The Invisible Hand's Blind Spot: A Demand Externality in Competitive Automation

Rev. Steven Milanese · ORCID 0009-0008-0442-208X

Milanese, S. (2026). The Invisible Hand's Blind Spot: A Demand Externality in Competitive Automation. Strange Quarks, vol. I. https://doi.org/10.5281/zenodo.22240302

Version 1.0 · 19 August 2026 · https://doi.org/10.5281/zenodo.22240302

ABSTRACT

Hemenway Falk and Tsoukalas (2026) identify a demand externality in competitive automation: the automating firm captures the full cost saving of replacing a worker, while the consumer demand that worker would have provided is destroyed for every firm in the market. The distortion's sign depends on the income-replacement rate η — the share of displaced earnings workers recover — but that rate has been treated as fixed. We endogenize it: absorption capacity is itself depleted by automation and rebuilt by reinvestment, so η moves with the state of the transition, generating a tipping frontier — for each state of reinstatement capacity, the automation level at which replacement crosses unity. Since automation is sunk, excursions past the frontier exhibit hysteresis. The dynamically optimal instrument carries a premium for the reabsorption capacity each automated task depletes, and its sign turns positive strictly before replacement fails — early action is a corollary of the model, not a preference imposed on it. And because past displacement has produced persistent earnings losses, economies may already sit on its over-automation side, making the sign of the optimal Pigouvian instrument — a tax in one regime, a subsidy in the other — an empirically live question rather than a hypothetical. Today's income-replacement rate is measurable; the reabsorption mapping is not — aggregate data cannot separate movement along it from shifts of it, so the frontier's location and the instrument's future path remain unidentified without cross-sectional variation in automation intensity. Displacement-conditioned transfers shift the curve upward — policy relocates the frontier, not merely the distance to it.

Disclosure statement. *Working draft. The formal proofs and the dynamic environment's solution (Appendix B) are in preparation; theorem statements are final, proofs forthcoming. Appendix A derives the benchmarks; Appendix C presents the two-sector case in part.*

Working Thesis

> *A formal proof of structural vulnerability, contingent on reabsorption speed, that market*

> *forces cannot self-correct.*

1.1 The Intuition

Consider a firm that employs a worker at wage w to perform a task an AI system can perform at cost $c < w$. The firm's calculation is simple, and within the four corners of that calculation it is correct: automate the task, keep the saving $s = w - c$. The saving is real, immediate, and entirely private. Nothing in standard price theory suggests the firm should hesitate — indeed, a firm that hesitates leaves money on the table for rivals who will not.

But the worker is not only an input. The worker is also a customer. Wages paid to labor return to firms as revenue — just not, in general, to the firm that paid them. When a displaced worker recovers only a fraction η of their former income, the sector loses demand equal to $\ell = \lambda(1-\eta)w$ for every task automated, where λ is the share of worker income spent in the sector. One scoping choice deserves to be visible: the worker was a customer of everything, not only of this market, and spending directed outside the sector is a loss this model does not count. Appendix C develops the two-sector case — on the reabsorption side, not the spending side: the sectors that would reabsorb displaced workers are themselves automating, so displacement anywhere depletes reabsorption everywhere — and the single-sector figure is exact where automation is symmetric, understating depletion for any sector outpaced by the sectors its displaced workers would turn to. The loss that remains within the sector lands on every firm in the market. The automating firm, one competitor among N, absorbs only its share: ℓ/N.

Here is the asymmetry on which everything that follows turns. The firm internalizes one hundred percent of the cost saving and one-N-th of the demand destruction, though both are consequences of the same act. If there are twenty competitors in the market, the CEO who automates keeps the entire wage saving and feels one-twentieth of the damage it causes. The first figure appears in the firm's own ledger, quarterly and unmistakable. The second appears in everyone else's, diffuse and unattributable. A rational agent responds to the first and not the second — not because the agent cannot see the second — the model grants full visibility — but because seeing a cost is not bearing it.

The price system cannot repair this. Prices coordinate production with extraordinary efficiency: they tell the firm precisely what the worker costs and what the AI costs, and the firm's response to that information is socially useful. But no price tells the firm what the worker's spending was worth to other firms. The demand that displaced workers would have directed at rivals has no market, no signal, no counterparty, no venue in which it is priced. Creating that price is what Section 5 does. The invisible hand coordinates production. It does not coordinate automation — not because it fails at its task, but because this was never its task.

And note what this implies about competition, because it inverts the usual reasoning. In most settings, fragmentation disciplines: many rivals mean no single actor's private calculus can distort the whole. Here fragmentation is the disease. As N rises, each firm's share of the demand loss falls toward zero while the cost saving remains whole. A monopolist capturing the entire market's revenue would internalize the full ℓ: under monopoly, the externality disappears — the divergence between private and cooperative incentives, though not every divergence from what a planner would choose. The more competitive the market, the wider the gap between what the firm sees and what the economy pays. The institution economists trust most to align private incentives with social costs is, in this setting, the one that guarantees their divorce.

Whether this asymmetry is a footnote or a fault line depends on η — on what happens to the displaced worker next. Existing work treats that question as a parameter. This paper treats it as the thing the economy itself determines — and shows what follows.

1.2 The Source Result — and Its Limit

The starting point for everything we do is the model of Hemenway Falk and Tsoukalas (2026), which gives the asymmetry of Section 1.1 its first formal statement. In their framework, N symmetric competitors — twenty, in the running example of Section 1.1 — each choose what fraction of a continuum of tasks to automate, at a quadratic adjustment cost governed by k. The decentralized equilibrium automates until the private saving from the marginal task equals the firm's private share of the demand loss plus the marginal integration cost; the cooperative benchmark, in which firms internalize the loss in full, stops earlier. The distance between the two is the over-automation wedge, $\ell(1 - 1/N)/k$: proportional to the demand loss, diluted by competition, tempered only by adjustment friction. The core comparative statics flow from that expression, and they are three: the wedge widens as markets fragment, widens with the demand loss — and so narrows as displaced workers recover their income — and narrows as adjustment frictions rise. The baseline formula contains no s, which is itself a result: cost savings per se cannot soften the wedge. AI capability enters through two other doors. Cheaper AI lowers the threshold $N^* = \ell/s$, expanding the set of markets in the trap — as costs fall toward zero, virtually any market with two or more firms is inside it — and at the extreme tips both benchmarks into the full-automation corner, where the wedge saturates and stops registering an externality that is still consuming demand. More productive AI widens the wedge outright: in the source's productivity extension, capability adds a market-share term to the firm's first-order condition, raising the equilibrium rate while leaving the cooperative optimum untouched.

Their Corollary 2 gives the wedge its sign. For $\eta < 1$ the wedge is positive and competition over-automates; at $\eta = 1$ it closes; for $\eta > 1$ — displaced workers reabsorbed into better-paying roles — the wedge turns negative and the identical logic indicts competition for automating too slowly. The same instrument corrects both regimes, they observe: a tax below unity, a subsidy above. Two features of this result matter for what follows, and precision about both is load-bearing. First, the sign reversal is theirs — a labeled corollary, presented as an extension beyond the model's baseline domain of $\eta \in [0, 1]$. It is not a possibility we discovered, and not one they left informal. Second, η enters their analysis as a parameter: a number the economy is assigned, not an outcome it produces. Precision matters here, because their paper comes closer to our move than a summary suggests: their model holds η fixed, and their discussion conjectures that it rises over time as displaced workers retrain and new occupations emerge — but in neither does automation itself act on η. Their model can say what competition will do at any given income-replacement rate. It cannot say what income-replacement rate the transition will generate.

That is the limit. It is also the opening.

1.3 Our Move: η Is Endogenous to the Transition Itself

We make one change to the framework: the income-replacement rate becomes a function of the state of the transition itself. Write $\bar{\alpha}$ for the aggregate automation rate and R for reinstatement capacity — the stock of absorptive roles, transferable skills, and still-hiring industries that time and investment rebuild — and let $\eta = \eta(\bar{\alpha}, R)$, decreasing in the first argument, increasing in the second. The motivation is deliberately mechanism-neutral: absorption capacity is itself depleted by automation. Nothing in our results depends on the precise channel by which the depletion occurs.

Four consequences follow, and they structure the rest of the paper.

First, the economy has a tipping frontier. Define $\bar{\alpha}^*(R)$ by $\eta(\bar{\alpha}^*(R), R) = 1$ — for each state of reinstatement capacity, the automation level at which reabsorption exactly replaces lost income. Below the frontier, reabsorption more than replaces lost income, the wedge is negative, and the optimal instrument is a subsidy; above it, displacement destroys demand on net, the wedge is positive, and the optimal instrument is a tax (Theorem 1). Note what this does not assert: that any economy crosses the frontier mid-transition. A crossing requires $\eta(0, R_0) > 1$ — the earliest displaced workers landing in better jobs — and the displacement literature, which the source paper itself cites, documents exactly the opposite: large, persistent earnings losses (Jacobson et al., 1993). The empirically honest statement is stronger and stranger: the frontier exists, and the evidence suggests advanced economies may already sit on its over-automation side. Whether they do is an empirically live question, not a hypothetical one.

Second, the tipping point is not a point of no return — it is worse: a point of no full return. The second state variable is what makes this claim coherent: with $\eta = \eta(\bar{\alpha}, R)$, reinstatement capacity can rebuild even though automation is sunk. But an excursion past the frontier destroys demand along the way, and when R later rebuilds and η recovers, that demand is not refunded. The economy that retreats keeps the scar (Theorem 2). Welfare depends on the path, not merely the endpoint — which is the formal content of the claim that timing is the policy variable. A qualification, developed in Section 3.4: crossing the frontier needs no technological trend — on the subsidy side the automation target can leap it in a single step — but sustained deepening past it does; left to itself, the myopic economy self-arrests.

