sketch. Plug the wage responses from (L) (for $R_B < R_A$, we have $dw_U/dp_X < 0$, $dw_S/dp_X > 0$) into the Hicksian responses da_i^j and evaluate ($\star$). Under (iii), $da_U^A/dp_X > 0$ and $da_S^A/dp_X < 0$ via the induced wage changes. Under (ii), $da_U^B/dp_X > 0$; da_S^B/dp_X is not needed to be signed if it does not overturn the two summands in ($\star$). Because $a_U^A/a_S^A < R_A^{-1} < R_B^{-1} = a_U^B/a_S^B$ (unskilled intensity of B), the bracketed terms deliver a negative total in ($\star$). □

Interpretation. When B is unskilled-intensive, a rise in p_X makes B less competitive and, via (L), lowers w_U and raises w_S. Hicksian coefficients then adjust so that each unit uses *more* U in B and (through wages) more U in A as well; to clear factor markets at fixed endowments, the output mix must tilt *away* from B (hence $dQ_B/dp_X < 0$). Conversely, for $dp_X < 0$, Q_B expands ($dQ_B > 0$). If B is skill-intensive, the induced coefficient changes pull in opposing directions and the sign becomes model-specific.

Remark (role of SS). Proposition 4 uses only the *weak* requirement $\partial a_U^B/\partial p_X > 0$ (Hicksian substitution between X and U). The *strong substitution* assumption (SS) is *sufficient* for this inequality and, separately, implies that as p_X falls the sectoral intensity ratio $R_B(p_X)$ increases; this can generate an intensity reversal (Rev) and hence a divergence between local and level conclusions (Section 4).

3 Formal Model 2: Model 1 with One Non-Traded Good

We now make sector A non-tradeable and close the model with Cobb–Douglas preferences over (A, B). This endogenizes the domestic relative price $p_A/\bar{p}_B$ and introduces a demand/price-index channel that interacts with the income and output-mix effects analyzed above.

3.1 Setup and market closure

Preferences and demand. A representative household has utility

$$U(C_A, C_B) = C_A^{1-\alpha} C_B^{\alpha}, \qquad \alpha \in (0, 1),$$

facing the price vector $(p_A, \bar{p}_B)$. Given nominal income Y, Cobb–Douglas demand implies

$$C_A = \frac{(1-\alpha)Y}{p_A}, \qquad C_B = \frac{\alpha Y}{\bar{p}_B}.$$

The exact consumption price index is $P = p_A^{1-\alpha} \bar{p}_B^{\alpha}$.

Non-traded market clearing and the p_A closure. Because A is non-traded, $C_A = Q_A$. Hence the equilibrium price of the non-traded good is

$$\boxed{p_A = \frac{(1-\alpha)\,Y}{Q_A}} \qquad \Longleftrightarrow \qquad d\ln p_A = d\ln Y - d\ln Q_A. \tag{NT-P}$$

Intuitively, demand pressure from higher income ($Y \uparrow$) and production scarcity in A ($Q_A \downarrow$) both bid up p_A.

3.2 Equilibrium

Indirect utility and a two-channel welfare map. With CD, indirect utility satisfies

$$\ln U = \text{const} + \ln Y - (1-\alpha)\ln p_A - \alpha \ln \bar{p}_B.$$

Using (NT-P), the welfare differential is

$$\boxed{d\ln U = \alpha\, d\ln Y + (1-\alpha)\, d\ln Q_A}. \tag{W}$$

For upstream shocks to p_X,

$$\frac{d\ln U}{dp_X} = -\alpha\frac{M_X}{Y} \;+\; (1-\alpha)\frac{1}{Q_A}\frac{dQ_A}{dp_X}, \tag{W–p_X}$$

combining the effect on income (applying the envelope theorem, $dY/dp_X = -M_X$) with the output response of A.

Real wages. Let $i \in \{S, U\}$. The real consumption wage is w_i/P. Because $d\ln P = (1-\alpha)\, d\ln p_A = (1-\alpha)(d\ln Y - d\ln Q_A)$,

$$\boxed{d\ln\!\left(\tfrac{w_i}{P}\right) = d\ln w_i - (1-\alpha)\, d\ln p_A = d\ln w_i - (1-\alpha)\,(d\ln Y - d\ln Q_A)}. \tag{RW}$$

Thus nominal Stolper–Samuelson–type changes from (L) are filtered through a cost-of-living term driven by the non-traded price.