Third, what aggregate data cannot identify is the mapping. Today's η is measurable — the displacement literature measures displaced cohorts' earnings losses over multi-year horizons — from which the replacement rate is recovered — and with it the current sign of the wedge. But the reabsorption mapping $\eta(\bar{\alpha}, R)$ — its slope, the frontier's location, the counterfactual η under any alternative automation or policy path — is not identifiable from aggregate series, for the textbook reason: such series cannot separate movement along the mapping from shifts of it (Proposition, Section 3.4). Everything forward-looking policy needs — tipping risk, timing, preemption, the instrument's path — is a property of the mapping, not the level. Identification requires cross-sectional variation in automation intensity across sectors and regions, under a common-curve assumption. This is not a counsel of despair; it is the design specification for the empirical program of Section 6. And the political-economy consequence sharpens rather than weakens: lobbies do not contest today's measured η; they contest the counterfactual — "our sector's reabsorption will hold" — which is precisely the object aggregate data cannot settle.

Fourth, policy can move the tipping point itself. Displacement-conditioned transfers — wage insurance, severance, targeted reemployment support — shift the $\eta(\bar{\alpha}, R)$ mapping upward, relocating the frontier rather than merely shrinking the economy's distance to it. The state space also formalizes a distinction the source draws informally in its revenue-allocation discussion: direct transfers raise η holding R fixed — mechanical, exposed to moral hazard, and sustainable only while funded — while retraining invests in R and moves the whole frontier. A static model cannot express any of this; an endogenous one cannot avoid it.

A corollary is worth stating here, because it converts the source's own conjecture into a regime of this model. Their discussion predicts that the income-replacement rate will rise as workers retrain, with the optimal tax shrinking accordingly. In our state space, that conjecture is the regime in which reinstatement outruns depletion: R-rebuilding is their rising-η force, formalized, and $\bar{\alpha}$-depletion is the countervailing force they do not model. The race between the two decides the path of the optimal instrument — non-monotone in general, self-limiting when reinstatement wins (Corollary, Section 3.4). Their Section 6 is not an objection to this framework; it is one of its regimes.

One clarification before proceeding. The source paper's own endogenous-wage extension lets aggregate automation depress wages; in their words, wage flexibility "raises the threshold at which the externality activates" and, short of collapsing wages to AI's cost, cannot close the wedge once it does. Their channel moves that threshold; ours operates on the sign. The two compose; they do not compete.

1.4 Contributions

Our contributions are four.

First, we endogenize the reabsorption margin in the demand-externality framework of Hemenway Falk and Tsoukalas (2026), converting their static sign corollary into dynamics: a tipping frontier in the automation path at which the optimal Pigouvian instrument reverses sign, hysteresis around it, and a depletion-corrected rate whose sign changes strictly before the frontier — the myopic rate flips at the frontier; the dynamic rate flips before it (Section 3.4; proofs in Appendix B).

Second, we prove that the reabsorption mapping is unidentified from aggregate data: the level of η is measurable, but aggregate series cannot separate movement along the mapping from shifts of it, so the frontier's location, the mapping's slope, and every counterfactual path remain unidentified however much aggregate data is gathered. This formalizes a political-economy problem that has been treated as rhetorical: what lobbies contest is the counterfactual — that their sector's reabsorption will hold — and the counterfactual cannot be settled by commissioning more of the same data (Sections 3.4 and 5.2).

Third, we re-examine each of the source's policy verdicts under endogenous reabsorption, and show which survive unchanged and which transform. Universal transfers still enter the demand level and never touch the wedge; capital taxation still cancels at the automation margin; bargaining still fails on both margins. But displacement-conditioned transfers now relocate the tipping frontier rather than merely shrinking the per-task loss, and retraining now builds reinstatement capacity — an investment in the state variable that moves the entire frontier. The same decomposition that organizes these mechanisms also disciplines our interpretive claim: the externality term is structural — present even when the planner assigns workers zero weight — while the distributional premium is not, and we confine "structural, not distributional" to the first (Section 4; decomposition in Section 3.2).

Fourth, we retarget the empirical program: the object to estimate is η as a function of automation intensity, identified from cross-sectional variation, with consumer-side, cross-sector spending signatures that separate the demand externality from ordinary competition — and with the current ambiguous evidence read, correctly, as what a latent failure mode looks like before the tipping point (Section 6).

1.5 Roadmap

Section 2 develops the mechanism informally and disarms the objection every economist reaches for first. Section 3 presents the model and its results. Section 4 walks the policy failures, mechanism by mechanism. Section 5 takes up the instrument and its political economy. Section 6 specifies the empirical program and confronts the skeptical evidence. Section 7 situates the paper, with particular care for the nearest neighbor. Section 8 concludes. Proofs are in the appendices; the body text carries the mechanisms. Throughout, the worked example is the twenty-firm market of Section 1.1 — the CEO who keeps the whole saving and feels one-twentieth of the damage. Everything this paper proves is, in the end, about that one-twentieth.

2. The Mechanism

This section develops the mechanism in words. Section 3 gives it a model; Section 4 gives it policy teeth. Everything here runs on the twenty-firm market of Section 1.1.

2.1 The Firm's Calculus

Firm i automates a fraction α_i of its tasks, at a quadratic adjustment cost governed by k. Each automated task saves $s = w - c$, and the saving is captured in full. Each automated task also displaces a worker, and displacement destroys sectoral demand of $\ell = \lambda(1-\eta)w$ per task — but that loss is borne by the market, not the firm. With twenty competitors, nineteen-twentieths of it lands in someone else's ledger.

Two accounting exercises now diverge. The private one weighs the whole saving against one-twentieth of the damage. The social one weighs the whole saving against the whole damage. The externality is the difference between them, and it has a property that runs against every instinct competition policy teaches: it grows with competition. As N rises, each firm's share of the damage falls toward zero while its share of the saving remains one; the fragmented market is not the discipline but the worst case. Under monopoly, the externality disappears — the divergence between private and cooperative incentives, that is. A planner who places weight on worker income would still choose differently from the monopolist; what vanishes at N = 1 is the wedge between one firm and its rivals' joint interest, not the wedge between the firm and society.

2.2 The Fordist Fallacy, Preempted

Every economist meeting this argument reaches for the same rebuttal, and it is a good one: Henry Ford did not pay five dollars a day so his workers could buy his cars. A firm that raises wages to manufacture demand for its own product is making a gift to its competitors, who sell into that demand without funding it. The firm-level claim is a fallacy, and we affirm it without reservation — because the fallacy is the mechanism. Precisely because no individual firm should internalize the demand its payroll creates, no individual firm does; the demand effect is external to every firm and therefore priced by no one. What is irrational for one firm is indispensable for all of them together: their payrolls are each other's revenues. The Fordist fallacy holds at the micro level and reverses at the macro level, and the boundary between the two is exactly where the externality lives. The reader who arrived with the objection now has the paper's thesis: the demand side of the labor market is a commons, and the rationality of each grazier is not the rationality of the group.

2.3 The Prisoner's Dilemma Structure

Suppose the twenty firms could agree to restrain automation at the cooperative level. The agreement would raise every firm's profit — restraint is Pareto-superior to the race. It would also be worthless. If nineteen firms restrain, the twentieth gains a cost advantage and market share by automating. If nineteen automate, the twentieth must follow or be crushed on price. Automation is a strictly dominant strategy: each firm's best reply does not depend on what the others do.

That is what makes the trap a trap rather than a mistake. Foresight does not rescue the firms — the model grants them full visibility of the demand mapping, and they over-automate anyway, because seeing a cost is not bearing it. Communication does not rescue them either: under dominance, announcements of restraint are cheap talk in the sense of Crawford and Sobel (1982) — the observation is the source paper's; the citation is ours. Every CEO in the market can see the cliff. The cliff is visible, quantified, and shared — and the gas pedal is still the rational choice, because the alternative is to arrive at the same cliff slightly poorer than one's rivals.

2.4 The Red Queen Effect

The two forces economists trust most to correct markets both feed the trap. Competition does so through Section 2.1's dilution. Technology does so through a channel that inverts the usual consolation — the idea that technological progress disciplines its own disruptions.

When AI becomes more productive, each firm perceives a market-share gain from automating beyond its rivals — a gain that is real if the others stand still. The others do not stand still. At the symmetric equilibrium, every firm has automated more, the market-share gains cancel to zero, and what remains is the additional demand destruction. The firms run faster to stay in the same place, and the running itself burns the commons. The source paper names this the

Red Queen effect and proves it: in their productivity extension, greater capability raises the equilibrium automation rate while leaving the cooperative optimum untouched, widening the wedge outright. Cheaper AI enters differently, widening not the wedge but the trap's jurisdiction: it lowers the market fragmentation at which automation begins, and at the extreme tips both benchmarks into the full-automation corner, where the wedge saturates and stops registering an externality that is still consuming demand. Better technology, more competition, faster decline — the triangle is complete.

2.5 Pedigree

One paragraph of lineage, with full positioning deferred to Section 7. Demand spillovers of this kind are among the oldest mechanisms in economics: Rosenstein-Rodan's big push (1943), formalized by Murphy, Shleifer, and Vishny (1989), set within the coordination-failure framework of Cooper and John (1988). That lineage is the source paper's own related-work framing, and we adopt it as they built it. What this paper adds to the lineage — history-dependence, tipping, and a threshold that policy itself can move — is Section 7's subject.

Section 3 makes all of this precise, and then sets it in motion.

3. The Model

This section does two things. Sections 3.1–3.3 restate the source model in its own notation, as compactly as self-containment allows — enough to fix the three objects everything turns on: the saving s, the loss ℓ*, and the wedge between private and cooperative automation. Section 3.4 then makes this paper's single change, and proves what follows from it.*

3.1 The Environment

The economy contains $N \geq 2$ symmetric firms — twenty in the running example — each endowed with a continuum of tasks of measure L. Firm i automates a fraction $\alpha_i \in [0, 1]$ of its tasks. Automation costs c per task against a wage of w, so each automated task saves $s = w - c$, and re-tooling is costly: the firm pays a quadratic integration cost $(k/2)L\alpha_i^2$, with $k > 0$ governing the friction.

The demand side is where the mechanism lives. Workers spend a share λ of their income in the sector. A displaced worker recovers a fraction η of their former wage through reemployment, transfers, or other sources; the remainder is lost. Each automated task therefore destroys $\ell = \lambda(1-\eta)w$ of sectoral demand. With symmetric automation rate $\bar{\alpha}$, aggregate sectoral demand is

$$D(\bar{\alpha}) = A + \lambda wLN - \ell LN\bar{\alpha},$$

where A is autonomous demand. Read the last term slowly: sectoral demand is a declining function of aggregate automation, and each firm's revenue is D/N. The worker-as-customer of Section 1.1 is now an equation.

Two baseline assumptions deserve explicit statement, because later sections lean on both. First, firm owners spend none of their income in the sector — the owner MPC is normalized to zero. This is the load-bearing demand assumption: it is what makes the saving s a destruction of sectoral demand rather than a transfer to agents who would spend it back. The source relaxes it through capital-income recycling (their Section 5.4) and gives it a structural foundation: under non-homothetic preferences, mass-market goods saturate at high incomes, so marginal owner income flows to a separate luxury segment (their appendix; Matsuyama 2002; Comin et al. 2021; Boppart 2014). Second, the environment assumes full transparency: every firm observes how automation maps into lost worker income and reduced aggregate spending. This is the assumption Section 2.3's "seeing is not bearing" leans on — and the source's own question is whether visibility alone curbs automation. It does not.

3.2 Three Benchmarks, One Wedge

Firm i chooses α_i to maximize profit, taking rivals' automation as given. The marginal calculus is

$s = \ell/N + k\alpha_i$:

the saving on the marginal task, weighed against the firm's own share of the demand loss, plus the marginal integration cost. The symmetric Nash equilibrium and the cooperative benchmark follow directly:

$\alpha^{NE} = (s - \ell/N)/k,\ \alpha^{CO} = (s - \ell)/k,$

and the over-automation wedge is their difference,

$\alpha^{NE} - \alpha^{CO} = \ell(1 - 1/N)/k.$

Section 1.2 has already walked through the comparative statics: the wedge widens with fragmentation, widens with the demand loss, narrows with friction — and contains no s. In the running example the externality term is nineteen-twentieths of ℓ/k: each firm internalizes one part in twenty of what its automation destroys, while the market collectively bears all of it. (Derivations: Appendix A.)

The third benchmark is a planner who places weight μ on worker income. The source's decomposition (their eq. 9) splits the distance to that benchmark into two terms:

$\alpha^{NE} - \alpha^{SP}(\mu) = \ell(1 - 1/N)/k + \mu\ell/(\lambda(1-\mu)k).$

Two terms, two diagnoses. The first is the externality term: present even at $\mu = 0$ — a planner who cares nothing for workers still disagrees with the market — and growing with N toward its maximum of ℓ/k. The second is a distributional premium: independent of N, and exploding without bound as $\mu \to 1$. This paper's claim that the problem is "structural, not distributional" attaches to the first term only, and Sections 4.6 and 7.4 hold it to that scope.

3.3 The Sign of the Wedge

One result of the source paper is load-bearing for everything we do, and we state it exactly as they prove it. Their Corollary 2 (Sign of the externality): in the interior regime, the wedge is positive for all $\eta < 1$, zero at $\eta = 1$ — where the decentralized and cooperative rates coincide — and negative for $\eta > 1$, when reabsorption into better-paying roles makes displacement a net creator of demand and competition automates too slowly. Their baseline defines $\eta \in [0, 1]$; the $\eta > 1$ case is presented as a natural extension beyond it, and their Section 4.1 observes — pointing forward to their 4.6 — that the same instrument corrects both regimes: a tax below unity, a subsidy above. One boundary observation matters for what follows: replacement income above full replacement arrives from outside the modeled flows — the $\eta > 1$ regime extends rather than closes their accounting, and the frontier that organizes this paper lives at its edge.

Carry two features forward. The sign reversal is theirs. And η is exogenous throughout their analysis — a parameter the economy is assigned. Their discussion conjectures that it rises over time as workers retrain; nowhere in their framework does automation itself act on it.

3.4 Endogenous Reabsorption

We replace the parameter with a state. The transition is described by two state variables. The first is aggregate automation $\bar{\alpha}$, which is sunk: investments in AI capacity are substantially irreversible, so $\bar{\alpha}$ is non-decreasing. The second is reinstatement capacity R — the stock of absorptive roles, transferable skills, and still-hiring industries — which time and investment rebuild. The income-replacement rate is

$\eta = \eta(\bar{\alpha}, R)$, *decreasing in* $\bar{\alpha}$, *increasing in* R. Throughout, depletion names this channel and no other: sunk automation crowds given capacity — η falls at any R — while the reinstatement stock itself is untouched by automation, by construction of its law of motion below.

R is the reinstatement channel of Acemoglu and Restrepo (2019) given an explicit state; the source's reduced-form treatment of wage adjustment (their Section 5.3) is the methodological precedent for constructing it this way.

To make statements about paths, the state space needs laws of motion; we specify them here and solve the environment in Appendix B. Time is discrete. Each period, firms replay the stage game of Section 3.2 at the current state: the symmetric equilibrium automation choice at $\eta_t = \eta(\bar{\alpha}_t, R_t)$ determines the flow of new automation. The states evolve by

$$\bar{\alpha}_{t+1} = \max\{\bar{\alpha}_t, \alpha^{NE}(\eta_t)\},\ R_{t+1} = (1 - \delta)R_t + \rho + I_t,$$

so automation is sunk and non-decreasing, while reinstatement capacity depreciates at rate δ, rebuilds at a baseline rate ρ, and responds to policy investment I_t. Per-period social payoff is the stage-game surplus net of the cost of reinvestment, $S(\bar{\alpha}_t, R_t) - \chi(I_t)$, with χ increasing and convex — without it the planner would purchase reabsorption without limit, and the trade-off this paper exists to study would be assumed away. Social welfare is the discounted sum $W = \Sigma_t\, \beta^t\, [S(\bar{\alpha}_t, R_t) - \chi(I_t)]$. Firms, by contrast, are taken to replay the static stage game each period rather than to manage their influence on the state. The assumption is justified by the model's own central fact: each firm's influence on the aggregate state is of order 1/N — the same dilution that generates the externality also guarantees that no firm's dynamic calculation can overturn its static one. The trap is self-reinforcing: the more fragmented the market, the more rationally myopic each competitor. Everything labeled a path result below — Theorems 2 and 3, and the Corollaries — is a statement about this environment.

Definition (Tipping frontier). $\bar{\alpha}^*(R)$ is the automation level solving $\eta(\bar{\alpha}, R) = 1$: for each state of reinstatement capacity, the point at which reabsorption exactly replaces lost income.

Theorem 1 (Tipping frontier). In the interior regime: below the frontier, $\eta > 1$ and the decentralized economy under-automates — the optimal static instrument is a subsidy; above the frontier, $\eta < 1$ and the economy over-automates — the static Pigouvian rate is a tax equal to the uninternalized demand loss per task,

$$\tau^* = (N-1)/N \cdot \lambda(1 - \eta(\bar{\alpha}, R))\, w,$$

the myopic benchmark that Theorem 3 corrects for dynamics. The frontier is strictly increasing in reinstatement capacity: by the implicit function theorem, $\bar{\alpha}^{*\prime}(R) = -\eta_R/\eta_{\bar{\alpha}} > 0$, so additional reabsorption capacity pushes the frontier outward — the formal content of the claim that policy relocates the frontier. (Proof: Appendix B.)

Remark. The frontier is endogenous to policy: displacement-conditioned transfers shift $\eta(\bar{\alpha}, R)$ upward and move $\bar{\alpha}^*(R)$ itself (Section 4.1).

What Theorem 1 does not assert is that any economy crosses the frontier in the course of the transition. A crossing from the subsidy side requires $\eta(0, R_0) > 1$ — the earliest displaced workers landing in better jobs — and the displacement literature, which the source itself cites, documents the opposite: large, persistent earnings losses (Jacobson et al., 1993). The empirically honest reading is that the frontier exists and that advanced economies may already sit on its over-automation side. Which side they sit on is an empirically live question, not a hypothetical one.