Outputs and factor markets. Production, factor prices, and Hicksian coefficients continue to be determined by (ZP)–(FE). Output responses follow

$$dY = -A^{-1}(dA)Y, \qquad dQ_B \text{ as in } (\star),$$

and where dQ_A implied by factor balance and Hicksian responses so that signs can be tracked by the same local intensity and substitution conditions used in Prop. 4.

3.3 Comparative statics: entry, reversals, and incidence

We interpret "entry" as a drop in the foreign upstream price p_X to a low level due to the arrival of an efficient supplier.

Low-price entry ($dp_X < 0$). Under the conditions of Prop. 4 with $R_B < R_A$ (sector B locally unskilled-intensive) and X a Hicksian substitute for U in B:

$$dp_X < 0 \;\Rightarrow\; dY > 0, \quad dQ_B > 0, \quad dQ_A < 0, \quad dp_A > 0.$$

Welfare then moves according to (W–p_X): the income gain $\alpha(-M_X/Y) > 0$ is partially offset by the non-traded scarcity channel $(1-\alpha)\, d\ln Q_A < 0$ (a "Dutch disease" effect).

Can both worker types be hurt locally by entry? Nominally, at fixed traded-good prices one wage rises and the other falls (by (Signs)). With a non-traded good, however, *real* wages can both fall if the cost-of-living term dominates:

$$\exists\, dp_X < 0 : \quad d\ln(w_S/P) < 0, \quad d\ln(w_U/P) < 0, \tag{BH}$$

even though one nominal wage rose. Sufficient conditions are a large α and a sharp contraction of Q_A (hence $dp_A \gg 0$) relative to the nominal gain of the winning factor.

Reversals at very low entry prices. Two reversals are possible as p_X falls:

1. An *intensity reversal* $R_B(p_X)$ crossing R_A (as in (Rev)), flipping the *nominal* incidence in (Signs).
2. A *welfare reversal* from the non-traded channel: even though $dY/dp_X = -M_X < 0$ everywhere, the sign of $d\ln U/dp_X$ in (W–p_X) can flip depending on dQ_A/dp_X.

Proposition 5 (Propagation of income contractions through the non-traded margin)**.** *Consider a marginal increase $dp_X > 0$ (an upstream squeeze). Then:*

1. *Income contracts by the envelope: $dY/dp_X = -M_X < 0$.*
2. *The non-traded price moves as $d\ln p_A/dp_X = d\ln Y/dp_X - d\ln Q_A/dp_X$.*

3. *If $R_B < R_A$ and X (weakly) substitutes for U in B (so $dQ_B/dp_X < 0$ by Prop. 4), then factor reallocation raises Q_A, i.e. $dQ_A/dp_X > 0$. Therefore*

$$\frac{d\ln p_A}{dp_X} \;<\; 0, \qquad \frac{d\ln P}{dp_X} = \alpha\,\frac{d\ln p_A}{dp_X} \;<\; 0.$$

4. *Consequently, the welfare effect decomposes as*

$$\frac{d\ln U}{dp_X} = -(1-\alpha)\frac{M_X}{Y} \;+\; \underbrace{\alpha\frac{1}{Q_A}\frac{dQ_A}{dp_X}}_{\textit{cost-of-living relief}} \qquad \in \;\Big(-\tfrac{M_X}{Y},\;0\Big).$$

That is, the income contraction is partially cushioned *by cheaper non-tradables.*

If instead $R_B > R_A$, the sign of dQ_A/dp_X is ambiguous and so is the direction of dp_A; the cushioning may vanish or become amplification.

Corollary 1 (Both workers can be hurt by entry)**.** *Under $R_B < R_A$ and weak substitution, a sufficiently large α and sufficiently elastic p_A (via $Q_A \downarrow$ for $dp_X < 0$) make* (BH) *hold. Equivalently,*

$$(1-\alpha)\Big(\big|\tfrac{d\ln Y}{d\ln p_X}\big| + \big|\tfrac{d\ln Q_A}{d\ln p_X}\big|\Big) \;>\; \max\Big\{\big|\tfrac{d\ln w_S}{d\ln p_X}\big|,\;\big|\tfrac{d\ln w_U}{d\ln p_X}\big|\Big\}$$

is a sufficient local condition for $dp_X < 0$ to reduce both real wages.