Theorem 2 (Hysteresis). Fix the policy sequence $\{I_t\}$, and consider two transition paths that share it and terminate at the same state $(\bar{\alpha}_T, R_T)$ — differing through a one-time impulse to the automation stock, both sharing the exogenous trend path (c(t), φ(t)): over an excursion window the first path's automation exceeds the state-contingent cooperative benchmark — ᾱ_t > α^CO(η_t), by more than the second path falls short of it — while the second never exceeds it. Because the law of motion of R does not involve ᾱ, the two paths carry identical reinvestment stocks and identical investment costs; the welfare comparison is therefore clean. Welfare along the first path is strictly lower, by exactly the surplus forgone over the course of the excursion; no recovery of R reimburses it. (Proof: Appendix B.) The second state variable is what makes this theorem statable at all: with sunk ᾱ alone, η could never recover, and "recovery" would have no meaning. With R, recovery is possible — and insufficient. The economy that retreats keeps the scar. The hysteresis is remanence, not metaphor: a trend that reverses leaves ᾱ at its historical maximum, so the long-run state is selected by the path, not merely the welfare along it — sunk-cost hysteresis in the sense of Dixit (1989) and Baldwin and Krugman (1989).

Theorem 3 (The depletion-corrected rate). In the dynamic environment, automation imposes a second externality the static rate does not price: each firm's automation today permanently lowers the replacement rate available from the common reabsorption pool, raising ℓ for every firm in every future period. The dynamically optimal instrument is

$$\tau^{**}(t) = (N-1)/N \cdot \lambda(1 - \eta(\bar{\alpha}(t), R(t)))\, w + \Lambda(t),$$

where Λ(t) ≥ 0 is the shadow price of reabsorption depletion — strictly positive wherever automation depletes, in either regime — augmented near the frontier by the precautionary margin that Theorem 2's irreversibility generates. On the over-automation side, the myopic rate of Theorem 1 systematically undertaxes. (Proof: Appendix B.) One consequence qualifies the source's robustness argument: that "even an imprecisely targeted tax yields a first-order gain" presumes a static externality; under endogenous reabsorption, the myopic rate is a floor, not the target.

Corollary (The dynamic boundary precedes the frontier). The dynamically optimal instrument changes sign where $\tau^{**}(t) = 0$, which is at $\eta = 1 + N\Lambda(t)/[(N-1)\lambda w] > 1$ — strictly on the subsidy side of the static frontier. The economy should be taxed into positive rates while measured replacement still exceeds unity. Early action is not a precautionary preference layered on the analysis; it is the analysis.

Remark (Self-arrest, and what escapes it). The myopic dynamics contain a stabilizing feedback: as $\bar{\alpha}$ rises, η falls, ℓ rises, and the equilibrium automation target $\alpha^{NE}(\eta)$ falls with it, until the ratchet binds where $\bar{\alpha} = \alpha^{NE}(\eta(\bar{\alpha}, R))$. Left to itself, the economy self-arrests at that fixed point. Sustained excursions deep past the frontier therefore require a moving target, which only an exogenous driver supplies — AI costs falling or capability rising through the transition. Appendix B characterizes the fixed point and the trend-driven paths; Section 6.4 draws the empirical moral, which cuts against complacency in both directions. Crossing needs no trend at all — on the subsidy side $\ell < 0$, so the automation target runs ahead of s/k and can leap the frontier in a single step, the ratchet arresting the economy beyond it. What the trend governs is ongoing deepening: only a moving target sustains it, and with the trend in place the warning signs steepen rather than jump.

Corollary (The optimal instrument's path is non-monotone in general). Along a transition path,

$$\tau^{**}(t) = (N-1)/N \cdot \lambda(1 - \eta(\bar{\alpha}(t), R(t)))\, w + \Lambda(t),$$

whose first term declines over time if and only if reinstatement outruns depletion — $d\eta/dt > 0$ along the path. The source's conjecture of a shrinking, self-limiting tax is exactly this R-wins regime. In general the race can go the other way, and the optimal tax can rise for long stretches before it falls. Their Section 6 is not an objection to this framework; it is one of its regimes.

Proposition (Aggregate non-identification of the reabsorption mapping). The level of η is observable, with a lag: displacement-cohort studies measure earnings losses over multi-year horizons (Jacobson et al., 1993), and the source itself notes that setting the rate requires only sector-level observables. The mapping is another matter. Let $\{\bar{\alpha}_t, \eta_t\}$ be any observed aggregate path. Then there exist reabsorption mappings $\eta(\bar{\alpha}, R)$, together with shift processes over them, that generate the same path while implying different slopes, different frontier locations, and different counterfactual paths of η under alternative automation or policy trajectories. Aggregate data cannot separate movement along the mapping from shifts of it: policy, demographics, and the cycle displace the mapping even as automation moves the economy along it. The sign of today's wedge is measured; everything forward-looking — tipping risk, timing, preemption, the instrument's path — depends on the mapping, and the mapping is unidentified without cross-sectional variation in automation intensity under a common-curve assumption. (Proof — observational equivalence over aggregate paths: Appendix B.) The political-economy consequence is exact: lobbies do not contest today's measured η; they contest the counterfactual — "our sector's reabsorption will hold" — and that is precisely the object aggregate data cannot settle.

One clarification to close. The source's endogenous-wage extension lets aggregate automation depress wages, which in their words "raises the threshold at which the externality activates" without closing the wedge once active. That channel moves their N^*. Ours moves the sign. The two compose without tension: wage adjustment can slow the economy's approach to automation even as endogenous reabsorption decides what automation means when it arrives.

4. Why Standard Remedies Fail: Mechanisms, Mapped to Parameters

The source paper evaluates the standard policy menu mechanism by mechanism, and their Table 1 states the organizing fact: each instrument enters the model through exactly one parameter — retraining through η, UBI through A, the capital income tax through t, worker equity through ε, bargaining through M — and only the automation tax, entering through τ, eliminates the wedge. Their ranking is stated with a care worth matching: it holds in the regime where the externality is active and uninternalized — spending leaks ($\lambda < 1$), replacement is incomplete ($\eta < 1$), automation rates are non-contractible and unobservable across firms, and the market is fragmented — and the apparent exceptions (full reabsorption, full recycling, full profit-sharing at $\varepsilon = 1/\lambda$, a binding grand coalition over observable automation) are boundary or knife-edge cases, each of which removes the very leakage or non-contractibility that constitutes the externality, confirming rather than overturning the ranking. The rest cushion the losers or partially shrink the distortion; none removes it. We re-examine each instrument under endogenous reabsorption. This changes the question. The static question is whether an instrument moves the wedge at a fixed income-replacement rate. The dynamic question is whether it moves the state. Under the second question the menu sorts into three classes: instruments that still fail on their own terms, instruments that still half-work, and two instruments that now do something the static model could not express — they move the frontier.

4.1 Universal Basic Income

A UBI funded from general revenue is unconditional: the employed and the displaced receive the same payment. It therefore enters the model as an increase in autonomous demand A — a constant added to aggregate spending that does not alter the marginal income loss from displacement. And the wedge does not contain A. The expression $\ell(1 - 1/N)/k$ is built from the demand loss, the market's fragmentation, and the adjustment friction; the level of baseline spending appears nowhere in it. UBI fails, then, not because it is small or badly funded but because of its universality: it raises the consumption floor and leaves the automation margin exactly where it was. A profit tax paired with an unconditional transfer fails for the same reason twice — both instruments move levels, and neither touches the margin where the externality resides.

The contrast that carries the point is the transfer UBI is usually confused with. Displacement-conditioned transfers — wage insurance, severance supplements, reemployment bonuses — are paid only to the displaced. They enter η directly, shrink $\ell = \lambda(1-\eta)w$, and narrow the wedge at its source. Universality is not a detail of program design; it is the difference between a transfer that the firm's calculus notices and one it never sees.

Endogenous reabsorption strengthens the contrast into something the static model cannot say. With $\eta = \eta(\bar{\alpha}, R)$, a displacement-conditioned transfer does more than shrink the wedge at a fixed frontier: it shifts the mapping upward and thereby relocates $\bar{\alpha}^*(R)$ itself. Policy moves the frontier, not merely the economy's distance to it. Two caveats keep the claim honest. First, this is a statement about UBI as modeled here — unconditional and general-revenue-financed — not a verdict on every design that travels under the name; a UBI variant conditioned on displacement is, in this taxonomy, a displacement-conditioned transfer. Second, UBI retains a real role: it matters for welfare, and it addresses the distributional premium of Section 3.2's decomposition. What it does not address is the externality term, and the externality term is the part that survives a planner who cares nothing for distribution at all.

4.2 Retraining and Upskilling

Retraining acts on the correct margin — it is the one entry on the standard menu whose parameter is η itself — and under endogenous reabsorption it acquires a deeper distinction: it is the only standard remedy that invests in the state variable. The source's own revenue-allocation discussion draws the informal version of this line: direct transfers raise η mechanically, hold R fixed, and are sustainable only while funded, with the moral-hazard exposure that conditioning on displacement invites; retraining builds the stock of absorptive roles and transferable skills — R — and moves the entire frontier, making the gain self-sustaining. In the state space of Section 3.4, the distinction is exact: transfers are a flow into $\eta(\bar{\alpha}, R)$ at given R; retraining is an investment I_t in the law of motion of R.

But retraining fights a moving target, and the target now moves for two reasons. The first is the source's: partial reabsorption leaves $\eta < 1$, and any permanent wage-loss residual keeps the wedge alive — retraining narrows but, short of full replacement at equal-or-better wages, does not close. The second is ours, and we state it conditionally, because we do not assume the curvature that would guarantee it: if depletion bites where retraining works — if the marginal training dollar buys less reabsorption as automation spreads — then the instrument's yield falls as the problem grows. What stands without any curvature assumption is the asymmetry of jobs: retraining narrows the wedge but never prices it. An instrument whose yield may fall as the problem grows, and which in any case leaves the externality unpriced, is a complement to the corrective tax, not a substitute for it. The self-limiting trajectory the source conjectures — tax revenue funding retraining, retraining raising η, the tax shrinking — is real, and Section 3.4 names it: it is the R-wins regime of the Corollary. It is a regime, not a guarantee.

4.3 Worker Equity and Profit-Sharing

Worker equity is the cleverest failure on the menu, because it reaches inside the firm's objective function. Give workers a profit share at rate ε and part of the demand destroyed by displacement returns to them as capital income, which they spend: the demand-loss parameter becomes $\ell_\varepsilon = \ell - \lambda\varepsilon s$, and the leakage divisor becomes $N_\varepsilon = N - \lambda\varepsilon(N-1)$. The Nash equilibrium moves toward the cooperative optimum; the cooperative optimum itself does not move, because a planner who controls every firm already internalizes the recycling — in the planner's first-order condition the profit-sharing terms cancel. The source's Proposition 3 gives the wedge under sharing: $\ell(N-1)(1-\lambda\varepsilon)/[kN_\varepsilon]$, strictly decreasing in ε and strictly positive for all $\varepsilon < 1/\lambda$.

That bound is the failure. Closing the wedge requires $\lambda\varepsilon = 1$, i.e. $\varepsilon = 1/\lambda$, which exceeds one whenever $\lambda < 1$: each dollar of profit recycled to workers generates only λ dollars of sectoral demand, so compensating for the leakage would require sharing more than the firm's entire profit. Even full profit-sharing leaves the wedge strictly positive. And the arrangement does not arise on its own: their Corollary 3 shows that a firm choosing its own sharing rate to maximize retained profit sets $\varepsilon_i = 0$ as a dominant strategy, because the firm captures only λ/N of the demand its sharing creates. The recycling runs within the firm; the externality runs across firms. No within-firm instrument reaches a cross-firm margin.

4.4 Capital Income Taxation

A proportional tax on capital income scales the entire profit function by $(1 - t)$, and a positive scalar cancels from the first-order condition: the equilibrium automation rate, the threshold N^*, and the wedge are all unchanged. The instrument operates on profit levels, not on the per-task margin where the externality lives — a failure structurally identical to UBI's. The distinction matters because the policy debate conflates the two instruments: the robot taxes studied in the literature are per-unit levies on adoption, which do operate on the margin; a capital income tax is a fundamentally different instrument wearing a similar name.

Capital-income *recycling* — the source's Section 5.4 relaxation, in which owners spend a share $\hat{\eta}$ of their income in the sector — does better, and its better is still not enough. Recycling narrows the wedge: it raises the fraction of the demand loss each firm internalizes from $1/N$ toward $1/\hat{N}$, with $\hat{N} = N(1-\hat{\eta}) + \hat{\eta}$ interpolating between the market and the monopolist.

But their Proposition 10 gives the boundary: the externality vanishes only when $\hat{\eta} \geq \ell/s$, and when $\ell > s$ — the regime in which firms automate where a planner would prefer none — the required recycling rate exceeds one. Recycling is impotent precisely where the externality is most harmful. Addressing how income is spent narrows the wedge; the dilution across firms persists, and the wedge persists with it.

4.5 Bargaining: Both Margins

If the externality cannot be taxed or transferred away, can it be bargained away? The question has two margins, and both fail — for different reasons.

Firm-to-worker bargaining cannot reach the externality at all. The uninternalized loss $\ell(1 - 1/N)$ falls on the revenue of *rival* firms, and rivals' retained workers are unharmed by firm i's automation — they earn w per task and have no basis for negotiation with i. The channel runs through the product market, not through any bilateral employment relationship, so no within-firm bargain can be written over it.

Firm-to-firm bargaining reaches the right margin and dies of free-riding. A coalition of M firms that jointly maximizes members' combined profit internalizes M/N of the aggregate demand loss and automates at $\alpha^M = (s - M\ell/N)/k$; the residual wedge $\ell(1 - M/N)/k$ stays strictly positive for all $M < N$ (their Proposition 4). Only the grand coalition replicates the cooperative optimum, and four features of this externality prevent the grand coalition from forming. Voluntary agreements are not self-enforcing: automation is strictly dominant, so a member gains from deviating whether or not others honor the agreement — no non-binding arrangement is stable. The externality is multilateral and diffuse: each firm's contribution to the aggregate loss is too small to motivate any single negotiation and too large in aggregate to ignore — precisely the large-numbers setting in which Coase himself acknowledged that private bargaining breaks down. The automation rate is not contractible among firms: an internal organizational choice rivals cannot observe or verify. And automation is sunk: in a repeated setting, trigger-strategy punishments cannot undo a deviation, and a firm that delays while rivals proceed loses share. The fundamental obstacle is incentive compatibility, not transaction costs — even with costless negotiation, the game retains its dominant-strategy structure. Communication does not rescue restraint either: announcements of restraint are cheap talk in the sense of Crawford and Sobel (1982), an observation we owe to the source paper, whose current version leaves it uncited. Correcting the externality requires an instrument that does not rely on voluntary agreement — one that changes each firm's marginal incentive directly.

4.6 The Meta-Failure

Stand back from the menu and a pattern emerges that no single row of the source's Table 1 displays. Every standard instrument does one of three things: it shifts a level (A, t), narrows the per-task loss (η, ε), or requires a coalition (M). None prices the per-task externality except τ. The failure is not that the standard remedies are badly designed; it is that they are answers to a different question. They ask how to cushion displacement, recycle its proceeds, or restrain it by agreement. The externality asks what price a firm faces when it destroys demand it does not own.

One scoping claim, held to its exact width. This paper's contention that the problem is "structural, not distributional" characterizes the externality term $\ell(1 - 1/N)/k$ — the term that survives a planner who weights workers at zero, complete credit markets, and full transparency. The total wedge against a planner who does weight workers additionally contains the distributional premium $\mu\ell/(\lambda(1-\mu)k)$, and transfers genuinely address that component; Section 3.2 assigns each term its diagnosis, and nothing here walks the assignment back.

Endogenous reabsorption sharpens the meta-failure rather than softening it. In the static menu, transfers and retraining are the half-measures: they shrink the wedge but never eliminate it, and only τ fixes it. In the dynamic environment they acquire a second job the static model could not assign them: transfers shift the η mapping, retraining builds R, and both move the frontier that determines which regime the economy is in. The corrective tax and the income-replacement policies are therefore not rival answers to one question. They are the two halves of the dynamic answer — τ prices the externality along the path; transfers and retraining move the path. Section 5 takes up the first; the second returns there as the instrument's financing.

5. The Pigouvian Instrument and Its Political Economy

5.1 The Formula — and What It Prices

The source's Proposition 5 establishes the corrective instrument: a per-task automation tax at rate $\tau^* = \ell(1 - 1/N)$ implements the cooperative optimum, and with an exogenous lump-sum rebate it delivers each firm exactly its cooperative profit. The rate has the interpretation the whole paper is built on: each firm already bears ℓ/N of the demand loss its automation creates, and the tax charges it for the remaining $\ell(1 - 1/N)$ — the portion imposed on rivals. The tax is not a levy on technology. It is the price the price system never generated: the market value, to other firms, of the spending the displaced worker no longer directs at them. Written out, $\tau^* = (N-1)/N \cdot \lambda(1 - \eta)w$, and for large N it approaches $\lambda(1-\eta)w$ — which is why the source can note that setting the rate requires only sector-level observables. The demand loss is built from quantities any statistical agency already tracks: wages, spending shares, replacement rates.

Endogenous reabsorption turns the formula into an instrument with a time path. The static rate becomes the myopic benchmark of Theorem 1, $\tau^*(\bar{\alpha}, R) = (N-1)/N \cdot \lambda(1 - \eta(\bar{\alpha}, R))w$ — state-contingent, and negative on the subsidy side of the frontier. What the source's Section 4.1 presents as a symmetry — the same instrument, a tax below full replacement and a subsidy above — becomes, under $\eta = \eta(\bar{\alpha}, R)$, a path rather than a switch: the economy carries a position relative to the frontier, and the optimal instrument tracks it. Theorem 3 then corrects the benchmark for dynamics: the dynamically optimal rate is $\tau^{**}(t) = \tau^*(t) + \Lambda(t)$, where $\Lambda(t) \geq 0$ prices the depletion externality — the future ℓ each firm's automation raises for every firm — and carries a precautionary margin near the frontier, where Theorem 2's irreversibility makes excursions un-refundable. On the over-automation side, the myopic rate is a floor, not the target. This qualifies the source's robustness argument at its foundation: the observation that "even an imprecisely targeted tax yields a first-order gain" rests on a quadratic welfare loss around a static optimum; under endogenous reabsorption there is a second margin the imprecise tax does not price, and systematic undertaxation compounds through the state.

Two features of the instrument survive endogenization intact. The first is financing: the source's revenue-allocation logic — directing proceeds to displacement-conditioned transfers or retraining — is, in our frame, the choice between shifting the mapping and investing in the state (Section 4.2), and the self-limiting trajectory they conjecture is the Corollary's R-wins regime. The second is Tinbergen's principle, which their Section 6 invokes: a distinct market failure requires a distinct instrument. Our extension adds a second distinct failure — the depletion externality — and the same principle then requires the second term. $\Lambda(t)$ is not an embellishment on the tax; it is the instrument the second failure requires.