Remark. The two elasticities $\frac{M_X}{Y}$ and $\frac{1}{Q_A}\frac{dQ_A}{dp_X}$ are sufficient statistics for the sign and size of $d\ln U/dp_X$ in (W–p_X). Likewise, $\alpha(d\ln Y - d\ln Q_A)$ governs whether both real wages fall in (RW)–(BH).

Proposition 6 (Local–level divergence with a non-traded good)**.** *Suppose p_X^* satisfies an intensity reversal* (Rev)*, and let U denote indirect utility under the Cobb–Douglas closure* (W)*.*

1. *If $p_X^\ell \downarrow p_X^*$ from above (so $R_B(p_X^\ell) < R_A$), then opening to cheap X globally reduces the real wage of one factor group (say unskilled), i.e. $w_U(p_X^\ell)/P < w_U(\infty)/P$. Yet at that same p_X^ℓ, the local comparative static satisfies*

$$\frac{d}{dp_X}\Big(\tfrac{w_U}{P}\Big)\Big|_{p_X^\ell} < 0,$$

so a marginal increase *in p_X would make the already-hurt group worse off again. In words: "entry hurts, and further price rises also hurt."*

2. *Conversely, if $p_X^\ell \uparrow p_X^*$ from below (so $R_B(p_X^\ell) > R_A$), then opening to cheap X globally reduces the real wage of the* other *factor group (say skilled), while at that same p_X^ℓ the local derivative satisfies*

$$\frac{d}{dp_X}\left(\tfrac{w_S}{P}\right)\Big|_{p_X^\ell} < 0,$$

so a marginal increase *in p_X again worsens outcomes for the already-hurt group.*

In both cases, aggregate income continues to satisfy $dY/dp_X = -M_X < 0$, while the welfare effect $d\ln U/dp_X$ may reverse sign depending on dQ_A/dp_X (cf. (W–p_X)). Thus the non-traded channel creates the possibility that one worker group loses globally from entry and also loses further locally from a subsequent upstream squeeze.

4 Formal Model 3: Two Non-traded Sectors

4.1 Price-index and variety identities

Under CES demand in sector B with elasticity $\sigma > 1$ and symmetric equilibrium prices $p_i = p_B = \mu m_B$, the sector-B price index and the free-entry condition imply

$$P_B \;=\; \mu\, m_B\, N^{\frac{1}{1-\sigma}}, \qquad N \;=\; \kappa\,\frac{Y}{F+\phi}, \quad \kappa \equiv \frac{\mu-1}{\mu}\,\alpha, \tag{33}$$

and

$$Q_B \;=\; \frac{\alpha}{\mu}\,\frac{Y}{m_B}. \tag{34}$$

4.2 Welfare ratio with two non-traded sectors

With Cobb–Douglas preferences $U = C_A^{1-\alpha}C_B^{\alpha}$ and both sectors non-traded, market clearing gives $C_A = Q_A$ and $C_B = Q_B$. Using (34), indirect utility can be written (up to a constant) as

$$Q_A^{1-\alpha}\left(\tfrac{\alpha}{\mu}\,\tfrac{Y}{m_B}\right)^{\alpha}. \tag{35}$$

Comparing the AI equilibrium to the no-AI baseline (denoted by superscripts ^{*AI} and ^{*0}, respectively), and normalizing $P_A = 1$, we obtain the welfare ratio (all evaluated at (Z, ϕ, p_X))

$$\frac{W^{*AI}}{W^{*0}} = \underbrace{\left(\frac{m_B^{*0}}{m_B^{*AI}}\right)^{\alpha}}_{\text{unit cost / } B\text{-side price index}} \cdot \underbrace{\left(\frac{N^{*0}}{N^{*AI}}\right)^{\frac{\alpha}{1-\sigma}}}_{\text{variety / entry}} \cdot \underbrace{\left(\frac{Y^{*AI}}{Y^{*0}}\right)^{\alpha}}_{\text{income}}. \tag{36}$$