5.2 The Hard Problem, Formalized

The political economy of this instrument has a familiar shape: the side that would pay contests the measurement. It is worth being exact about what can and cannot be contested, because the

Proposition of Section 3.4 draws the line in a surprising place.

Today's measurement is not the vulnerable object. The income-replacement rate is observable with a lag — the displacement literature measures displaced cohorts' earnings losses over multi-year horizons — and the source itself notes that setting the rate requires only sector-level observables. A lobby that disputed the current value of η would be disputing administrative data. What is not measurable from aggregate series is the mapping: the slope of $\eta(\bar{\alpha}, R)$, the location of the frontier, the counterfactual replacement rate under any alternative automation or policy path. And that is exactly the object a forward-looking instrument regime needs. The tax set today is a bet about the path of η; the path is a property of the mapping; the mapping is unidentified from the data the contesting parties can be compelled to produce. One asymmetry is worth naming, because it is the sharpest version of the point: the static rate is computable from today's measured η; the dynamic correction $\Lambda(t)$ integrates the mapping's slope forward — and the mapping is the object aggregate data cannot identify. The measurable instrument is the myopic floor; every dollar of the depletion premium above it rests on what cannot currently be measured.

The political-economy consequence follows with uncomfortable precision. Lobbies do not contest today's measured η. They contest the counterfactual — the claim that their sector's reabsorption will not hold — and the counterfactual is precisely the object aggregate data cannot settle. The contestability of the instrument is not a measurement shortfall that more funding would fix; it is a property of the system, proved rather than asserted. No volume of aggregate time series separates movement along the reabsorption curve from shifts of it.

The same proposition supplies the way out, and it is the bridge to Section 6. What aggregate data cannot identify, cross-sectional variation can: displacement outcomes observed at different points of automation intensity, across sectors and regions, trace out the mapping under a common-curve assumption. A durable instrument regime therefore cannot be organized around a point estimate of η — any such estimate is one reading on an unidentified function, hostage to the next contested counterfactual. It must be organized around the measurement infrastructure that estimates the function itself: the panels, the spending data, the cross-sectoral design. The instrument and the empirical program are not sequential policy steps. They are the same policy, at different maturities.

5.3 Implementation

Three practical problems attend the instrument, and it is worth separating them, because they have different owners.

The first is levying. Setting the rate requires sector-level observables; collecting it requires firm-level ones — an authority must observe automation rates at the level of the individual firm. The source judges this feasible and getting easier: AI adoption generates procurement records, and a tax authority — unlike a rival firm in a Coasean negotiation — can compel disclosure through mandatory reporting, payroll records, and procurement audits. The asymmetry is worth pausing on: the same non-contractibility that kills private bargaining in Section 4.5 is overcome by the state's audit power, which is one reason the correction is public rather than private.

The second is definition. A statute that taxes "automation" must not tax ordinary productivity investment, and the line between them is a legal object, not an economic one. The model offers the drafting principle: the externality attaches to the displacement of wage income that would have been spent in the sector, so the taxable event is task substitution with displacement, and the base should track displaced earnings rather than technology adoption as such. Getting this wrong in either direction — taxing productivity, or exempting displacement — reintroduces at the statutory level exactly the wedge the instrument exists to close.

The third is leakage across borders. The model is a closed-sector game; a unilateral tax can push adoption offshore, and the source's conclusion — multilateral coordination or border-adjustment mechanisms analogous to carbon policy — stands, with one amendment from the dynamics. Carbon leakage relocates emissions; automation leakage relocates the demand destruction *and* the depletion of reabsorption capacity, which is harder to repatriate than a factory. The case for coordination is stronger here than in the climate analogue the source reaches for.

5.4 The Reframe

The standard architecture of Pigouvian policy is: measure the externality, set the price, collect. This paper inverts the order. The measurement is not a preliminary to the policy; it is the policy — the only component that determines whether the instrument can survive the contest it will provoke. A tax set on a point estimate of an unidentified mapping will be litigated into paralysis by parties who cannot be proved wrong with the data at hand. A regime built on the mapping itself — estimated, monitored, and revised as the economy moves along it — converts the instrument from a number into an institution.

The source closes by naming two dynamic forces that pull in opposite directions — the irreversibility of automation investments, strengthening the case for early action, and the tendency of the income-replacement rate to rise as workers retrain and new occupations emerge, shrinking the optimal tax over time — and declines to model them. Modeling them is what this paper does. The forces become state variables, the conjecture becomes a regime of the Corollary, and the instrument becomes a path with a sign that cannot currently be read. The measurement problem is the policy problem. That used to be a slogan about this literature. It is now a theorem.

6. The Empirical Program: Estimating η as a Function, Not a Scalar

6.1 What the Externality Would Look Like

Begin where the source begins: with the firm side of the ledger. Their proposed signature is profit erosion — standard competitive models predict that cost-reducing technology raises profits, so profit erosion coinciding with mass layoffs would be difficult to rationalize without the demand externality (their Proposition 2). We adopt the signature and add the caveat that disciplines this section: it is neither necessary nor sufficient. Not sufficient, because intensified pass-through of a cost shock — Bertrand competition doing exactly what it should — also compresses margins alongside adoption; the firm side of the ledger cannot separate "rivals competing away the gains of cheaper production" from "the consumer base shrinking beneath every firm at once." Not necessary, because the externality is active whenever $N > N^*$, and absolute profits erode only when the saving is small enough relative to the demand loss: at the interior equilibrium, $\pi(\alpha^{NE}) - \Pi_0 = (L\alpha^{NE}/2)\cdot[s - \ell(2 - 1/N)]$, so erosion requires $s < \ell(2 - 1/N)$ — at twenty firms, savings below 1.95 times the per-task demand loss. (The crisper $s < \ell$ boundary holds only in the frictionless limit, where full automation is dominant and profit is $\Pi_0 + L(s-\ell)$.) Above that boundary the equilibrium can be Pareto dominated by the cooperative optimum — every firm's profit below its cooperative profit — while absolute profits still rise relative to the pre-automation baseline, and no erosion appears. And the band the boundary opens is the signature's sharpest territory: for $\ell < s < \ell(2 - 1/N)$, cooperative automation would be strictly profitable, yet competitive over-automation drives absolute profits *below* the pre-automation baseline — erosion appearing precisely where restraint would have raised profits. Both stories predict falling prices and squeezed margins; the signature is informative — its presence is strong evidence, its absence none — and the discriminating evidence must come from the demand side.

6.2 The Consumer-Side Signature

A caveat first, because it governs what counts as evidence. The source's baseline holds output per task fixed — an automated task and a human-performed task each produce one unit — so the model has no quantity margin at all. Any identification table's "quantities" row is meaningful only under their Section 5.1 relaxation, in which AI is $\varphi > 1$ times as productive. The paper says this explicitly rather than letting the table imply otherwise.

Under the baseline, identification runs through spending levels and income flows: aggregate expenditure D and the spending behavior of displaced cohorts. Under the relaxation, a subtlety appears that cuts against naive quantity tests: $\varphi > 1$ gives automating firms a market-share motive, and both rival explanations — healthy competition and the externality — predict quantity *reallocating toward* automating firms. Reallocation is therefore not discriminating. The discriminating object is the aggregate: competition predicts total quantities and total spending rise as prices fall; the externality predicts they stagnate or fall as the consumer base hollows out from under the price cuts.

	BERTRAND PASS-THROUGH	DEMAND EXTERNA
Prices	Fall	Fall
Aggregate quantities *(requires the φ > 1 relaxation)*	Rise	Fall or sta
Aggregate spending D	Rises *(requires elastic sectoral demand)*	Falls
Cross-sector spending of displaced cohorts	Neutral-to-positive	Falls

The third row discriminates only where sectoral demand is elastic: under inelastic demand even healthy pass-through reduces sectoral spending, and the row falls silent. The last row is the workhorse — the sharpest and the most testable with current data. Pass-through has no channel by which a displaced worker spends less on everything else; the externality has exactly that channel, and it is the channel this paper is about.

6.3 Estimating the Mapping

The Proposition of Section 3.4 does the design work: it tells us not merely that aggregate data fail, but exactly what variation succeeds. The mapping $\eta(\bar{\alpha}, R)$ is identified by displacement outcomes observed at different points of automation intensity — cross-sectional variation across sectors and regions — under a common-curve assumption: that $\eta(\bar{\alpha}, R)$ is an economy-wide mapping rather than a collection of sector-specific ones. That assumption is doing real work and must be treated accordingly: stated, defended, and robustness-checked against sector-specific mappings wherever the data permit.

The defense has two parts. First, the mapping's arguments are economy-wide states, and the displacement literature's central stylized fact — large, persistent earnings losses replicating across very different sectors and episodes — is itself evidence that reabsorption outcomes share structure: occupational mix, licensing, and local market thickness shift the curve's position, but the dependence on the two states is the common object, and the shifters are observable and conditionable. Second, the design degrades gracefully where the assumption fails: within-sector variation across regions — common occupational structure, differing automation intensity and local capacity — identifies sector-specific mappings directly, so the cross-sector design is the first instrument, not the only one. The paper's own two-sector extension (Appendix C) maintains the common mapping: the sectors differ in exposure, not in the curve.

The design follows. Displaced-cohort panels — reemployment duration and wage-on-reemployment, the Jacobson et al. (1993) apparatus — linked to sectoral automation intensity; AI-capability shocks as instruments for the intensity a sector would otherwise choose — restricted to shocks that displace without creating new tasks, since capability that creates tasks moves R as well and fails exclusion; cross-sector and cross-time variation to trace the mapping rather than to re-estimate its level; scanner and credit-card data for the spending side of Table 6.1. The identification is two-dimensional, and the design must be read accordingly: the cross-section at a date varies automation intensity against an approximately common reinstatement stock, identifying the mapping's $\bar{\alpha}$-slope; the panel's time dimension moves R as capacity depreciates, rebuilds, and is invested in, identifying the R-dimension. A cross-section alone recovers a slice; the panel recovers the surface. The deliverable is not a point estimate of η. It is an estimate of the mapping itself and a confidence region for the frontier $\bar{\alpha}^*(R)$ — the empirical object that Section 5's durable instrument regime requires, and the only object that can settle the counterfactuals the political process will contest.

6.4 The Skeptical Evidence

The strongest argument against the externality is that we have not seen it. The visible record of AI-era layoffs shows corrections and reversals rather than demand collapse: findings that attribute recent tech layoffs to overhiring correction rather than displacement, analyst readings of high-profile cuts — Block's reduction of roughly forty percent of its workforce among them — as firm-specific restructuring, and the partial rehiring that followed Klarna's much-cited automation of customer service. This evidence deserves engagement, not deflection, and the engagement has three steps.

First, all of it is consistent with a latent failure mode. The source's own framing, quoted directly: the signature "requires displacement at a scale and speed beyond what has materialized so far," and the contribution is "identifying a structural vulnerability rather than diagnosing an active crisis." Second, under endogenous reabsorption the current calm is not merely consistent with the model — it is what the model predicts on the near side of the frontier. The wedge is continuous through $\eta = 1$: the frontier is a sign change, not a cliff, so near it the wedge is small and the signature faint by construction. The dynamics sharpen the point in two directions. Left to itself, the economy self-arrests (Section 3.4's remark) — but nothing places the arrest on the near side: on the subsidy side the automation target runs ahead of s/k and can leap the frontier in a single step, so even a trendless economy can wake up on the wrong side. What a trendless economy cannot do is keep deepening: sustained movement through and past the frontier requires a moving target — falling AI costs or rising capability — and with that trend in place the signature does not jump; it steepens, growing as the economy moves along the mapping precisely because measurement is hardest where the wedge is smallest. Today's ambiguity is not the absence of evidence; it is what evidence looks like on the near side of a frontier the trend is pushing toward. Third, the source's conclusion goes further than this paper needs: past displacement episodes have consistently produced persistent earnings losses, "placing most economies firmly in the over-automation regime." We claim less and it suffices: the frontier exists, advanced economies may already sit on its over-automation side, and which side they sit on is measurable in principle — by the program of Section 6.3, and by nothing cheaper.

The skeptical evidence is thus an argument for the empirical program, not for waiting. The cost of discovering the frontier by crossing it is Theorem 2's surplus, forgone and unrefunded.

7. Related Work

7.1 Task-Based Automation

The framework within which all of this sits is the task-based model of Acemoglu and Restrepo (2018), and the source's endogenous-wage extension (their Section 5.3) is that framework's wage-adjustment feedback transplanted into the demand-externality setting. The channel closest to our state variable is the reinstatement effect of Acemoglu and Restrepo (2019): new tasks and occupations in which labor holds a comparative advantage rebuild the demand for displaced workers. The source invokes reinstatement as a conjecture — the income-replacement rate rises over time as displaced workers retrain and new occupations emerge — and thereby treats the race between reinstatement and displacement as something that happens to the model from outside. Our contribution to this lineage is to put the race inside it: R is the reinstatement channel given an explicit state, $\bar{\alpha}$ is the displacement channel given one, and $\eta(\bar{\alpha}, R)$ is the outcome of the two running against each other. Where the lineage asks whether reinstatement will arrive in time, the model here asks what each runner does to the other's track.

7.2 Demand Spillovers and Coordination

The demand-spillover pedigree — Rosenstein-Rodan (1943), Murphy, Shleifer, and Vishny (1989), Cooper and John (1988) — is the source paper's own related-work framing, and we adopt it as such, citations and all; the cheap-talk observation of Section 2.3 is likewise theirs, though the Crawford and Sobel (1982) citation is ours, the source's current version having dropped it. What the big-push lineage and the source share, for all their differences, is a static architecture: multiplicity of equilibria in the former, a parameter-dependent wedge in the latter. What endogenous reabsorption adds is history. With $\eta = \eta(\bar{\alpha}, R)$ and automation sunk, where the economy sits depends on the path it took; the instrument's sign can flip along a single trajectory; and excursions past the frontier leave scars that recovery does not refund. Tipping and hysteresis are not refinements of the spillover story. They are what the spillover story becomes when the spillover feeds back onto the parameters that generate it.

7.3 The Closest Neighbor

The nearest paper is Beraja and Zorzi (2025), who show that automation is inefficient when displaced workers face borrowing constraints during reallocation. The source's own distinction stands, and we adopt it because it survives endogenization unchanged: their friction operates through the labor market — firms ignore the welfare cost imposed on credit-constrained workers — while this one operates through the product market, with firms ignoring the demand they destroy for rivals; their inefficiency arises even for a single firm in isolation, while this one requires competition and vanishes under monopoly; and their planner corrects automation to protect worker welfare, while this paper's planner would restrain automation with zero weight on workers, because over-automation harms firm profits themselves.

Our addition to the comparison is that endogenous reabsorption makes the two channels potentially compounding rather than merely adjacent. Borrowing constraints bind hardest precisely when replacement has failed — when η has fallen — and in our frame a low η is not a static misfortune but a state the automation process itself deepens. The labor-market friction and the product-market externality then feed the same state variable from different sides: depletion tightens the constraints that make displacement costly, and displacement depletes the capacity that would relax them. A formal treatment of the compound case is left to future work; the architecture here — two states, one mapping — is built to receive it.

The nearest concurrent work is Bayraktar (2026), which builds on the same source and embeds the demand externality in a stationary heterogeneous-agent economy — incomplete markets, equity ownership of automation rents, market-clearing factor prices, and explicit fiscal closure — computing the regimes in which decentralized automation lowers stationary consumption and capital. The division of labor is exact, and stated in their text: the analysis is stationary by design, with transition-path evaluation outside the model solved. This paper's objects are precisely the transition — the state ($\bar{\alpha}$, R), the frontier, the hysteresis, the instrument's time path — and its conditioning variable is the reabsorption state rather than the wealth distribution. The two compose: theirs quantifies the externality's incidence across households in a stationary economy; ours prices its dynamics through the transition.

7.4 Corrective Taxation Traditions

Robot taxes have been justified by transitional frictions (Guerreiro et al., 2022) and distributional concerns (Costinot and Werning, 2023) — both noted in the source's related work, and both genuinely different from what is taxed here. Our wedge decomposition is where the differentiation lives: the distributional premium $\mu\ell/(\lambda(1-\mu)k)$ is the component that traditions of distributional taxation address, and the externality term $\ell(1-1/N)/k$ is the component that survives complete credit markets, costless transitions, and a planner who weights workers at zero. Guerreiro et al.'s finding that the optimal robot tax declines to zero as displaced cohorts retire shares the shape of our R-wins regime in spirit — their tax relaxes through demography as cohorts retire, ours through reabsorption as capacity rebuilds; the mechanisms are distinct and we do not conflate them — and our Theorem 3 adds the qualification their transitional framing omits: while automation continues to crowd given reabsorption capacity, the optimal instrument carries the depletion shadow price $\Lambda(t)$, and the myopic rate is a floor. The tradition prices the transition's victims. The instrument here prices the transition's shadow on the market itself.

8. Conclusion: Structural Vulnerability, Endogenous Frontier

Competitive automation creates a demand externality: the firm that replaces a worker with AI captures the full cost saving and bears only 1/N of the demand the displacement destroys. That is the source paper's result, and it is conditional — the distortion's sign depends on the income-replacement rate, reversing at full replacement. This paper began from the observation that the conditioning variable is not a number the economy is assigned but an outcome the transition produces, and it ends with the consequences. Absorption capacity is itself depleted by automation, so the replacement rate declines as automation spreads even as reinstatement capacity rebuilds. The economy therefore possesses a tipping frontier at full replacement, and its position relative to that frontier is endogenous to the automation process itself. Because automation is sunk, excursions past the frontier exhibit hysteresis: the surplus forgone along the way is not refunded when reabsorption recovers. Because the instrument must look forward, its correct level carries a term no static analysis prices — the shadow value of the reabsorption capacity that today's automation crowds. And because aggregate data cannot separate movement along the reabsorption mapping from shifts of it, everything the instrument needs to know — the frontier's location, the mapping's slope, the counterfactuals the political process will contest — is unidentified in the series we currently collect. Today's replacement rate is measurable, and with it today's sign. Tomorrow's is a property of a mapping we have not yet built the instruments to see.

The policy agenda follows in the order the theorems impose. Measurement infrastructure first: the displaced-cohort panels, the spending data, and the cross-sectional design that estimate the mapping rather than the level — the only component of the regime that can settle contested counterfactuals. A state-contingent instrument second: the Pigouvian rate, corrected for depletion, financed through the transfers and retraining that shift the mapping and rebuild the state — the two halves of the dynamic answer working as one policy.

International coordination throughout, because automation leakage carries off demand destruction and reabsorption capacity together, and neither repatriates easily.