Using (33) to substitute $N^{*r} = \kappa Y^{*r}/(F + \phi^r)$ with $\phi^0 \equiv 0$ (no access fee paid at the baseline), the variety term becomes

$$\left(\frac{N^{*0}}{N^{*AI}}\right)^{\frac{\alpha}{1-\sigma}} = \left(\frac{Y^{*0}}{Y^{*AI}}\right)^{\frac{\alpha}{1-\sigma}} \left(\frac{F+\phi}{F}\right)^{\frac{\alpha}{1-\sigma}} = \left(\frac{Y^{*AI}}{Y^{*0}}\right)^{\frac{\alpha}{\sigma-1}} \left(\frac{F}{F+\phi}\right)^{\frac{\alpha}{\sigma-1}}. \tag{37}$$

Combining (36) and (37) yields the decomposition

$$\boxed{\frac{W^{*AI}}{W^{*0}} = \left(\frac{m_B^{*0}}{m_B^{*AI}}\right)^{\alpha} \left(\frac{F}{F+\phi}\right)^{\frac{\alpha}{\sigma-1}} \left(\frac{Y^{*AI}}{Y^{*0}}\right)^{\alpha+\frac{\alpha}{\sigma-1}}.} \tag{38}$$

The first term captures how cheaper AI lowers the sector-B price index through unit costs; the second captures how access fees reduce entry/variety; the third folds in income effects, which themselves depend on leakage abroad. Unlike Model 2, the CPI burden falls mainly through P_B (via m_B and N), not through p_A. Note that here we have normalized $p_A = 1$, so the price effects in that sector are captured in changes in relative income. Without the normalization, there would be a fourth term capturing the relative price changes (or, applying equilibrium identities, equivalently the output ratio on Sector A).

4.3 Linear pricing: optimal usage fee at the adoption bound

Set $\phi = 0$ and let the supplier choose $p_X \geq 0$ subject to the adoption constraint $p_X \leq p_{X,\max}^{\text{mode}}(Z)$, where "mode" is either *partial* or *full* automation:

$$p_{X,\max}^{\text{part}}(Z) = u_B\, w_U^{*AI}(Z, 0, p_X), \qquad p_{X,\max}^{\text{full}}(Z) = s_B\, w_S^{*AI}(Z, 0, p_X) + u_B\, w_U^{*AI}(Z, 0, p_X).$$

The supplier's revenue is

$$\Pi^{\text{lin}}(Z, p_X) = p_X\, Q_B^{*AI}(Z, 0, p_X) = \frac{\alpha}{\mu}\, Y^{*AI}(Z, 0, p_X)\, \frac{p_X}{m_B^{*AI}(Z, 0, p_X)}, \tag{39}$$

with $m_B^{*AI}(Z, 0, p_X) = p_X$ (*full*) or $m_B^{*AI}(Z, 0, p_X) = s_B w_S^{*AI}(Z, 0, p_X) + p_X$ (*partial*).

Proposition 7 (Boundary optimality under linear pricing)**.** *Suppose $\phi = 0$ and that along the feasible set $p_X \in [0, p_{X,\max}^{\text{mode}}(Z)]$ the aggregate income is weakly decreasing in p_X:*

$$\frac{dY^{*AI}}{dp_X}(Z, 0, p_X) \;\leq\; 0.$$

Then any interior first-order condition for (39) *cannot deliver a revenue maximum strictly*

below the adoption cap. In particular, under partial automation, if

$$m_B^{*AI}(Z,0,p_X) \;=\; s_B\, w_S^{*AI}(Z,0,p_X) + p_X \;\geq\; p_X, \tag{40}$$

the objective $p_X \mapsto \Pi^{\text{lin}}(Z,p_X)$ is (weakly) increasing on the feasible set, and the optimum is at the boundary $p_X = p_{X,\max}^{\text{part}}(Z)$.