The invisible hand coordinates production with a precision no planner has matched. It was never given the instruments to coordinate automation, because the demand that displaced workers would have spent has no price, no venue, no counterparty. What this paper adds to that diagnosis is that the blind spot is not static. Every task automated without the correction deepens it — depleting the capacity that would have absorbed the next displacement, moving the frontier, and doing so in a way the market's own signals cannot register and aggregate statistics cannot yet locate. The hand is myopic, and the myopia deepens the faster it runs.

Appendix A. Benchmarks and the Wedge Decomposition

This appendix derives the three benchmarks of Section 3.2 and restates the source's decomposition in the notation of Section 3.1. Everything here is the source's — their Propositions 1 and 2, including their eq. (9) — restated for self-containment. The interior regime ($\ell < s < k + \ell/N$) is maintained throughout, matching the body.

The firm's problem. Firm i chooses α_i to maximize profit. Automating the marginal task saves $s = w - c$ in full; it also destroys ℓ of sectoral demand, of which the firm bears only its revenue share $1/N$; and it costs $k\alpha_i$ at the margin in integration. The private first-order condition is therefore $s = \ell/N + k\alpha_i$. Rivals' automation enters the firm's payoff only through a term independent of α_i, so the equilibrium rate is a strictly dominant strategy — foresight about rivals does not change it (their Proposition 1).

Benchmarks. The symmetric Nash equilibrium and the cooperative benchmark — a planner setting a common rate, who faces the full ℓ rather than ℓ/N — are $\alpha^{NE} = (s - \ell/N)/k$ and

$\alpha^{CO} = (s - \ell)/k$, each pinned to [0, 1] outside the interior regime. The over-automation wedge is their difference, $\ell(1 - 1/N)/k$: strictly increasing in N and ℓ, decreasing in k, and approaching its maximum ℓ/k as $N \to \infty$. At $N = 20$, the running example, the externality term is nineteen-twentieths of ℓ/k. Firms automate at all only when $N > N^* = \ell/s$ (their eq. 7); below the threshold the saving does not cover even the firm's private share of the loss. Two corners carry the source's own caution: for $s \leq \ell$ the cooperative benchmark is zero and the wedge is α^{NE} itself; for $s \geq k + \ell$ both rates pin at full automation and the measured wedge reads zero — not because the externality has vanished, but because the $\alpha^{NE} - \alpha^{CO}$ comparison can no longer detect it.

The decomposition. A planner placing weight $\mu \in [0, 1)$ on worker income automates at $\alpha^{SP}(\mu) = (s - \ell)/k - \mu\ell/(\lambda(1-\mu)k)$, which reduces to α^{CO} at $\mu = 0$. The total wedge splits in two (their eq. 9): $\alpha^{NE} - \alpha^{SP}(\mu) = \ell(1 - 1/N)/k + \mu\ell/(\lambda(1-\mu)k)$. The first term is the demand externality — present at $\mu = 0$, growing with N toward ℓ/k. The second is the distributional premium — independent of N, growing without bound as $\mu \to 1$, with the planner prohibiting automation entirely at $\bar{\mu} = \lambda k\alpha^{CO}/(\ell + \lambda k\alpha^{CO})$. Section 3.2's scoping — "structural, not distributional" attaches to the first term only — is read directly off this decomposition. The surplus loss is quadratic in the total wedge and scales with NL, which is what licenses the source's remark that even an imprecisely targeted tax yields a first-order gain — the remark Section 5.1 qualifies dynamically.

Appendix C. The Two-Sector Case: Cross-Sector Depletion

Two sectors, $i \in \{1, 2\}$, each with the structure of Section 3.1: sector i has automation rate $\bar{\alpha}_i$ and replacement rate η_i, decreasing in automation and increasing in reinstatement capacity R. The single difference from the one-sector model is the reach of the depletion primitive.

Displaced workers from either sector draw on a shared reabsorption market, so sector i's effective exposure is to the economy's automation, not only its own: $\eta_i = \eta((\bar{\alpha}_i + \kappa\bar{\alpha}_j)/(1 + \kappa), R)$, where $\kappa > -1$ is sector i's net spillover exposure to sector j's automation — the sharedness of the reabsorption pool, net of the cross-sector income channel discussed below. The maintained assumption of this appendix is $\kappa \in (0, 1]$: positive net exposure, bounded above by a fully common labor market. Under the maintained assumption, the normalized argument is a convex combination of the two sectors' automation rates — an effective economy-wide automation rate, lying in [0, 1], so η is evaluated nowhere outside its Section 3.4 domain. $\kappa \to 0$ recovers the one-sector model; $\kappa = 1$ weights the two sectors' automation equally — the fully common market with no offsetting income effect. R remains an independent state, exactly as in Section 3.4: its law of motion involves no automation variable, the property Theorem 2's welfare comparison leans on. Depletion enters, as it does throughout the paper, through η's first argument. This is the formal content of the sentence the one-sector model could not state: the sectors that would reabsorb displaced workers are themselves automating.

Lemma (Cross-sector depletion). For $j \neq i$, $\partial\eta_i/\partial\bar{\alpha}_j = [\kappa/(1 + \kappa)] \cdot \eta_{\bar{\alpha}}((\bar{\alpha}_i + \kappa\bar{\alpha}_j)/(1 + \kappa), R) < 0$: automation in sector j lowers the effective replacement rate in sector i, holding sector i's own automation fixed.

Proof. Chain rule: $\eta_{\bar{\alpha}} < 0$ by the depletion primitive of Section 3.4, and $\kappa/(1 + \kappa) > 0$ by the maintained assumption. ■

The exposure κ is net of the one channel that runs the other way. Automation in sector j lowers prices there; the real-income gain can raise demand for sector i's product and speed reabsorption in i. The lemma holds whenever the depletion channel dominates this cross-sector income channel — equivalently, whenever κ is positive, which is the maintained assumption stated above rather than a property of the definition. That assumption is the two-sector analogue of the maintained regime the source states in its footnote 2. Strong enough complementarities could make $\varkappa \leq 0$; such sector pairs would be absorbing each other's workers *because* of automation, and the knife-edge confirms the rule by naming its exception. Under the maintained assumption, each sector's automation taxes the other's reabsorption, and the wedge, the frontier, and the hysteresis of the one-sector model all survive. Under symmetric automation the two-sector model reproduces the one-sector figure exactly — a consistency property of the shared pool, not a gap in it: the one-sector economy is the symmetric special case. The compounding is cross-sectional. Where automation runs asymmetric, depletion concentrates on the lagging sector, whose effective exposure exceeds its own automation rate — the single-sector reading understates what that sector's displaced workers face. And the marginal fact is unconditional: from any symmetric position, the first sector to automate faster begins depleting the other's reabsorption.

What this appendix does not yet do is solve the two-sector equilibrium: cross-sector comparative statics of the frontier and the instrument path remain in preparation, as the working-draft disclosure notes.

References

Acemoglu, Daron, and Pascual Restrepo. 2018. "The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment." American Economic Review 108(6): 1488–1542.

Acemoglu, Daron, and Pascual Restrepo. 2019. "Automation and New Tasks: How Technology Displaces and Reinstates Labor." Journal of Economic Perspectives 33(2): 3–30.

Baldwin, Richard, and Paul Krugman. 1989. "Persistent Trade Effects of Large Exchange Rate Shocks." Quarterly Journal of Economics 104(4): 635–654.

Bayraktar, Erhan. 2026. "Automation, Income Incidence, and Capital Accumulation in Incomplete Markets." Working paper. arXiv:2605.05127.

Beraja, Martin, and Nathan Zorzi. 2025. "Inefficient Automation." Review of Economic Studies 92(1): 69–96.

Boppart, Timo. 2014. "Structural Change and the Kaldor Facts in a Growth Model with Relative Price Effects and Non-Gorman Preferences." Econometrica 82(6): 2167–2196.

Comin, Diego, Danial Lashkari, and Martí Mestieri. 2021. "Structural Change with Long-Run Income and Price Effects." Econometrica 89(1): 311–374.

Cooper, Russell, and Andrew John. 1988. "Coordinating Coordination Failures in Keynesian Models." Quarterly Journal of Economics 103(3): 441–463.

Costinot, Arnaud, and Iván Werning. 2023. "Robots, Trade, and Luddism: A Sufficient Statistic Approach to Optimal Technology Regulation." Review of Economic Studies 90(5): 2261–2291.

Crawford, Vincent P., and Joel Sobel. 1982. "Strategic Information Transmission." Econometrica 50(6): 1431–1451.

Dixit, Avinash. 1989. "Entry and Exit Decisions under Uncertainty." Journal of Political Economy 97(3): 620–638.

Guerreiro, João, Sergio Rebelo, and Pedro Teles. 2022. "Should Robots Be Taxed?" Review of Economic Studies 89(1): 279–311.

Hemenway Falk, Brett, and Gerry Tsoukalas. 2026. "The AI Layoff Trap." Working paper, University of Pennsylvania and Boston University. Version of June 3, 2026 (first version

March 2, 2026). arXiv:2603.20617, https://arxiv.org/abs/2603.20617.
SSRN: https://ssrn.com/abstract=6448898.

Jacobson, Louis S., Robert J. LaLonde, and Daniel G. Sullivan. 1993. "Earnings Losses of Displaced Workers." American Economic Review 83(4): 685–709.

Matsuyama, Kiminori. 2002. "The Rise of Mass Consumption Societies." Journal of Political Economy 110(5): 1035–1070.

Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny. 1989. "Industrialization and the Big Push." Journal of Political Economy 97(5): 1003–1026.

Rosenstein-Rodan, Paul N. 1943. "Problems of Industrialisation of Eastern and South-Eastern Europe." Economic Journal 53(210/211): 202–211.