Proof. Differentiate (39):

$$\frac{d\Pi^{\text{lin}}}{dp_X} = \frac{\alpha}{\mu}\left[\frac{p_X}{m_B^{*AI}}\,\frac{dY^{*AI}}{dp_X} + \frac{Y^{*AI}}{(m_B^{*AI})^2}\left(m_B^{*AI} - p_X\,\frac{dm_B^{*AI}}{dp_X}\right)\right].$$

Under *partial* automation, $dm_B^{*AI}/dp_X = 1$ and (40) gives $m_B^{*AI} - p_X \geq 0$; with $dY^{*AI}/dp_X \leq 0$, the first term is non-positive and the second is non-negative, so $d\Pi^{\text{lin}}/dp_X \geq 0$. Hence Π^{lin} is (weakly) increasing until the adoption constraint binds. The same calculation shows that under *full* automation $m_B^{*AI} = p_X$ implies $d\Pi^{\text{lin}}/dp_X = \frac{\alpha}{\mu}\frac{p_X}{m_B^{*AI}}\frac{dY^{*AI}}{dp_X} \leq 0$, so any interior extremum is weakly decreasing in p_X and cannot dominate the boundary where feasibility ends. In both modes, the best feasible p_X is the largest compatible with adoption. □

Remark. The sufficient condition $dY^{*AI}/dp_X \leq 0$ is the natural GE response to a higher upstream usage fee; it holds whenever the induced contractions in B dominate any offsetting variety or A-side effects.

5 Optimal Two-Part Tariffs

Fix primitives Z, elasticity $\sigma > 1$, markup $\mu = \sigma/(\sigma-1)$, and license fee $F > 0$. In the AI regime, with access fee $\phi \geq 0$ and usage fee $p_X \geq 0$, the monopolist chooses

$$\Pi(Z,\phi,p_X) \;=\; \underbrace{\phi\, N^{*AI}(Z,\phi,p_X)}_{\text{access revenue}} \;+\; \underbrace{p_X\, Q_B^{*AI}(Z,\phi,p_X)}_{\text{usage revenue}}, \tag{41}$$

subject to the adoption constraint $p_X \leq p_{X,\max}(\phi)$. Under CES/free entry,

$$N^{*AI} = \kappa\,\frac{Y^{*AI}(Z,\phi,p_X)}{F+\phi}, \quad Q_B^{*AI} = \frac{\alpha}{\mu}\,\frac{Y^{*AI}(Z,\phi,p_X)}{m_B^{*AI}(Z,\phi,p_X)}, \quad \kappa = \tfrac{\mu-1}{\mu}\alpha,$$

with $p_B = \mu m_B^{*AI}$. The adoption frontier $p_{X,\max}(\phi)$ is strictly decreasing and continuously differentiable on its domain.

Assumption 2 (Regularity and monotonicity)**.** *(a) Y^{*AI}, m_B^{*AI}, N^{*AI}, Q_B^{*AI} are continuously differentiable on the feasible set.*

(b) Income and scale are weakly decreasing in fees: $\partial Y^{*AI}/\partial\phi \le 0$, $\partial Y^{*AI}/\partial p_X \le 0$, *and* $\partial m_B^{*AI}/\partial p_X > 0$.
(c) For any ϕ *with adoption feasible at* $p_X = 0$, $Q_B^{*AI}(Z,\phi,0) > 0$.
(d) The reduced-form profit $\widehat{\Pi}(\phi) := \Pi(Z,\phi,p_{X,\max}(\phi))$ *is strictly concave on an interval* Φ *with* $p_{X,\max}(\phi) > 0$ *for* $\phi \in \Phi$.

Proposition 8 (Zero usage fee is never optimal)**.** *For any* ϕ *with adoption feasible at* $p_X = 0$,

$$\left.\frac{\partial \Pi}{\partial p_X}\right|_{(Z,\phi,0)} = Q_B^{*AI}(Z,\phi,0) > 0,$$

so $p_X = 0$ *cannot be optimal.*

Proof. Differentiate Π: $\partial\Pi/\partial p_X = Q_B^{*AI} + p_X\,\partial Q_B^{*AI}/\partial p_X$. At $p_X = 0$ the second term vanishes and Assumption 2(c) gives $Q_B^{*AI} > 0$. □

Proposition 9 (Adoption frontier binds)**.** *Any profit-maximizing* $(\phi^\star, p_X^\star)$ *satisfies* $p_X^\star = p_{X,\max}(\phi^\star)$.

Proof. Consider the Lagrangian $\mathcal{L} = \Pi(Z,\phi,p_X) + \lambda\,[p_{X,\max}(\phi) - p_X]$. KKT conditions are

$$\frac{\partial \Pi}{\partial \phi} + \lambda\, p'_{X,\max}(\phi) = 0, \qquad \frac{\partial \Pi}{\partial p_X} - \lambda = 0, \qquad \lambda \ge 0, \qquad \lambda\,[p_{X,\max}(\phi) - p_X] = 0.$$

If the constraint were slack, $\lambda = 0$ and stationarity would require $\partial\Pi/\partial p_X = 0$. By Assumption 2(b)–(c) this cannot hold in a neighborhood of $p_X = 0$, and globally slackness would imply one fee could be raised to increase profit. Hence $\lambda > 0$ and $p_X^\star = p_{X,\max}(\phi^\star)$. □

Proposition 10 (Generic interiority of both fees)**.** *Under Assumption 2, there is a unique* $\phi^\star \in \Phi$ *solving*

$$\frac{d\widehat{\Pi}}{d\phi}(\phi^\star) = \frac{\partial \Pi}{\partial \phi} + \frac{\partial \Pi}{\partial p_X} \cdot p'_{X,\max}(\phi^\star) = 0,$$

with $p_X^\star = p_{X,\max}(\phi^\star) > 0$. *Thus both* $\phi^\star > 0$ *and* $p_X^\star > 0$ *generically.*

Proof. By Proposition 9 the problem reduces to $\max_{\phi\in\Phi} \widehat{\Pi}(\phi)$. Assumption 2(d) gives strict concavity and hence a unique maximizer. Since $p'_{X,\max}(\phi) < 0$ and both partial derivatives are finite and continuous, the FOC pins down $\phi^\star$ interior, with $p_X^\star > 0$ because $p_{X,\max}(\phi) > 0$ on Φ. □

Proposition 11 (Rent shifting under unilateral caps)**.** *Let* $(\phi^\star, p_X^\star)$ *be the unconstrained optimum.*

(i) If access is capped, $\phi \le \bar{\phi} < \phi^\star$, then the constrained optimum is

$$\phi^\dagger = \bar{\phi}, \qquad p_X^\dagger = p_{X,\max}(\bar{\phi}) > p_X^\star.$$

(ii) If usage is capped, $p_X \le \bar{p}_X < p_X^\star$, let $\phi_{\min}(p_X)$ be the inverse of $p_{X,\max}(\phi)$ (well-defined since $p'_{X,\max} < 0$). Then the constrained optimum satisfies

$$p_X^\dagger = \bar{p}_X, \qquad \phi^\dagger \ge \phi_{\min}(\bar{p}_X) > \phi^\star.$$

Proof. (i) On the frontier, $p_X = p_{X,\max}(\phi)$. With $p'_{X,\max} < 0$, if $\bar{\phi} < \phi^\star$ then $p_{X,\max}(\bar{\phi}) > p_{X,\max}(\phi^\star) = p_X^\star$. Hence the constrained solution is $(\bar{\phi}, p_{X,\max}(\bar{\phi}))$.
(ii) Feasibility requires $\bar{p}_X \le p_{X,\max}(\phi)$. Since $p_{X,\max}$ is strictly decreasing, the smallest feasible ϕ is $\phi_{\min}(\bar{p}_X) > \phi^\star$ when $\bar{p}_X < p_X^\star$. Thus the constrained solution is at least $(\phi_{\min}(\bar{p}_X), \bar{p}_X)$, with $\phi^\dagger > \phi^\star$. □

Corollary 2 (Comparative statics on the frontier)**.** *For full automation,*

$$p_{X,\max}^{full}(\phi) = m_{\text{dev}}\left(\frac{F}{F+\phi}\right)^{\frac{1}{\sigma-1}}, \qquad p'_{X,\max}(\phi) = -\frac{1}{\sigma-1}\frac{m_{\text{dev}}}{F+\phi}\left(\frac{F}{F+\phi}\right)^{\frac{1}{\sigma-1}} < 0.$$

Hence the inverse $\phi_{\min}(p_X)$ exists, is C^1, and satisfies $\phi'_{\min}(p_X) < 0$. The partial-automation case follows by subtracting $s_B w_S^{*AI}$ inside the cap.

Summary. The monopolist's optimum lies on the adoption frontier and, under mild conditions, both fees are strictly positive. If regulators cap one instrument, extraction shifts into the other; disciplining both fees is necessary to restore efficiency.